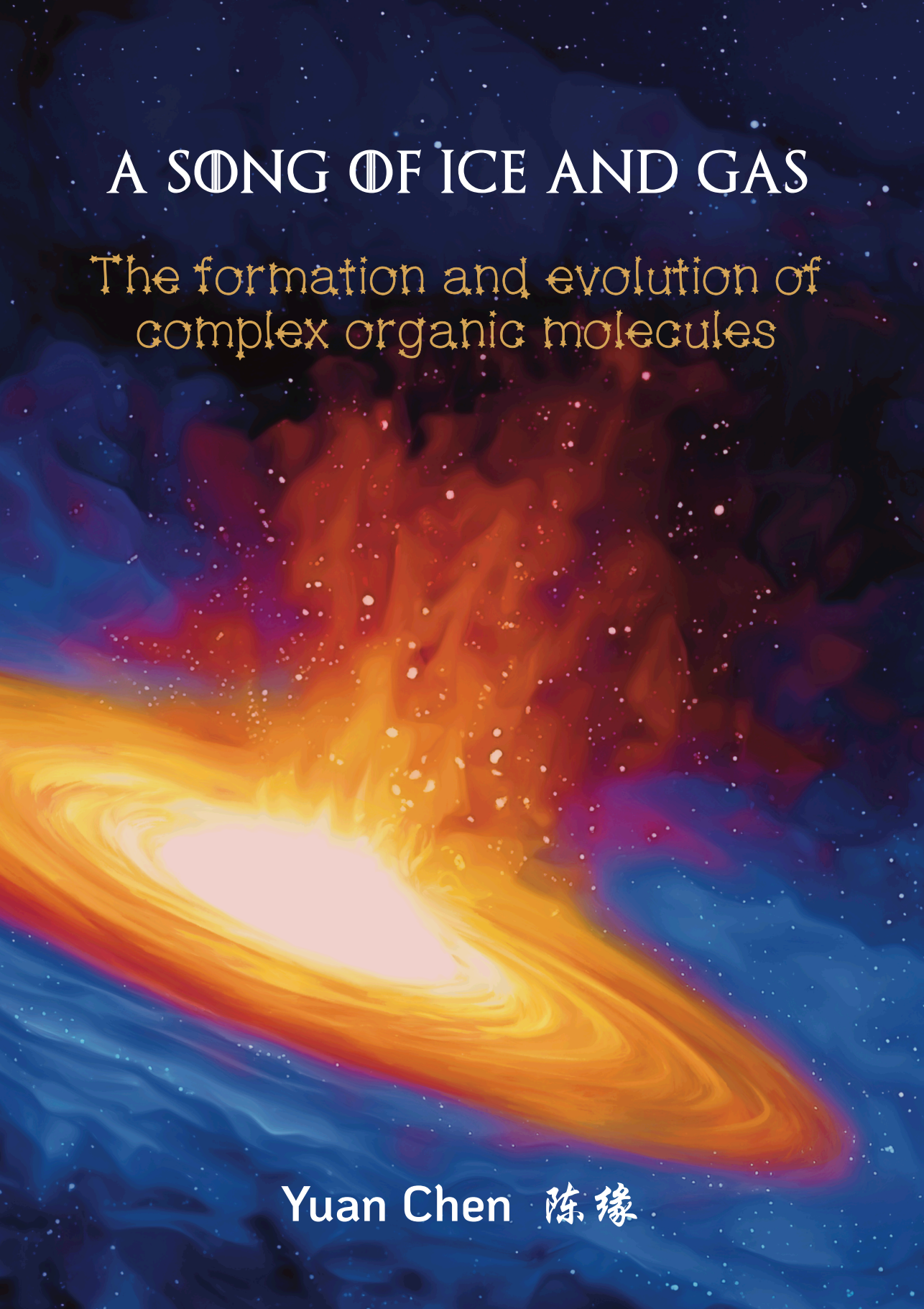
A SONG OF ICE AND GAS

The formation and evolution of complex organic molecules



Yuan Chen 陈缘

Propositions

accompanying the thesis

A Song of Ice and Gas

The formation and evolution of complex organic molecules

1. The physical differentiation between low- and high-mass protostars does not necessarily lead to chemical differentiation of complex organic molecules. (*Chapter 2*)
2. Large complex organic molecules can be abundant if they have certain functional groups. (*Chapter 3*)
3. Both inheritance and reprocessing can be responsible for shaping the chemistry during star formation. (*Chapter 4*)
4. Systematic measurement of peak wavelength and width of absorption bands is a powerful tool for identifying the contributing species and their mixing environments. (*Chapter 5*)
5. The explosive growth of spectral observations of interstellar molecules calls for more powerful fitting tools that can handle line blending and uncertainty estimation automatically.
6. Manual inspection is more reliable and interpretable than machine-learning algorithms when the data volume is limited.
7. Weak spectral features can contain rich science, but are also prone to misinterpretation.
8. Facing the staggering oversubscription rates of ALMA and JWST, it is better to focus on crafting one or two high-quality proposals rather than taking a shotgun approach.
9. It is harder but more fulfilling to take on a challenge than to circumvent it.
10. Good science rests on honesty and patience, while scientific success depends more on ambition and pragmatism.
11. For those accustomed to rich food cultures, a long and exclusive contract with a terrible canteen vendor is the no.1 enemy of academic productivity.
12. Studying abroad, especially in a distinct cultural context, is not just about academics; it is about breaking down and rebuilding one's life and worldview.

Yuan Chen
Leiden, 24 Feb 2026

A Song of Ice and Gas

The formation and evolution of complex organic molecules

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr. S. de Rijcke,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 24 februari 2026
klokke 14:30 uur
door

Yuan Chen
陈缘

geboren te Guangzhou, China
in 1997

Promotor:

Prof. dr. E. F. van Dishoeck

Co-promotor:

Dr. M. L. van Gelder	Netherlands Organisation for Applied Scientific Research (TNO)

Promotiecommissie:

Prof. dr. I. A. G. Snellen	
Prof. dr. S. Viti	
Prof. dr. J. K. Jørgensen	University of Copenhagen, Denmark
Prof. dr. N. Sakai	RIKEN Pioneering Research Institute, Japan
Prof. dr. J.-E. Lee	Seoul National University, South Korea

ISBN: 978-94-6534-206-1

Cover: A protostellar system with distinct colors highlighting the hot core region and the cold envelope. Designed by Yuan Chen using Midjourney (AI-generated draft), Procreate (hand-drawn refinement), and Keynote (text editing and formatting).

时间向前走，一定只有路口没有尽头。
As time moves forward,
there are always crossings ahead and never an end to the road.

—陈奕迅《让我留在你身边》
—Eason Chan, *Let Me Stay by Your Side*

Table of Contents

1 Introduction **1**
- 1.1 Star formation . . . 2
  - 1.1.1 Prestellar phase . . . 2
  - 1.1.2 Low-mass star formation . . . 3
  - 1.1.3 High-mass star formation . . . 4
- 1.2 Chemistry in early stages of star formation . . . 5
  - 1.2.1 Cold prestellar and collapsing phase . . . 5
  - 1.2.2 Warm-up phase . . . 8
  - 1.2.3 Hot core phase . . . 8
- 1.3 Complex organic molecules . . . 9
  - 1.3.1 Categories . . . 9
  - 1.3.2 Formation theories . . . 9
  - 1.3.3 Observations of gas-phase COMs . . . 12
    - 1.3.3.1 Rotational spectroscopy . . . 12
    - 1.3.3.2 Millimeter facilities . . . 14
    - 1.3.3.3 Line surveys . . . 16
  - 1.3.4 Observations of solid-phase COMs . . . 18
    - 1.3.4.1 Vibrational bands of interstellar ices . . . 18
    - 1.3.4.2 Laboratory IR spectroscopy . . . 19
    - 1.3.4.3 Observations in the pre-JWST era . . . 21
    - 1.3.4.4 Findings and challenges of JWST observations . . . 22
- 1.4 This thesis . . . 24
  - 1.4.1 Overview of chapters . . . 24
  - 1.4.2 Outlook . . . 26

2 CoCCoA: Complex Chemistry in hot Cores with ALMA Selected oxygen-bearing species **29**
- 2.1 Introduction . . . 31
- 2.2 Observations . . . 33
  - 2.2.1 Overview . . . 33
  - 2.2.2 Data reduction . . . 34
- 2.3 Methods . . . 36
  - 2.3.1 Spectral analyses . . . 36
  - 2.3.2 Isotope ratio calibration for methanol . . . 40
- 2.4 Results . . . 41
  - 2.4.1 Morphology . . . 41
  - 2.4.2 Spectra and detection . . . 41
  - 2.4.3 Column density and excitation temperature . . . 42
- 2.5 Discussion . . . 43
  - 2.5.1 Column density ratios . . . 43
  - 2.5.2 Observations versus simulations . . . 46

2.5.2.1 CH_3OCH_3 (DME) and CH_3OCHO (MF) 47
2.5.2.2 C_2H_5OH 48
2.5.2.3 CH_3CHO 49
2.5.2.4 CH_2OHCHO (GA) and $(CH_2OH)_2$ (EG) 50
2.5.3 The influence of energetic processes 51
2.5.4 O-COMs from clouds to comets 52
2.5.5 JWST ice observations 53
2.6 Conclusions 54
Appendices 56
2.A Additional figures 56
2.B Additional tables 63

3 CoCCoA: Complex Chemistry in hot Cores with ALMA The chemical evolution of acetone from ice to gas 79
3.1 Introduction 81
3.2 Observations 82
3.3 Methods 83
3.3.1 Acetone (CH_3COCH_3) 83
3.3.2 Ketene (CH_2CO) 86
3.3.3 Propyne (CH_3CCH) 87
3.3.4 Propanal (C_2H_5CHO) and other three-carbon O-COMs 88
3.3.5 Methanol (CH_3OH) 89
3.4 Results 89
3.4.1 Emission morphology 89
3.4.2 Physical properties 91
3.5 Discussion 92
3.5.1 Acetone and other O-COMs 92
3.5.2 Acetone: Ice versus gas 94
3.5.3 Chemical evolution of acetone 96
3.5.3.1 Insights from astrochemical simulations 96
3.5.3.2 Insights from computational chemistry 99
3.5.3.3 Insights from experiments 101
3.5.4 Correlations among O-COMs and acetone relevant species . . . 101
3.6 Conclusions 103
Appendices 105
3.A Supplementary figures for Section 3.5 105
3.B Best-fit physical parameters of molecules involved in this work in addition to acetone 106
3.C Transitions covered in the CoCCoA dataset 108
3.D Figures for spectral fitting results 111

4 JOYS+: The link between the ice and gas of complex organic molecules Comparing JWST and ALMA data of two low-mass protostars 125
4.1 Introduction 127
4.2 Observations and data reduction 130
4.2.1 ALMA 130
4.2.2 JWST 131
4.3 Methods 132

4.3.1 ALMA . . . 132
4.3.2 JWST . . . 133
4.3.2.1 Global continuum . . . 133
4.3.2.2 Silicate bands . . . 135
4.3.2.3 Local continuum . . . 135
4.3.2.4 Gas-phase lines . . . 136
4.3.2.5 Absorption features of complex organic molecule ices 136
4.3.2.6 Ice bands of CH_3OH and H_2O . . . 140
4.3.2.7 Ice column density . . . 140
4.4 Results . . . 141
4.4.1 ALMA . . . 141
4.4.1.1 Emission maps . . . 141
4.4.1.2 Spectra . . . 141
4.4.2 JWST . . . 144
4.4.2.1 Constituents of ice mixtures . . . 144
4.4.2.2 Temperature of ice mixtures . . . 147
4.4.2.3 Decomposition of the B1-c spectrum in 6.8–8.8 µm . . 148
4.4.2.4 Fitting statistics . . . 150
4.4.2.5 CH_3OH and H_2O bands . . . 152
4.5 Discussion . . . 155
4.5.1 Column density ratios: Ice versus gas . . . 155
4.5.2 Origin of the similarity in complex organic molecule ratios between ice and gas . . . 155
4.5.3 Origin of the difference in complex organic molecule ratios between ice and gas . . . 156
4.5.3.1 Gas-phase reprocessing . . . 156
4.5.3.2 Spatial distribution . . . 157
4.5.4 Ice ratios in protostars and comets . . . 160
4.6 Conclusions . . . 161
Appendices . . . 163
4.A Images of ALMA and JWST observations . . . 163
4.B The silicate features . . . 166
4.C Selection of ice mixtures . . . 166
4.D Baseline correction of lab spectra . . . 168
4.E Fitting results of the ALMA spectra of IRAS 2A and B1-c . . . 169
4.F H_2CO bands in the CH_3OCHO mixture . . . 172
4.G Temperature analysis of COM ices . . . 174
4.H Band strengths . . . 174
4.I Statistics of fitting the JWST/MIRI-MRS spectrum of B1-c . . . 174
4.J The CH_3OH band at 3.9 µm . . . 177
4.K Additional figures for the JWST fitting results . . . 178
4.L Additional table . . . 181

5 JOYS+: An overview of the ice features in the COM fingerprint range (6.8–8.8 µm) 191
5.1 Introduction . . . 194
5.2 Observations . . . 195

5.3 Methods . . . 196
5.3.1 Removal of gas-phase lines . . . 197
5.3.2 Interactive fitting of global continuum, silicates, and H_2O . . . 200
5.3.3 Interactive fitting of local continuum . . . 201
5.3.4 Determination of band properties . . . 202
5.4 Results and discussion . . . 203
5.4.1 Silicates and simple ices . . . 203
5.4.2 COM ice fingerprint range (6.8–8.8 µm) . . . 207
5.4.2.1 Detection of ice bands . . . 207
5.4.2.2 Band properties I: area . . . 208
5.4.2.3 Band properties II: peak and width . . . 209
5.5 Conclusions . . . 211
Appendices . . . 213
5.A Interactive fitting tool . . . 213
5.B Additional figures . . . 214
5.C Table of band properties in observations . . . 214
5.D Uncertainty estimation of band properties . . . 222
5.D.1 Observed spectra . . . 222
5.D.2 Laboratory spectra . . . 222
5.E Laboratory spectra of $HCOO^-$ ices . . . 223
5.E.1 Experiments . . . 223
5.E.2 Baseline correction and band property measurements . . . 224

Bibliography **227**

Nederlandse samenvatting **239**

English summary **245**

Publications **251**

Curriculum Vitae **255**

Acknowledgements **257**

1 Introduction

The chemical world on Earth is astonishingly diverse. From water, air, and minerals to organic compounds, from biotic molecules such as proteins and nucleic acids to human-made polymers and medicines, the variety of substances is virtually limitless. This richness of chemistry forms the basis of the natural environment on Earth, nurtures and sustains life, and has been further amplified by human ingenuity. In contrast, when we turn our eyes to the vast universe, the picture seems remarkably simple at first glance. The majority of baryonic matter is composed of only two elements, hydrogen and helium, synthesized in the first few minutes after the Big Bang (Peebles 1993). All heavier elements, grouped together by astronomers under the umbrella term "metals", are produced much later in the interiors of stars and dispersed through stellar winds and supernovae (Burbidge et al. 1957), which only constitute a small fraction of cosmic material. From this perspective, one might think the universe is a chemically barren space.

However, this impression exists only because we have not lifted the veil on the true complexity of interstellar chemistry. Over the past half century, advances in molecular spectroscopy and observational techniques have revealed that interstellar chemistry is far more diverse than initially expected. As of Nov 2025, around 340 molecules have been detected in the interstellar medium (ISM) and circumstellar environments[1]. These species range from simple diatomics to complex organics composed of ten or more atoms, including but not limited to alcohols, aldehydes, acids, ethers, esters, ketones, and even aromatic compounds.

An interesting aspect of these discoveries is their strong connection with star formation. The majority of known interstellar molecules have been identified in the dense, shielded environments of molecular clouds and in the warm, chemically rich zones surrounding protostars — the so-called **hot cores** for high-mass protostars and **hot corinos** for low-mass ones (see reviews by Herbst & van Dishoeck 2009; Jørgensen et al. 2020; Ceccarelli et al. 2023); here we collectively refer to both as "hot cores". These regions act as natural laboratories where low temperatures and high densities combine to drive unique chemical processes that lead to the formation of a variety of interstellar ices on dust grains (Sects. 1.2.1–1.2.2) followed by their subsequent release into the gas phase (Sect. 1.2.3). The recognition that the interstellar medium is not merely a dilute mixture of atoms and simple molecules but rather a chemically complex environment has transformed our understanding of cosmic evolution, linking the physics of star formation with the origins of molecular complexity and, ultimately, the ingredients necessary for life.

To deepen our understanding of how chemical complexity arises and evolves in the Universe, this thesis makes use of the most powerful astronomical facilities to date, the

[1] https://cdms.astro.uni-koeln.de/classic/molecules

Atacama Large Millimeter/submillimeter Array (ALMA; Sect. 1.3.3.2) and the *James Webb* Space Telescope (JWST; Sect. 1.3.4.4), to study **complex organic molecules (COMs)**[2] in protostellar systems. Particular focus is placed on a subset of COMs that are sufficiently abundant to be routinely detected in space (see Table 1.1). These observations provide critical insights into the processes that shape the chemistry of the interstellar medium and its evolution through the protostellar and protoplanetary stages, and potentially to the siblings of our own Solar System.

1.1 Star formation

Star formation is a process in which cold molecular gas assembles and collapses under self-gravity to form protostars, disks, and eventually planetary systems. In addition to self-gravity, other factors including turbulence, magnetic fields, thermodynamics (heating and cooling), angular momentum transport, and stellar feedback (e.g., winds, outflows, and radiation) also regulate the dynamical processes of star formation. Relevant studies are often conducted under two regimes: low-mass and high-mass, with a dividing line[3] of $\sim$8 $M_\odot$. However, the recent review by Beuther et al. (2025) has shown multiple similarities in the properties of low- and high-mass star-forming regions, which weaken the bimodality of star formation to some extent.

1.1.1 Prestellar phase

Molecular clouds are the cradles of star formation. Cloud complexes span a few to hundreds of parsecs (pc) in size and typically have total masses of 10^3–10^5 $M_\odot$ (e.g., Kennicutt & Evans 2012). Zooming from cloud scales to sub-parsec scales, the seeds of star formation are embedded in the coldest (10–20 K) and densest (10^4–10^6 cm^{-3}) regions of molecular clouds called prestellar or starless cores (e.g., Bergin & Tafalla 2007). In the low-mass regime, these cores are smaller (0.01–0.1 pc) and less massive ($\sim$1 $M_\odot$) than their high-mass analogs (0.1–1 pc, 10^2–10^3 $M_\odot$), which are often embedded in infrared dark clouds (IRDCs) and difficult to be identified due to their short lifetime, large distances, and crowded environments (e.g., Motte et al. 2018).

Prestellar cores have a centrally concentrated density profile and will meet the fate of gravitational collapse once the Jeans criterion (Jeans 1902) is met, that is, the density and mass are high enough for self-gravity to overcome the resistance of internal gas pressure and other support from turbulence and magnetic fields. During the collapse of a prestellar core, its inner regions contract faster than the outer layers, because the free-fall time $t_{\rm ff} \propto \rho^{-1/2}$ is shorter where the density is higher. As a result, the center quickly becomes opaque to its own cooling radiation, raising the internal pressure and forming a small hydrostatic object called the first hydrostatic core (FHSC; Larson 1969; Masunaga et al. 1998). As material continues to rain in, the

[2] COMs, or interstellar COMs (iCOMs), are initially defined by Herbst & van Dishoeck (2009) as molecules containing six or more atoms (i.e., complex), one of which is carbon (i.e., organic). In practice, however, some five-atom molecules such as formic acid (HCOOH) and ketene (CH_2CO) are also considered part of the broader class of COMs because of their strong chemical connections to larger organic molecules (e.g., Sect. 1.3.1).

[3] There is also a more stringent dividing line of 2 $M_\odot$ for low-mass stars, and those with masses between 2 and 8 $M_\odot$ are classified as intermediate-mass stars, also known as Herbig Ae/Be stars. Here we adopt a loosely defined dividing line of 8 $M_\odot$ between low- and high-mass stars.

central density and temperature rise to a point ($\rho_c \sim 10^{-7}\,\mathrm{g\,cm^{-3}}$ and $T_c \sim 2000\,\mathrm{K}$) where hydrogen molecules begin to dissociate and cool down the core, triggering a second, rapid collapse that leads to the formation of a protostar (Masunaga & Inutsuka 2000). At the same time, angular momentum naturally redirects some infalling gas into a rotating disk while bipolar outflows remove excess angular momentum and clear envelope cavities, marking the transformation from a cold prestellar core to an actively accreting protostellar system.

Given the differences in evolutionary timescales and physical properties (e.g., radiation intensity), the following protostellar and pre-main-sequence phases of star formation will be introduced separately for the low- and high-mass regimes.

1.1.2 Low-mass star formation

The processes of low-mass star formation after the prestellar phase are well described by the observational **Class** system and the physical **Stage** framework. Back to the 1970s, analytic models such as "inside-out collapse" of a singular isothermal sphere (Shu 1977) provided a canonical picture of how a molecular cloud core evolves into a protostar, which was further synthesized into a sequence of phases including the collapse of dense cores, the growth of protostars via accretion, the formation of circumstellar disks, and the eventual dispersal of envelopes (Shu et al. 1987). These theories were not tied to concrete observational diagnostics until the launch of the Infrared Astronomical Satellite (IRAS) in 1983, which provided unprecedented amounts of infrared (IR) photometry data of young stellar objects (YSOs) that are hardly visible at optical wavelengths. An observational classification system was then built by Lada (1987) upon these IR surveys using the slope of the spectral energy distribution (SED) between 2 and 25 μm:

$$\alpha_{\mathrm{IR}} = \frac{\mathrm{d}\log(\lambda F_\lambda)}{\mathrm{d}\log\lambda}. \tag{1.1}$$

In general, the IR sources were divided into three classes: Class I objects have rising SEDs longward of 2 μm that are apparently broader than blackbodies; Class II objects have flatter or declining SEDs with a still prominent IR excess; Class III objects have nearly photospheric SEDs with little IR excess. The IR excess in Class I, II, and III objects was suggested to arise from the infalling envelopes, the accreting circumstellar disks, and the leftover grains of the dispersed disks, respectively.

Soon afterward, submillimeter (sub-mm) continuum observations revealed sources with strong long-wavelength emission but essentially no near-IR counterpart, and a new "Class 0" category was introduced to describe these deeply embedded protostars (André et al. 1993). Class 0 objects are characterized by their large contribution of sub-mm luminosity ($L_{\mathrm{submm}}/L_{\mathrm{bol}} > 0.5\%$), implying that the bulk of the system mass still resides in a cold envelope rather than in the central protostar.

In the mid-1990s, the Class system was further refined by Greene et al. (1994), who carefully recalibrated the boundaries in spectral slope and introduced the "flat-spectrum" category to bridge the gap between Class I and Class II sources. They also codified the ranges of the IR spectral index α_{IR} for different classes: Class I ($\alpha_{\mathrm{IR}} > 0.3$), flat-spectrum ($-0.3 \leq \alpha_{\mathrm{IR}} \leq 0.3$), Class II ($-1.6 \leq \alpha_{\mathrm{IR}} \leq -0.3$), and Class III ($\alpha_{\mathrm{IR}} < -1.6$). Meanwhile, Chen et al. (1995) proposed to use bolometric

temperature ($T_{\rm bol}$) as another diagnostic of evolutionary stages, where $T_{\rm bol}$ was defined as the effective temperature of a blackbody with the same mean frequency as the observed SED. Their thresholds, <70 K for Class 0, 70–650 K for Class I, and 650–2800 K for Class II, became widely adopted in statistical studies such as the large surveys with the Infrared Space Observatory (ISO) and *Spitzer* Space Telescope (e.g., Evans et al. 2009; Enoch et al. 2009).

Although this empirical scheme offered a practical way to organize the newly accessible population of YSOs and mapped neatly onto the theoretical picture of evolving protostellar systems, a critical limitation was revealed that the observational classification did not always reflect the true physical state of a source. The apparent SED of a protostellar system is sensitive to its inclination, extinction, and outflow cavity geometry (e.g., Whitney et al. 2003b; Crapsi et al. 2008). For example, an edge-on disk around a T Tauri star (intrinsically Class II) could mimic the rising SED of a Class I protostar, while a pole-on Class I source could resemble a flat-spectrum or Class II object. This degeneracy motivated theoretical studies with radiative transfer models that could separate appearance from physical conditions. Based on Whitney et al. (2003a), Robitaille et al. (2006) created a large grid of modeled SEDs and introduced the **Stage** classification system, in which Stage 0/I objects are characterized by significant envelope accretion rates ($\dot{M}_{\rm env}/M_\star > 10^{-6}\,{\rm yr}^{-1}$), Stage II objects have little envelope but retain substantial disks ($M_{\rm disk}/M_\star > 10^{-6}$ at least, usually observed to be 10^{-3}–10^{-2}), and Stage III objects have negligible envelopes and disks ($M_\star \gg M_{\rm disk}, M_{\rm env}$). André et al. (2014) further separated Stage 0 from Stage I by the relative mass of envelope versus star and disk ($M_{\rm env} > M_\star + M_{\rm disk}$), giving a physical foundation to the empirical Class 0 definition. The timescales of different stages were estimated by large IR surveys with support and verification by theoretical models. Statistically speaking, the embedded protostellar Stages (0 and I) are relatively short (∼0.5 Myr in total; André et al. 2000; Enoch et al. 2009; Evans et al. 2009) compared to the disk-bearing Stages (II and III), which typically last for 1–3 Myr and 5–10 Myr, respectively (e.g., Haisch et al. 2001; Hernández et al. 2007; Fedele et al. 2010; Williams & Cieza 2011; Ribas et al. 2014).

In summary, the classification of YSOs into observational *Classes* and physical *Stages* converted a once qualitative picture of low-mass star formation into a coherent evolutionary sequence from cold dense cores to pre-main-sequence stars. Despite the distinctions between the two frameworks, they both contribute to advancing our understanding of low-mass star formation, with the Class system offering a practical observational taxonomy and the Stage system providing the physical context needed to interpret those observations.

1.1.3 High-mass star formation

Compared to the low-mass regime, massive stars form in denser, more distant, and highly clustered environments over shorter timescales ($\sim 10^5$ yr), making the observations of embedded prestellar cores very challenging (see reviews by Beuther et al. 2007; Zinnecker & Yorke 2007; Tan et al. 2014; Motte et al. 2018). Although the broad pictures of gravitational collapse, disk-mediated accretion, and bipolar outflows are shared with low-mass star formation, massive protostars typically reach the main sequence while still accreting, and their radiative and mechanical feedback rapidly reshapes their surroundings. As a result, observers often speak of an approximate

sequence from infrared-dark clumps and massive dense cores, through chemically enriched hot molecular cores, to hyper/ultracompact H II regions, and finally compact H II regions around emerging OB clusters.

How massive stars form has been under active debate and development. One of the dominant theories is **core accretion** (or **monolithic collapse**), where massive stars condense from high-surface-density cores and their envelopes collapse through massive, self-gravitating disks at high accretion rates (McKee & Tan 2003). This scenario is sometimes considered as a scale-up version of low-mass star formation, but only to the first order. Another popular theory is **competitive accretion**, where massive stars begin as lower-mass seeds within a common gravitational potential, and gain the bulk of their mass from the surrounding clump via Bondi–Hoyle accretion (Bonnell et al. 2001; Bonnell & Bate 2006). The competitive accretion theory was further extended into the **global hierarchical collapse** (GHC) theory to explain the observed "ridge and hub" structures in high-mass star-forming (HMSF) regions (Motte et al. 2018). In GHC, the classical Bondi–Hoyle accretion from a quasi-uniform background is upgraded to a non-homologous, multi-scale gravitational contraction of the entire cloud, in which collapse proceeds faster on smaller scales, and gas inflows from clouds through filaments (ridges) to clumps and cores (hubs; Vázquez-Semadeni et al. 2017, 2019; Motte et al. 2018).

The extreme physical conditions associated with high-mass star formation (e.g., high densities, intense radiation fields, strong turbulence, and rapid evolutionary timescales) can give rise to chemical environments that are very different from those of low-mass protostars. In particular, the high temperatures and ultraviolet (UV) radiation rates around massive protostars are expected to facilitate a wide range of chemical processes. In this context, high-mass star-forming regions can serve as a testbed for investigating how chemistry responds to different physical conditions during star formation.

1.2 Chemistry in early stages of star formation

In this section, chemical processes during the early stages of star formation are introduced. Considering the scope of this thesis and the fact that a large fraction of interstellar molecules were detected toward prestellar cores and hot cores around protostars, the chemistry occurring in other components (e.g., disks and outflows) and later stages (e.g., later than Class I) are left out. The chemical evolution will be divided into three phases, which are applicable in both low- and high-mass regimes, with the main difference lying in timescales.

1.2.1 Cold prestellar and collapsing phase

As the diffuse ISM evolves into dense clouds ($n \gtrsim 10^4$ cm^{-3}), the interstellar UV radiation is efficiently shielded. With the attenuation of photoionization and photodissociation, the dominant carbon species in the gas phase shifts from C^+ to C, and finally to CO (Leung et al. 1984; van Dishoeck & Black 1986; Snow & McCall 2006). Without photoelectric heating, the shielded interiors of dark clouds can be cooled down to ∼10 K by both dust and molecular gas (Goldsmith 2001).

At such low temperatures, gas-phase chemistry is mostly driven by cosmic rays (CRs), given that ion-neutral reactions are usually exothermic and barrierless (Herbst & Klemperer 1973). Neutral-neutral reactions are usually endothermic or have considerable activation barriers, which are very inefficient under cold conditions. A typical starting point of CR-induced gas-phase chemistry is H_3^+ produced by ionization of H_2 via $H_2 + CR \rightarrow H_2^+ + e^- + CR$ and $H_2^+ + H_2 \rightarrow H_3^+ + H$. More complex molecular ions are then formed through $H_3^+ + X \rightarrow XH^+ + H_2$. These proton-transfer reactions are exothermic and barrierless, and therefore can proceed very efficiently at ∼10 K. In particular, the proton transfer with HD leads to isotopic exchange (Watson 1974):

$$H_3^+ + HD \rightleftharpoons H_2D^+ + H_2, \tag{1.2}$$

which is exothermic in the forward direction by 232 K. The H_2D^+ ions can further transfer D to many other molecules via ion-neutral reactions, and hence promote the overall deuteration in the chemical network (see review by Ceccarelli et al. 2014). On the other hand, many hydrogenation reactions of molecular ions in the form of $X^+ + H_2 \rightarrow XH^+ + H$ are endothermic or possesses a potential barrier, thus are suppressed at low temperatures. As a result, many of the molecules present in cold dense clouds are unsaturated, which has been observed toward several prestellar cores (e.g., Spezzano et al. 2017; Xue et al. 2025).

As gas-phase chemistry evolves, a fraction of the gas-phase material begins to freeze onto dust grains at low temperatures; only the lightest species (e.g., H, H_2, and He) can sublimate efficiently at ∼10 K and thus remain in the gas phase. The adsorbed atoms and molecules, either via physisorption or chemisorption, then contribute to ice formation through three main reaction mechanisms: Langmuir-Hinshelwood (L-H), Eley-Rideal (E-R), and hot-atom. In the L-H mechanism, also known as the diffusive mechanism, reactants that are originally adsorbed at different potential minima (i.e., binding sites) have sufficient mobility to meet each other by thermal hopping or quantum tunneling over the energy barrier between binding sites. In the E-R mechanism, a gas-phase species is directly adsorbed on top of an existing adsorbate and reacts with it. In the hot-atom mechanism, a gas-phase species lands on a vacant binding site but is able to encounter a nearby adsorbate after brief diffusion. Among the three mechanisms, L-H (diffusive) has been the most studied and implemented one, but recent modeling studies call attention to the E-R (non-diffusive) mechanism in the ice formation of complex molecules, given the limited diffusion efficiency of heavy radicals at low temperatures (Jin & Garrod 2020; Garrod & Herbst 2006; also see Sect. 1.3.2).

Figure 1.1 summarizes the grain-surface chemistry in the prestellar and protostellar stages. Water (H_2O) is the first ice species that forms and accumulates on dust grains[4]. As the visual extinction (A_V) reaches the threshold of ∼1.6 mag (Whittet et al. 2001), H_2O ice can form efficiently via hydrogenation of O atoms (Tielens & Hagen 1982; Dulieu et al. 2010) and builds up a H_2O-rich layer on bare grains. A H_2O-rich environment is also referred to as a **polar** environment, indicating that the ice matrix is made of molecules with large permanent dipole moments. In the same epoch, NH_3 and CH_4 form by successive H-addition to the accreted N and C atoms (e.g., Hidaka et al. 2011; Qasim et al. 2020). Once CO forms abundantly in the gas

[4]Molecular hydrogen is the first molecule that forms on dust grains, but it promptly desorbs after forming exothermically and never builds up an ice mantle.

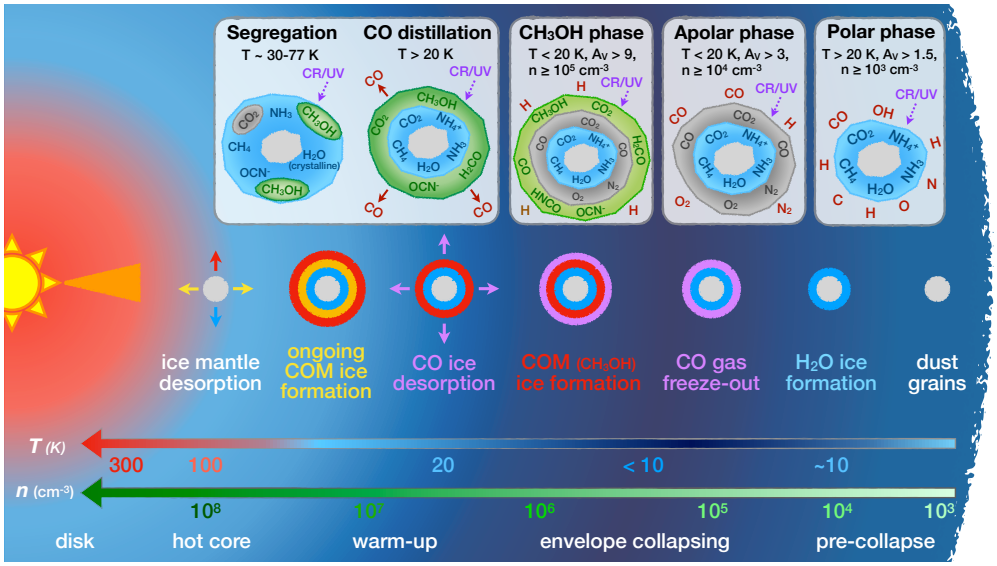


Figure 1.1: Schematic of the grain-surface chemistry in the prestellar and protostellar stages. From right to left, the diagram illustrates the formation and evolution of interstellar ices on dust grains from early to late stages. Different chemical phases are shown in greater detail at the top. Physical stages during star formation along with their typical temperatures and densities are indicated at the bottom. Adapted from Fig. 14 in Herbst & van Dishoeck (2009) and Fig. 10 in Boogert et al. (2015).

phase and gradually accretes onto dust grains, CO_2 also begins to form in the H_2O-rich layer via CO + OH → CO_2 + H[5] and CO + O (Watanabe & Kouchi 2002; Ioppolo et al. 2011; Minissale et al. 2013; Garrod & Pauly 2011). The freeze-out of gas-phase CO speeds up as density increases toward the cloud center, and when density exceeds $\sim 10^5$ cm^{-3}, CO freeze-out becomes so efficient that a CO-rich (**apolar**) ice layer is built upon the H_2O-rich one, with the gas-phase CO depleted significantly (Caselli et al. 1999). This process, known as "catastrophic CO freeze-out", leads to substantial accumulation of CO ices, which further boosts the formation of CO_2 (Pontoppidan 2006) and a series of hydrogenated products such as H_2CO and CH_3OH (e.g., Watanabe & Kouchi 2002; Fuchs et al. 2009; Cuppen et al. 2009). During this cold phase, the grain-surface chemistry is mostly driven by the diffusion of H atoms and CR-induced chemistry (e.g., bombardment of energetic particles and irradiation of secondary UV fields), which produce most of the commonly observed interstellar ices, except that ions such as NH_4^+, OCN^-, and $HCOO^-$ are formed via acid-base reactions between NH_3 and HNCO or HCOOH (Raunier et al. 2004; van Broekhuizen et al. 2004).

[5]Although this reaction has long been considered the main formation pathway of CO_2 ice, recent theoretical and laboratory studies show that it mainly produces HOCO and requires subsequent radical reactions to yield CO_2 (Molpeceres et al. 2023; Ishibashi et al. 2024).

1.2.2 Warm-up phase

Once a protostar is born at the center of the collapsed prestellar core (see Sect. 1.1.1), the envelope is gradually heated from the inside out. As temperature increases inward from ∼10 K to over 100 K, thermal desorption (or sublimation) of mixed ices proceeds in successive steps rather than in a one-off manner, which has been observed in temperature-programmed desorption (TPD) experiments[6] (e.g., Collings et al. 2004; Minissale et al. 2022; Kruczkiewicz et al. 2024). Species with higher binding energies[7] (E_b) tend to sublimate later at higher temperatures, although in some cases more volatile species (i.e., those with smaller E_b) can be trapped by H_2O—one of the least volatile common ices. As temperature reaches ∼20 K, apolar ices such as CO, CH_4, N_2, and O_2 start to leave their original ice matrices and sublimate from dust grains. In particular, the distillation of CO from CO_2 ice is observable via the CO_2 band at 15.2 µm, which develops a double-peak profile due to the emergence of the narrow band of pure CO_2. At higher temperatures (30 K to ∼77 K; Boogert et al. 2000; Öberg et al. 2009a), CO_2 further segregates from H_2O-rich ices but sublimates later at ∼45–90 K due to the trapping of H_2O (Fayolle et al. 2011).

While volatile species successively desorb from dust grains, the increased temperature also stimulates the formation of more complex molecules via the diffusion of heavier radicals. For example, CH_3OCHO can be formed through the recombination between HCO and CH_3O radicals (Garrod & Herbst 2006); similar routes are suggested for CH_3OCH_3 (via CH_3 + CH_3O) and HCOOH (via HCO + OH). However, the picture of COM formation is more complicated than diffusive grain-surface chemistry in the warm-up phase; a more detailed introduction is provided in Sect. 1.3.2. Regardless of how complex molecules are formed, they are likely trapped in mixtures with H_2O and CH_3OH (see Chapter 4). As a result, these trapped ices are released into the gas phase via co-desorption with H_2O and CH_3OH at ∼100 K, at which temperature all volatiles sublimate from dust grains.

1.2.3 Hot core phase

In compact regions around protostars where the temperature exceeds 100 K (∼10^2–10^3 au; i.e., hot cores), no volatile ices are left on dust grains, and the chemistry shifts decisively to the gas phase. This is the most popular phase for astrochemical observations in which a large fraction of gas-phase molecules, especially COMs, are discovered and studied to infer their formation mechanisms and evolution histories (see Sect. 1.3.3.3).

The high temperature in hot cores (100–300 K) permits a richer network of gas-phase reactions to occur than in cold prestellar cores, including those that are endothermic and have energy barriers. However, many gas-phase reactions proposed for COM formation were found to be inefficient to explain the observed abundances (e.g., Horn et al. 2004). There has also been a long-standing debate on the relative importance of gas-phase and grain-surface chemistries to COM formation, which is introduced in more detail in Sect. 1.3.2. Nevertheless, it is possible that the ice inven-

[6] In a TPD experiment, molecules are deposited first on a cold surface and then heated steadily, during which their desorption is monitored by a mass spectrometer or other detectors.

[7] Binding energy, also known as desorption energy (E_des), is the energy barrier that a molecule must overcome to leave the surface.

tories released from dust grains will be reprocessed by the gas-phase chemistry, which should be taken into account when interpreting the observational results of gas-phase molecules.

Hot cores are known for their richness of COMs. A few protostars, however, are found to harbor rich carbon-chain molecules that are dominant in prestellar clouds but deficient in hot cores (e.g., Sakai et al. 2008, 2009). This distinct chemical inventory is explained by the warm carbon chain chemistry (WCCC; Sakai & Yamamoto 2013), which is suggested to originate from a short prestellar phase. If the gravitational collapse proceeds fast enough, a considerable fraction of carbon atoms in the gas phase will be directly accreted onto dust grains before being converted into CO. These C atoms are then hydrogenated into CH_4 molecules on grain surfaces, which sublimate at 20–30 K early in the warm-up phase and trigger a series of gas-phase reactions to form carbon-chain molecules (Sakai & Yamamoto 2013). The lack of CO in the ingredients of grain-surface chemistry also lowers the production of common O-bearing COMs via hydrogenation of CO.

Toward the end of the hot core phase, materials in the protostellar envelope are either dissipated or accreted onto a circumstellar disk. With the development of observational techniques, many interstellar molecules including COMs, though with lower abundances and less complexity, have been observed toward protoplanetary disks (see Sect. 1.3.3.3). It is suggested that both inheritance from the protostellar phase and reset/reprocessing in later evolutionary stages are likely to contribute (Öberg et al. 2023).

1.3 Complex organic molecules

1.3.1 Categories

The earliest appearance of the term "complex organic molecules" in the astronomical context can be traced back to the 1970s (e.g., Sagan 1972), but the first explicit definition (carbon-bearing molecules with at least six atoms) was given by Herbst & van Dishoeck (2009) and was widely adopted by the following studies. So far, more than 150 COMs among ∼340 interstellar molecules have been detected in the gas phase, and the increasing rate of new detections has remained at a fairly high level since the 21st century, especially with the advent of ALMA (e.g., McGuire 2018, 2022). The detection of solid-phase COMs (or COM ices) was only considered robust for methanol (CH_3OH), the simplest and most abundant COM, before the latest JWST observations showcased the feasibility of detecting COM ices larger than CH_3OH (see Sect. 1.3.4.4 and Chapter 4). Table 1.1 lists a number of detected gas-phase COMs with up to three carbons categorized by their functional groups. The most studied COMs are the oxygen or nitrogen-bearing ones with one to two carbons, whereas the detection limit has gone far beyond this.

1.3.2 Formation theories

How COMs are formed is still an open question under active investigation. Early chemical models for simple molecules mainly consider gas-phase exothermic ion-molecule reactions initiated by cosmic-ray ionization (e.g., Herbst & Klemperer 1973). After

Table 1.1: Selected interstellar COMs with up to three carbon atoms.

Category	Functional group	1-carbon	2-carbon	3-carbon
Alcohol	C–OH	CH_3OH	C_2H_5OH C_2H_3OH $(CH_2OH)_2$ CH_3OCH_2OH	n/i-C_3H_7OH
Aldehyde	–C=O–H	H_2CO	CH_3CHO CH_2OHCHO	C_2H_5CHO
Ether	–C–O–C–		CH_3OCH_3	$C_2H_5OCH_3$
Ester	–O–C=O–		CH_3OCHO	C_2H_5OCHO
Acid	–C=O–OH	HCOOH	CH_3COOH	
Ketone	C–(C=O)–C		CH_2CO	CH_3COCH_3
Epoxide	cyclic –(C–O–C)–		c-C_2H_4O	c-C_3H_6O
Cyanide	–C≡N	HCN	CH_3CN	C_2H_5CN
Amine	$–NH_2$, –NH–	CH_3NH_2 NH_2CHO	$C_2H_3NH_2$ CH_3NHCHO	
Cyanate	–O–C≡N	HNCO	CH_3NCO	C_2H_5NCO
Alkane	saturated hydrocarbon	CH_4	C_2H_6	C_3H_8
Alkene	–C=C–, hydrocarbon		C_2H_4	C_3H_6
Alkyne	–C≡C–, hydrocarbon		C_2H_2	C_3H_4

Notes: (1) species in red are those that have been studied in this thesis (Chapters 2 & 3); (2) species in gray are simple molecules by definition (<6 atoms); HCOOH and CH_2CO are not strictly COMs, but they are usually studied together with O-COMs; (3) uncommon O/N-bearing COMs and all S-bearing species, carbon chains, and aromatic compounds are not included; (4) CH_2OHCHO is also considered as the simplest sugar; (5) CH_2CO is not strictly an ketone since one side of its carbonyl group is ended with a hydrogen.

the discovery of COMs other than CH_3OH in the 1970s (e.g., Snyder et al. 1974; Brown et al. 1975) and following wideband line surveys toward Sgr B2 and a few hot cores in Orion in the 1980s (e.g., Sutton et al. 1985; Cummins et al. 1986; Blake et al. 1987), the gas-phase network was extended to COMs, of which the formation is based on simple molecules released from grain mantles (e.g., H_2CO, NH_3, and at most CH_3OH; Charnley et al. 1992). The grain surface chemistry at that time only included cosmic-ray induced formation of radicals at low temperatures, followed by recombination reactions and thermal desorption in the warm-up phase.

However, both theoretical and experimental studies in the 2000s challenged the efficiency of gas-phase formation of COMs, and instead suggested significant contribution from grain-surface (solid-phase) chemistry. A study combining quantum chemical calculations and experiments showed that the previously adopted gas-phase formation route of CH_3OCHO (via $CH_3OH_2^+$+H_2CO) is inefficient due to the high activation energy barrier, and its product in laboratory experiments is prone to an isomer of CH_3OCHO; even with alternative gas-phase pathways, the observed abundances could not be achieved (Horn et al. 2004). Meanwhile, the gas-grain chemical model proposed by Garrod & Herbst (2006) provided results in closer agreement with

the observations, which emphasized the importance of including grain-surface chemistry and its coupling with gas-phase chemistry. More support for solid-phase COM formation came from cryogenic experiments. Watanabe & Kouchi (2002) showed the feasibility of efficiently forming CH_3OH by successive hydrogenation of CO at 10 K, which was confirmed by independent experiments and cross-checked with Monte Carlo simulations by another group (Fuchs et al. 2009). Energetic experiments also showed that a variety of larger O-COMs can be further produced by UV photolysis and warm-up starting from CH_3OH-rich ices (Öberg et al. 2009b).

Since the 2010s, grain-surface chemistry has been forced to undergo a series of revisions in response to the unexpected detection of gas-phase COMs in cold ($\lesssim 10$ K) prestellar cores (e.g., Bacmann et al. 2012; Vastel et al. 2014; Jiménez-Serra et al. 2016a; Soma et al. 2018). The presence of gas-phase COMs under such low temperatures, albeit in much lower abundances, casts doubt on the exclusive role of grain-surface chemistry and challenges the warm diffusion theory to form large COMs. Meanwhile, to explain the new observational results, Balucani et al. (2015) proposed a new gas-phase formation route for CH_3OCH_3 and CH_3OCHO, which is made up of low-barrier neutral–neutral and radiative-association reactions starting from non-thermally desorbed CH_3OH. Experimental studies also found that photodesorption of CH_3OH ices is far less efficient than previously assumed in chemical models ($\sim 10^{-5}$ molecules per photon; Bertin et al. 2016; Cruz-Diaz et al. 2016), potentially supporting the proposition of Balucani et al. (2015), whereas chemical/reactive desorption and cosmic-ray induced desorption seem to be more plausible explanations.

The questioning of the necessity of warm grains to form large COMs was mitigated by non-energetic experiments at low temperatures (10–15 K). Fedoseev et al. (2015) showcased the possibility of forming backbones (i.e., carbon–carbon bonds) of two-carbon O-COMs through recombination between two HCO radicals, with CH_2OHCHO and $(CH_2OH)_2$ observed in the product of their CO hydrogenation experiments at 13 K. More cryogenic experiments without energetic irradiation were able to produce CH_2CO, CH_3CHO, C_2H_3OH, and C_2H_5OH via hydrogenation of C, CO, H_2CO, and C_2H_2 (input ingredients varied among different experiments), with the potential contribution of H-abstraction and radical-radical recombination (Chuang et al. 2016, 2020; Fedoseev et al. 2022). However, CH_3OCH_3 has long been absent from the product of non-energetic experiments, and the yields of some species represented by CH_3OCHO in UV-photolysis experiments are significantly lower than the observed levels (Chuang et al. 2017).

Although our understanding of COM formation is still incomplete (also seen in *Chapter 4*), recent development of **non-diffusive chemistry** further addresses the questions posed by the detection of cold COM gas and reinforces the contribution of grain-surface chemistry (Jin & Garrod 2020; Garrod et al. 2022). Non-diffusive chemistry lifts the temperature constraint on the diffusive reactions involving radicals larger than H atoms, which is the case for most COM formation. Instead of waiting for heavy species to thermally diffuse and meet each other, reactions in non-diffusive chemistry can occur between reactants produced adjacent or very close to each other, hence enhancing the reaction rates. There are four main mechanisms: Eley–Rideal (E-R), three-body (3-B), three-body excited-formation (3-BEF), and photodissociation-induced (PDI) reactions. Three-body reactions are those that occur immediately after an initiating reaction forms a product adjacent to a potential partner. Three-body

excited-formation considers a special case of basic 3-B, where the product of an initiating reaction is excited and hence able to overcome the barrier of following reactions with other reactants. This mechanism plays an important role in producing enough CH_3OCHO in ice mantles to match the observed abundances in the gas phase. Photodissociation-induced reactions are similar to 3-B, except that the reactants are produced by UV photodissociation. The involvement of non-diffusive chemistry allows large COMs to form earlier at low temperatures without requiring mobility of heavy radicals (e.g., via 3-B reactions) and in more efficient ways (e.g., 3-BEF for CH_3OCHO). The reframed grain chemistry is now more promising to reconcile with the cold-COM detections and bridge the gap between simulations/experiments and observations.

There are other hotly debated topics in COM studies besides the formation routes in different phases; for example, the differentiation between O- and N-bearing COMs as suggested by differences in their emission distributions (e.g., Csengeri et al. 2019; Mininni et al. 2023; Nazari et al. 2023b; Busch et al. 2024). We do not dive deeper into them here considering that this thesis only involves O-COMs and the main focus is on gas-to-ice comparisons. Nevertheless, it is clear that our understanding of COM formation will continue to advance through the collaboration of observers, theorists, and experimentalists.

1.3.3 Observations of gas-phase COMs

1.3.3.1 Rotational spectroscopy

COMs in the gas phase are mostly observed at submillimeter (submm) or millimeter (mm) wavelengths for their rotational transitions. There are a few reasons to look at (sub)mm wavelengths: (1) pure rotational lines of molecules with permanent dipole moments have high Einstein A coefficients (10^{-6}–10^{-3} s^{-1}), making their emission bright in dense hot cores under quasi-local thermodynamic equilibrium (LTE); (2) the rotational transitions of large molecules such as COMs mainly fall at (sub)mm wavelengths, or equivalently tens to hundreds of GHz in frequency; (3) the Earth's atmosphere has good transmission windows in the 80–900 GHz range, making it possible for modern radio telescopes (see Sect. 1.3.3.2) to deliver spectra with high spectral and angular resolution as well as wide bandwidths.

An observed spectrum records three principal properties of each spectral line: (i) the central frequency, (ii) the integrated intensity, and (iii) the profile (width and shape). The rest frequency of a transition is determined by the energy difference between its upper and lower levels. However, the rest frequency usually deviates from the observed central frequency of this line because of the radial velocity of the emitting gas with respect to the observer (usually denoted as $v_{\rm lsr}$, where "lsr" refers to local standard of rest). The intensity is dependent on the column density of this molecule and its level populations, which are determined by the kinetic temperature ($T_{\rm kin}$) under strict LTE conditions. In practice, however, molecular emission in hot cores may still remain sub-thermally excited (i.e., $T_{\rm ex} < T_{\rm kin}$) at densities of $< 10^9$ cm^{-3} (e.g., Fig. 6 of Johnstone et al. 2003), and the LTE assumption usually indicates that the observed level populations can be characterized by a single temperature, which is not necessarily the same as $T_{\rm kin}$. The line profile reflects the velocity field (thermal and non-thermal) and kinematic components of the probed gas. The line width is

usually characterized by the full width half maximum (FWHM), which is typically smaller in prestellar cores (< 1 km s^{-1}) and low-mass protostars (< 3 km s^{-1}) than in HMSF regions with stronger turbulence ($\gtrsim 5$ km s^{-1} and up to a few tens of).

Due to the larger amount of atoms, COMs exhibit additional spectral complexity than simpler molecules. The two most relevant conceptions are torsion and conformer. **Torsion** is the internal rotation of a group that can rotate almost freely relative to the molecular frame through a periodic potential. The most common case in astrochemistry is the methyl group (–CH_3), which has a threefold rotational symmetry and vertical mirror planes (also denoted as C_{3v} internal rotors). When the internal rotation (i.e., torsion) couples to the overall rotation, energy levels are better described by a torsion–rotation Hamiltonian rather than a simple rigid rotor. In databases for rotational spectra, CH_3-bearing molecules are usually labeled with an additional torsion quantum number (v_t=0, 1, ...) and grouped into two symmetry species A and E, which represent the two lowest-energy torsional states. In observations, the A and E families of some COMs (e.g., CH_3OH) can be fitted separately, and their ratios can be used to infer the formation temperature given E/A ~ 1 at $T \gtrsim 20$ K and E/A ~ 0.6–0.7 at $T \sim 10$ K. Besides different torsional states, many COMs also have more than one **conformer** (also known as conformational isomer) that is distinguished by different rotations of a single bond in a molecule. Common cases in astrochemistry are the gauche- and trans-conformers of alcohols (e.g., a-$(CH_2OH)_2$ and g-$(CH_2OH)_2$ studied in Chapter 2) and ethers. In principle, the most abundant conformer corresponds to the lowest potential energy, but in interstellar environments, temperatures may not be high enough to overcome the interconversion barriers among the low-energy conformers, and the abundance of more than one conformer can remain at a detectable level. In astronomical databases, conformers are often listed separately and treated as different species.

For molecules that include nitrogen (^{14}N) or deuterium (D), their rotational transitions (e.g., $J+1 \rightarrow J$) are further split by interactions with nuclear spins (denoted by I) into **hyperfine** transitions, since the total angular momentum $\vec{F} = \vec{J} + \vec{I}$ is split into $F = |J-I|$, $|J-I|+1$, ..., $J+I$. The most relevant coupling in astrochemistry is with the electric quadrupole, which is the case for ^{14}N and D (both with $I = 1$ and each rotational transition usually splits into a triplet $J-1$, J, and $J+1$). Although the frequency separation of hyperfine lines is usually small (typically kHz–MHz), modern millimeter facilities have enough spectral resolution to resolve them.

Another case of transition splitting occurs in symmetric top molecules that have one special principal axis, also known as the symmetry axis. For these molecules, another quantum number K is introduced to describe the projection of the total rotational angular momentum $\vec{J}$ onto the symmetry axis. If the dipole is parallel to the symmetry axis (e.g., CH_3CN and CH_3CCH), the selection rule is $\Delta K = 0$; hence for a single $J+1 \rightarrow J$ transition, there is a series of lines for each $K = 0, 1, ..., J$, also known as the **K-ladder**. In astrochemical studies, K-ladders of CH_3CN and CH_3CCH are frequently used as thermometers because their K-ladder transitions sample a wide range of $E_{\rm up}$ while keeping almost all the other parameters (e.g., Einstein A coefficient) the same. Therefore, their relative intensities can provide robust constraints on the temperature assuming LTE (Boltzmann distributions). Most COMs are not perfect symmetric tops and two (pseudo–)quantum numbers K_a and K_c are introduced to track the prolate and oblate limits, respectively, but they do not strictly represent the

projections of $\vec{J}$.

Rotational transitions of interstellar molecules are measured in the laboratory under isolated and well-controlled conditions that allow individual transitions to be resolved with uncertainties far below those achievable in astronomical spectra. The measured spectroscopic parameters are compiled into publicly accessible molecular line databases. Two of the most widely used resources are the Cologne Database for Molecular Spectroscopy (CDMS[8]; Müller et al. 2001, 2005) and the Jet Propulsion Laboratory (JPL[9]) molecular spectroscopy catalog (Pickett et al. 1998). These databases provide standardized line lists including transition frequencies (with uncertainties), upper-state energies, line intensities (usually in the form of Einstein A coefficients), and partition functions over a range of temperatures. Such information is essential for line identification and for deriving physical parameters through spectral modeling. As astronomical observations continue to develop toward higher sensitivity, broader bandwidths, and more complex molecules, the accuracy and completeness of laboratory spectroscopic data remain a key limiting factor in observational studies.

1.3.3.2 Millimeter facilities

Observations of rotational lines emitted by gas-phase interstellar molecules at millimeter wavelengths were initiated by the first detection of CO J=1–0 line at 115 GHz back in 1970 (Wilson et al. 1970). Millimeter astronomy was further boosted by the development of single dishes and interferometer arrays in the following decades. Currently, many millimeter facilities are still in active service (see Table 1.2), but ALMA is undoubtedly the flagship among them for studying protostellar sources given its unmatched combination of high sensitivity and resolution (both spatial and spectral).

Sitting on the high, dry Chajnantor plateau in northern Chile, ALMA consists of the main 12-m array with fifty antennas, the Atacama Compact Array (ACA) with twelve 7-m antennas, and four total-power (TP) antennas. Since Cycle 10 (2023), nine out of ten receiver bands covering 35–950 GHz (0.3–8.5 mm) have been available for use, and the remaining Band 2 (67–90 GHz) will become available soon. ALMA has the longest baseline of 1.6 km among the ground-based interferometer arrays, providing sub-arcsecond to milli-arcsecond angular resolution in its longest configuration (C-10). On the other hand, the short spacings of ACA and TP can recover extended emission on arcminute scales that are filtered out by ALMA, which is important for low-$E_{\rm up}$ lines or abundant molecules. In terms of spectroscopic performance, ALMA offers multiple choices for channel spacing (down to 15 kHz) and bandwidth (up to 1.875 GHz per window, $4\times1.875 = 7.5$ GHz instantaneously), which can be customized flexibly for different targets and science cases. These features combine to make ALMA the most powerful facility for COM observations. Its high angular resolution alleviates beam dilution effects and separates different chemical and kinematical components, which is crucial for COMs emitting from compact hot core regions. Moreover, its high sensitivity enables detection of rare isotopologs and higher-excitation lines, yielding more accurate column density measurements for optically thick molecules (e.g., the main isotopologs).

Other interferometer arrays such as NOEMA and SMA have coarser spatial reso-

[8] `https://cdms.astro.uni-koeln.de/`
[9] `https://spec.jpl.nasa.gov/`

Table 1.2: Selected millimeter telescopes in service

Facility	**Type & Aperture**	**Location**	**Frequency coverage**	**Angular resolution (beam size)**	**Spectral resolution**	**Instantaneous bandwidth**
ALMA[a]	Interferometer array (50×12 m + 12×7 m + 4×TP)	Chajnantor plateau, Chile (∼5000 m)	35–950 GHz (0.3–8.5 mm)	down to 0.006–0.14″ (Config. 10)	15.3 kHz to 15.6 MHz	up to 7.5 GHz (4 × 1.875)
NOEMA[b]	Interferometer array (12×15 m)	Plateau de Bure, Hautes-Alps, French (∼2550 m)	80–370 GHz (0.8–3.7 mm)	down to ∼0.1–1.0″ (Config. A)	native 2 MHz, can zoom to 62.5 kHz	up to 31 GHz (PolyFiX)
SMA[c]	Interferometer array (8×6 m)	Maunakea, Hawaii, USA (∼4080 m)	180–420 GHz (0.7–1.7 mm)	∼0.25–5″ at 345 GHz (Very Extended Config.)	140 kHz	up to 44 GHz
GBT[e]	Single dish (100 m)	Green Bank, WV, USA (∼800 m)	0.29–50.5 & 67–116 GHz (2.6 mm–1 m)	$\sim 780''/\nu$ [GHz]; ∼ 7.8″ at 100 GHz	down to 0.02 kHz (VEGAS)	up to 4–6 GHz
IRAM 30-m[d]	Single dish (30 m)	Pico Veleta, Spain (2850 m)	80–370 GHz (0.8–3.7 mm)	$2460''/\nu$ [GHz]; 6.6–31″	50/195 kHz (FTS)	up to 32 GHz (FTS)
Yebes 40-m[g]	Single dish (40 m)	Yebes, Castilla–La Mancha, Spain (∼980 m)	31.5—-50 & 72—-90.5 GHz (3.3–9.5 mm)	$1763''/\nu$ [GHz]; ∼20–56″	38 kHz	up to 18.5 GHz
JCMT[f]	Single dish (15 m)	Maunakea, Hawaii, USA (4092 m)	325–375 GHz (0.8–0.92 mm; HARP)	∼ 14″ at 345 GHz	30.5 kHz to ∼1 MHz	up to 1.8 GHz (ACSIS)

The full name and the link for technical information (up to Sep 2025) of each facility are listed below:
[a] Atacama Large Millimeter/submillimeter Array (`https://almascience.eso.org/documents-and-tools`)
[b] NOrthern Extended Millimeter Array (`https://www.iram.fr/IRAMFR/GILDAS/doc/html/noema-intro-html/node1.html`)
[c] SubMillimeter Array (`http://sma1.sma.hawaii.edu/status.html#arrayconf`)
[d] Institut de Radioastronomie Millimétrique 30-meter telescope (`https://www.iram.fr/GENERAL/calls/w25/30mCapabilities.pdf`)
[e] Green Bank Telescope (`https://www.gb.nrao.edu/scienceDocs/GBTpg.pdf`)
[f] James Clerk Maxwell Telescope (`https://www.eaobservatory.org/jcmt/instrumentation/`)
[g] `https://rt40m.oan.es/rt40m_en.php` and Tercero et al. (2021)

lution and lower line sensitivity than ALMA due to shorter maximum baselines and smaller collecting area, but they also provide important complementary capabilities in observing northern sources that are unreachable by ALMA and offering wider instantaneous bandwidths for line surveys. Single-dish telescopes, though with much lower angular resolution, are still actively used for COM observations. Their high spectral resolution (kHz-level), wide instantaneous bandwidth, large recoverable scale, and decent sensitivity make them more favorable than ALMA for surveys of narrow lines and extended emission toward prestellar cores (Sect. 1.3.3.3). They also remain indispensable in mapping large-scale chemistry and preselecting promising regions for follow-up observations with interferometer arrays.

1.3.3.3 Line surveys

To better identify and quantify the abundances of gas-phase COMs, it is preferable to observe a wide frequency range to maximize line coverage; therefore, such observations are also referred to as "line surveys". With single dishes and early interferometer arrays, (sub)millimeter line surveys toward line-rich YSOs such as Orion KL and Sgr B2 had already revealed dozens of lines per GHz (e.g., Fig. 1.2), many of which are attributable to gas-phase COMs emitting from hot cores (see review by Herbst & van Dishoeck 2009). Long carbon chains were also discovered in cold prestellar cores, represented by the Taurus Molecular Cloud 1 (TMC-1). However, many of these observations suffered from beam dilution, in which case the molecular emission is not spatially resolved and the actual source size of each emitting molecule is unknown, causing uncertainties in the derived column densities and relative ratios.

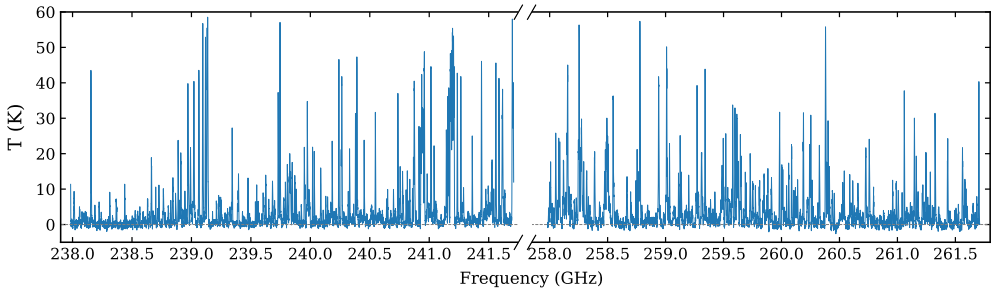


Figure 1.2: ALMA spectrum of G19.88-0.53, one of the high-mass hot cores studied in Chapters 2 & 3. Most of the displayed lines are emitted by gas-phase COMs.

The high sensitivity and angular resolution of ALMA significantly improve the depth, accuracy, and efficiency of line surveys. In early cycles, several dedicated line surveys were conducted for well-known hot corino/cores: the ALMA Protostellar Interferometric Line Survey (PILS) for the Class 0 binary system IRAS 16293-2422 (Jørgensen et al. 2016), the Exploring Molecular Complexity with ALMA (EMoCA[10] for Sgr B2(N), a massive star-forming region located in the central molecular zone (Belloche et al. 2016), and the G31.41+0.31 Unbiased ALMA sPectral Observational

[10]EMoCA was later upgraded to Re-exploring Molecular Complexity with ALMA (ReMoCA) with a higher sensitivity and angular resolution (Belloche et al. 2019).

Survey (GUAPOS) for the massive hot core G31.41+0.31 (Mininni et al. 2020). Covering a continuous ∼30 GHz bandwidth around 345 GHz (0.87 mm; ALMA Band 7) or 100 GHz (3 mm; ALMA Band 3), these programs delivered the most complete COM gas inventories for the three prototypical hot corino/cores. Many new species were discovered, including but not limited to chloromethane (CH_3Cl; Fayolle et al. 2017), urea (NH_2CONH_2; Belloche et al. 2019), propyl cyanide (C_3H_7CN; Belloche et al. 2014), propanol (C_3H_7OH; Belloche et al. 2022), and a number of deuterated COMs (e.g., Coutens et al. 2016; Manigand et al. 2019; Richard et al. 2021; Ilyushin et al. 2022; Drozdovskaya et al. 2022; Ferrer Asensio et al. 2023). The rich detection lists of these dedicated line surveys enable detailed investigations of isotopologs, isomers, and chemical similarities and differentiations between different species and different cores (e.g., Belloche et al. 2016; Calcutt et al. 2018; Mininni et al. 2023; Busch et al. 2024; Belloche et al. 2025). This is only available with sufficient depth and frequency coverage, but such observations are usually very expensive or even unrealistic with mm facilities other than ALMA.

Thanks to the high sensitivity of ALMA, deep and unbiased line surveys (i.e., those with low noise levels and large bandwidths) have become much more affordable, promoting the transition from case studies to large-sample surveys. Instead of staring at individual sources, the ubiquity of COMs has been examined from a broader perspective by exploring statistics of larger samples. For low-mass sources, the Perseus ALMA Chemistry Survey (PEACHES) targeted 50 embedded (Class 0/I) protostars and reported a detection rate of ≤58 % for COM emission (Yang et al. 2021). A sibling program of PEACHES, the ORion ALMA New GEneration Survey (ORANGES), detected a smaller fraction (∼26 %) of CH_3OH emission in 5 out of 19 targeted sources in the Orion Molecular Complex (OMC), which is a more illuminated environment than the Perseus (low-mass) star-forming region (Bouvier et al. 2022). A similar low detection rate for CH_3OH in OMC was found by another program, the ALMA Survey of Orion Planck Galactic Cold Clumps (ALMASOP), where only 11 out of 56 (∼20 %) Class 0/I sources exhibit CH_3OH emission (Hsu et al. 2022). Although these statistics hint at the non-ubiquity of COMs in star-forming regions, the non-detection of COM emission does not necessarily mean the absence of COMs; the interpretation of non-detection can be affected by factors such as detection sensitivity, disk shadowing, and dust attenuation (e.g., van Gelder et al. 2022; Nazari et al. 2022b).

In the high-mass regime, a substantial amount of ALMA data useful for COM studies has been obtained by a number of large-sample surveys, albeit not primarily intended for chemical purposes. These surveys include two ALMA Large Programs: the ALMA-IMF program (where IMF refers to initial mass function; Motte et al. 2018) targeting 15 HMSF regions in the Galactic plane, and the ALMA evolutionary study of high-mass protocluster formation in the GALaxy (ALMAGAL; Molinari et al. 2025) targeting over 1000 intermediate/high-mass dense clumps in the Galaxy. Another two ALMA programs, ALMA Three-millimeter Observations of Massive Star-forming regions (ATOMS; Liu et al. 2020) and Querying Underlying mechanisms of massive star formation with ALMA-Resolved gas Kinematics and Structures (QUARKS; Liu et al. 2024), observed the same sample of 146 HMSF regions at different bands (ALMA Bands 3 and 6). *Chapters 2 and 3* of this thesis also introduce a chemically oriented program, the Complex Chemistry of hot Cores with ALMA (CoCCoA), which exclusively focuses on line-rich hot cores with customized spectral setups for COM lines.

These large-sample surveys of HMSF regions not only increased the sample size of hot cores, but also enabled systematic and comparative studies of the abundances and distributions of COMs (e.g., Qin et al. 2022; Nazari et al. 2022a; Baek et al. 2022; Bonfand et al. 2024; Li et al. 2024, 2025; *Chapters 2 and 3* of this thesis). In particular, a similar trend of column density ratios between CH_3OH and some other COMs (especially CH_3OCH_3 and CH_3OCHO) is found for sources with luminosities across multiple orders of magnitude, shedding light on the histories and mechanisms of COM formation (e.g., Coletta et al. 2020; Nazari et al. 2022a; Chapter 2).

The investigation of gas-phase COMs is not limited to protostellar sources. Intensive line surveys have also been carried out for prestellar cores and protoplanetary disks (Class II objects). Numerous first detections of hydrocarbons (including long carbon chains, aromatics, and ions) were made toward the famous prestellar cloud TMC-1 by single dishes (e.g., Cernicharo et al. 2021; Wenzel et al. 2024, 2025). Two representative programs are the Q-band Ultrasensitive Inspection Journey to the Obscure TMC-1 Environment (QUIJOTE, with Yebes 40-m and IRAM 30-m radio telescopes; Cernicharo et al. 2022) and the GBT Observations of TMC-1: Hunting for Aromatic Molecules (GOTHAM; McGuire et al. 2020). These single-dish surveys excelled in detecting narrow lines of large molecules at low frequencies with their wide bandwidths ($\lesssim$50 GHz) and high spectral resolutions (kHz-level, $\lesssim$0.1 km s^{-1}).

For protoplanetary disks, it is already difficult to detect the most abundant COMs such as CH_3OH and CH_3CN, as these detections usually require a long integration time to achieve sufficient noise levels and angular resolutions. Early COM detections were reported for several disks that are either nearby (TW Hya; Walsh et al. 2016) or warm due to outbursts (V883 Ori; Lee et al. 2019b) or surrounding Herbig Ae/Be stars (MWC 480, HD 100546, and Oph IRS 48; Öberg et al. 2015; Booth et al. 2021; van der Marel et al. 2021; Brunken et al. 2022). In recent ALMA Cycles, follow-up line surveys have been conducted and more COMs in more disks are found (e.g., Booth et al. 2023, 2024a,b, 2025). The presence of COMs in protoplanetary disks, which is suggested to originate from sublimation of ice mantles on dust grains, provides evidence for ice inheritance across different star formation stages. Whether these observed COMs were inherited from earlier stages or formed in situ depend on many factors (e.g., thermal history and disk geometry), and recent studies suggest that both scenarios can play a role in shaping the disk chemistry (Öberg et al. 2023).

1.3.4 Observations of solid-phase COMs

1.3.4.1 Vibrational bands of interstellar ices

In dense star-forming regions, ices on dust grains serve as important reservoirs for interstellar molecules, where more than 50% of the heavy elements can be locked up. In the solid phase, molecules cannot rotate freely and emit rotational lines at (sub)millimeter wavelengths like those in the gas phase; instead, they can only emit or absorb energy through their vibrational transitions which fall in the near–to mid-infrared (∼2–20 µm). Because the dust extinction is lower in the infrared than in the optical, infrared light from a background star or a protostar can pass through dense cores and be absorbed by ices along the line of sight, making the IR vibrational absorptions a clean probe of interstellar ices.

For a free molecule, the total number of fundamental vibrational modes depends

on the number of atoms (N) and its structure; non-linear molecules have $3N-6$ and linear molecules have $3N-5$ fundamental modes, where $3N$ is the total number of mechanical degrees of freedom (DoF), and the subtracted degrees are for translations (i.e., moving as a whole in three directions) and rotations (around three or two axes, depending on the linearity). For symmetric molecules (e.g., H_2O and CO_2), some vibrational modes are degenerate with each other. In reality, molecules are not perfect harmonic oscillators, and non-fundamental modes such as overtones and combinations may appear due to anharmonicity (of which the observations have been reported by, e.g., Carr 1989; Keane et al. 2001; GRAVITY Collaboration et al. 2020; Brunken et al. 2024b). In ice mantles, interactions with environments can also produce additional features such as the H_2O libration mode at $\sim$13 µm and the O–H dangling mode of porous H_2O ice at $\sim$2.7 µm (Ehrenfreund et al. 1996; Noble et al. 2024).

With a larger number of atoms, COMs can have dozens of fundamental vibrational modes, but not all of them are detectable in IR observations. In principle, the intrinsic band strength A scales with $(\mathrm{d}\mu/\mathrm{d}Q)^2$, the square of the change in dipole along the mode, where μ and Q refer to the dipole moment and the vibrational coordinate, respectively. That is, modes that cause more charge movements are stronger (e.g., stretches of X–H, C–O, and C=O bonds). There are many strong COM bands, mostly CH_3 deformation and C–O stretching modes, sitting within the 6.8–8.8 µm range, which is also known as the **COM ice fingerprint range**. This range will be the focus of the COM ice studies presented in Chapters 4 and 5.

Unlike narrow gas-phase lines of rotational transitions, vibrational absorption features of ices are much broader, hence called "vibrational bands". The broadening effect is induced by interactions with neighboring molecules that shift the original vibrational transition by different amounts of energy. As a result, the profile of a vibrational band (including peak positions, widths, and shapes) can vary significantly with environments, with the two dominant factors being matrix polarity (polar/H_2O-rich or apolar/CO-rich) and temperature (amorphous/low-T or crystalline/high-T), as shown in Fig. 1.3. The diversity of band profiles was already observed in early detections of H_2O ices toward different sources (e.g., Merrill et al. 1976) and revealed the fact that interstellar ices are present in different mixtures rather than in pure form.

1.3.4.2 Laboratory IR spectroscopy

To infer ice compositions from observed vibrational bands, IR spectra of different ices are measured in the laboratory and compared with observations (see e.g., reviews by Allodi et al. 2013; Linnartz et al. 2015; Cuppen et al. 2024). Astrochemical experimentalists grow thick (10^1–10^3 monolayers) interstellar ice analogs by vapor-depositing gases onto cryogenic substrates ($\gtrsim$10 K) in ultrahigh vacuum ($\lesssim 10^{-9}$ mbar). Transmission or reflection IR spectra of deposited ices are then recorded using Fourier-transform infrared spectroscopy (FTIR) or reflection adsorption infrared spectroscopy (RAIRS), respectively. Parameters such as composition, thickness, temperature, and radiation processing are varied per experiment to investigate different topics. In practice, transmission spectroscopy is the default choice for characterizing IR features of ices (which is also the case relevant to this thesis), and reflection spectroscopy is preferred for studying surface chemistry. Three important properties can be collected from transmission FTIR for each ice species or mixture: (1) peak wavelengths (λ_{peak}) and shapes (including FWHM) of vibrational bands, (2) intrinsic band strengths A,

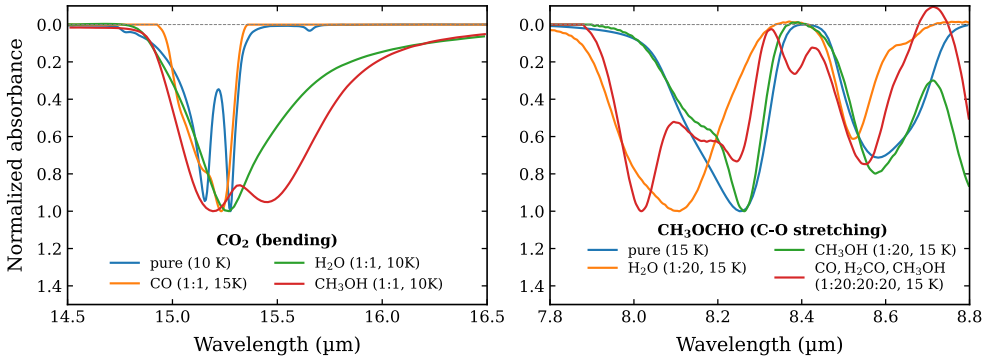


Figure 1.3: Laboratory-measured IR spectra of pure and mixed ices of CO_2 (left) and CH_3OCHO (right). The mixing constitute(s), mixing ratio, and temperature are indicated in the legend. In the right panel, some features do not belong to CH_3OCHO but the mixing constitutes (e.g., CH_3OH and H_2CO have a band at ∼8.9 μm and ∼8.0 μm, respectively).

and (3) optical constants n and k serving as required inputs for radiative transfer modeling. In this thesis, $\lambda_{\rm peak}$ and FWHM are used to identify ice compositions from the observed IR bands, and A is used to calculate ice column densities ($N_{\rm ice}$) of detected species. The relation between A and $N_{\rm ice}$ is given by

$$N_{\rm ice} = \frac{1}{A}\int_{\tilde{\nu}_1}^{\tilde{\nu}_2} \tau(\tilde{\nu})\,{\rm d}\tilde{\nu} = \frac{\ln 10}{A}\int_{\tilde{\nu}_1}^{\tilde{\nu}_2} Abs(\tilde{\nu})\,{\rm d}\tilde{\nu}, \tag{1.3}$$

where $\tilde{\nu}$ is the wavenumber, $\tilde{\nu}_1$ and $\tilde{\nu}_2$ indicate the range of the targeted vibrational band, τ is the optical depth, and Abs is the absorbance.

Driven by the demand for analyzing observational data, great efforts have been made in laboratories since the detection of the first interstellar ice band at 3.1 μm (Gillett & Forrest 1973), which was later attributed to amorphous (instead of crystalline) H_2O ice by Léger et al. (1979). This identification highlighted the need for laboratory spectra measured under astrophysical (i.e., low-temperature and low-pressure) conditions. More systematic experiments measured IR spectra and optical constants for various pure and mixed ices, while showing how temperature, dilution, and mixture composition can alter the IR features of these ices (e.g., Hagen et al. 1983; Hudgins et al. 1993). In parallel, band strengths of common interstellar ices were revisited and calibrated to improve the accuracy of ice column density determinations from observations (e.g., Gerakines et al. 1995; Mastrapa et al. 2009; Bouilloud et al. 2015). As IR spectroscopy observations developed (especially with the advent of space-based telescopes, see Sect. 1.3.4.3), ice experiments were also advanced to include warm-up and energetic irradiation, in purpose of studying the evolution of interstellar ice analogs and characterizing the processing products such as ions and more complex species (e.g., Allamandola et al. 1988; Ehrenfreund et al. 1997; Öberg et al. 2009b).

Before the launch of JWST, the laboratory group at Leiden conducted systematic IR spectral measurements for various COM ices under different mixing conditions and temperatures. The selected COMs are those that are abundantly detected in the gas phase, including CH_3CHO, C_2H_5OH, CH_3OCH_3, CH_3OCHO, CH_3COCH_3, CH_3CN, CH_3NH_2, and NH_2CHO (Terwisscha van Scheltinga et al. 2018, 2021; Rachid

et al. 2020, 2021, 2022; Slavicinska et al. 2023), of which the IR spectra are publicly available in the Leiden Ice Database for Astrochemistry (LIDA[11]; Rocha et al. 2022). IR spectra of other COMs (pure or mixed with H_2O) such as CH_2OHCHO, $(CH_2OH)_2$, CH_3COOH, C_3H_7OH and hydrocarbons with ≥ 6 atoms are provided by the Cosmic Ice Laboratory of NASA (Hudson et al. 2014; Hudson & Gerakines 2019). These laboratory measurements provide an indispensable toolkit for observers to identify interstellar ices, derive reliable column densities, and infer their thermal and chemical processing histories from observations.

1.3.4.3 Observations in the pre-JWST era

It is not trivial to observe at IR wavelengths due to the absorption and contamination of the Earth's atmosphere. IR facilities have to be built on high mountains or carried on aircraft to minimize the influence of water vapors and thermal emission of the atmosphere itself. The ideal way is to observe from space, which was finally realized in the 1980s by IRAS (Neugebauer et al. 1984).

Ice mantles were suggested to be present on dust grains about 80 years ago (van de Hulst 1946; Oort & van de Hulst 1946), but early attempts of searching for the 3.1 μm H_2O ice band turned out vain toward the diffuse ISM (Danielson et al. 1965; Knacke et al. 1969). It was not until the 1970s that the first detection of this band was made toward embedded protostars in the Orion BN-KL complex (Gillett & Forrest 1973), followed by more detections toward YSOs associated with dense molecular clouds (Merrill et al. 1976; Willner et al. 1982). In addition to H_2O, other key ice species were identified subsequently via their strong vibrational bands, including CO at 4.67 μm (Soifer et al. 1979; Lacy et al. 1984), XCN (OCN^-) at 4.62 μm (Geballe 1986; Grim & Greenberg 1987), NH_3 at 2.97 μm (Knacke et al. 1982; Leger et al. 1983), CO_2 at 15.2 μm (D'Hendecourt & Jourdain de Muizon 1989) and 4.27 μm (de Graauw et al. 1996; more convincing multi-band detections), CH_3OH at 3.53 μm (Grim et al. 1991), OCS at 4.9 μm (Palumbo et al. 1995), H_2CO at 3.47 μm (Schutte et al. 1996), and CH_4 at 7.67 μm (Boogert et al. 1996).

These early detections were made toward massive protostars and mostly in the near-IR with ground-based and airborne IR telescopes, except that mid-IR detections of CO_2 and CH_4 were realized by the ISO (Kessler et al. 1996). The Short Wavelength Spectrometer (SWS) of ISO provided continuous observations between 2.4 and 45 μm band at moderate spectral resolution ($\lambda/\Delta\lambda$ ∼1500–2000), delivering the first near-complete inventories of common ices and the complex absorption structure over the entire 2.5–20 μm range (e.g., Gibb et al. 2004). In parallel, near-IR spectrometers with higher spectral resolution ($\lambda/\Delta\lambda$ up to a few 10^4) started to serve on ground-based optical telescopes such as the Very Large Telescope (VLT), the Keck Observatory (McLean et al. 1998), and the Subaru 8.2 m Telescope. These observations, in synergy with models and experiments, gradually established an analytical framework that links band shapes to thermal and chemical processing of interstellar ices (e.g., Pontoppidan et al. 2003; van Broekhuizen et al. 2005).

As time moved to the 21st century, the launch of the *Spitzer* Space Telescope was a major step toward large-sample IR studies of YSOs, mostly owing to its high sensitivity enabled by the cryogenic environment. In particular, the Cores to Disks (c2d) Legacy

[11] https://icedb.strw.leidenuniv.nl

Program investigated more than one thousand YSOs in five nearby low-mass star-forming regions, with high-SNR spectra obtained for hundreds of YSOs between 5 and 38 µm (Evans et al. 2003, 2009). The c2d survey built the first large, homogeneous ice census toward low-mass YSOs, providing valuable statistics for common interstellar ices (H_2O, CO, CO_2, NH_3, CH_4, CH_3OH) and the 5–8 µm complex (Boogert et al. 2008; Pontoppidan et al. 2008; Öberg et al. 2008; Bottinelli et al. 2010; Öberg et al. 2011).

IR observations of the universe were continued with ground-based and airborne facilities; new space telescopes such as AKARI (Murakami et al. 2007), the *Herschel* Space Observatory (Pilbratt et al. 2010), and the Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010) were launched, but they did not provide higher-quality mid-IR spectroscopic observations after the decommissioning of *Spitzer*. The explored ice inventory thus did not extend far beyond simple species, with CH_3OH being the only COM that had been firmly detected in the solid phase before the advent of JWST. Nevertheless, substantial achievements have been made in understanding the icy universe in the pre-JWST era (see review by Boogert et al. 2015).

1.3.4.4 Findings and challenges of JWST observations

The launch of JWST on the Christmas Day of 2021 marked the beginning of a renaissance of IR astronomy. JWST is a 6.5-meter, cryogenic space telescope orbiting the Sun at the second Lagrange point (L2). With comprehensive improvements in sensitivity (μJy–mJy), spatial resolution (sub-arcsec), and resolving power (R up to 4000, wavelength-dependent), JWST offers unprecedented high-quality observations between 0.6 and 28 µm. In particular, the Near Infrared Spectrograph (NIRSpec) and the Medium Resolution Spectroscopy (MRS) mode of the Mid-Infrared Instrument (MIRI) provide integrated field unit (IFU) spectroscopic observations at sub-arcsec spatial resolutions and moderate resolving powers (1000–4000). In comparison, the resolving power of the InfraRed Spectrograph (IRS) of *Spitzer* is several times to order-of-magnitude lower (R ∼60–120 for ∼5–40 µm and $R = 600$ for ∼9.9–37 µm).

The advanced capabilities of JWST's mid-IR spectroscopy open a new window on the study of interstellar ices (see Fig. 1.4 for a typical IR spectrum of protostellar sources). In addition to the key vibrational bands of common ices (H_2O, CO, CO_2, NH_3, CH_4, CH_3OH, etc.), JWST is also capable to detect weak features of minor isotopologs of simple species and COMs larger than CH_3OH in the solid phase. The Early Release Science (ERS) program Ice Age toward background stars behind the Chamaeleon I cloud reveals a rich inventory of interstellar ices (including ^{13}C isotopologs and ions), providing an elemental budget along prestellar sightlines for the first time (McClure et al. 2023). The decent sensitivity and spectral resolution of JWST also enable systematic measurements of solid-phase $^{12}C/^{13}C$ ratios in solar-mass protostars (Brunken et al. 2024c) and secure detection of HDO ices in multiple YSOs (Slavicinska et al. 2024, 2025b).

With JWST, the fingerprint range of COM ices (6.8–8.8 µm) can be observed in greater detail. The COMs ORigin Investigated by the Next-generation Observatory in Space (CORINOS) program detects the 7.24 and 7.4 µm bands with high SNR toward the Class 0 source IRAS 15398-3359 (Yang et al. 2022). These two bands were originally noticed in the spectrum of a high-mass YSO (W33A) and studied by Schutte et al. (1999), who attributed them to $HCOO^-$, HCOOH, and CH_3CHO, but

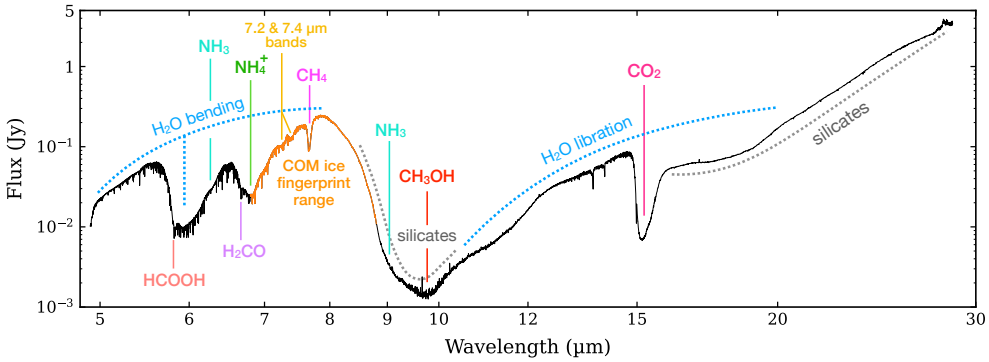


Figure 1.4: JWST/MIRI-MRS spectrum of B1-c, a low-mass protostar studied in Chapter 4. The major ice bands and the COM ice fingerprint range are labeled in different colors.

without claiming a firm detection. Yang et al. (2022) revisited the attribution of the aforementioned species plus C_2H_5OH and CH_3OCHO by comparing the peak wavelengths of the absorption bands at ∼7.24, 7.4, and 11 μm between observations and laboratory spectra. However, no firm identification was claimed until dedicated fitting was performed by Rocha et al. (2024) for one low-mass and one high-mass protostars (NGC 1333 IRAS 2A and IRAS 23385+6053), where laboratory spectra of candidate ices including CH_4, $HCOO^-$, HCOOH, and several two-carbon O-COMs were fitted to the observed COM ice fingerprint range. In *Chapter 4*, similar fitting is performed for another low-mass protostar B1-c; although with different fitting strategies, the identification of O-COMs larger than CH_3OH remains robust in both studies.

The exciting fact that quantitative analysis of COM ices is now available with JWST comes with challenges and caveats. One of the challenges lies in the isolation of weak COM ice bands from the entire spectrum. In Rocha et al. (2024) and Chapter 4, this process is divided into several steps, of which the performance is subjective and case-dependent. Although Chapter 5 tries to optimize and standardize these steps for a larger sample, it is impossible to rule out subjectivity completely. As a consequence, the isolated COM fingerprint range can be different when analyzed by different people/groups, and these differences may be propagated and amplified in the results and conclusions. Therefore, it is important for COM ice studies to elaborate on the methodology employed and provide conservative uncertainties for key quantities such as optical depths and column densities. Another challenge concerns the quantitative fitting of COM ices using laboratory spectra. With current laboratory databases, although significant development has been made in the past decade (Sect. 1.3.4.2), it is still difficult to obtain unambiguous decompositions for observed absorption features, especially those in the COM ice fingerprint range, because of the intrinsic complexity and degeneracy lying in the vibrational bands of mixed ices (Sect. 1.3.4.1). A specific example is given in Chapter 4, where different laboratory spectra are compared to observations, showing the non-uniqueness of COM ice fitting; that is, plausible fitting can be achieved by laboratory spectra with a range of mixing conditions and temperatures. Overall, the quantitative study of COM ices is still an emerging area in astrochemistry that requires further development and investigation.

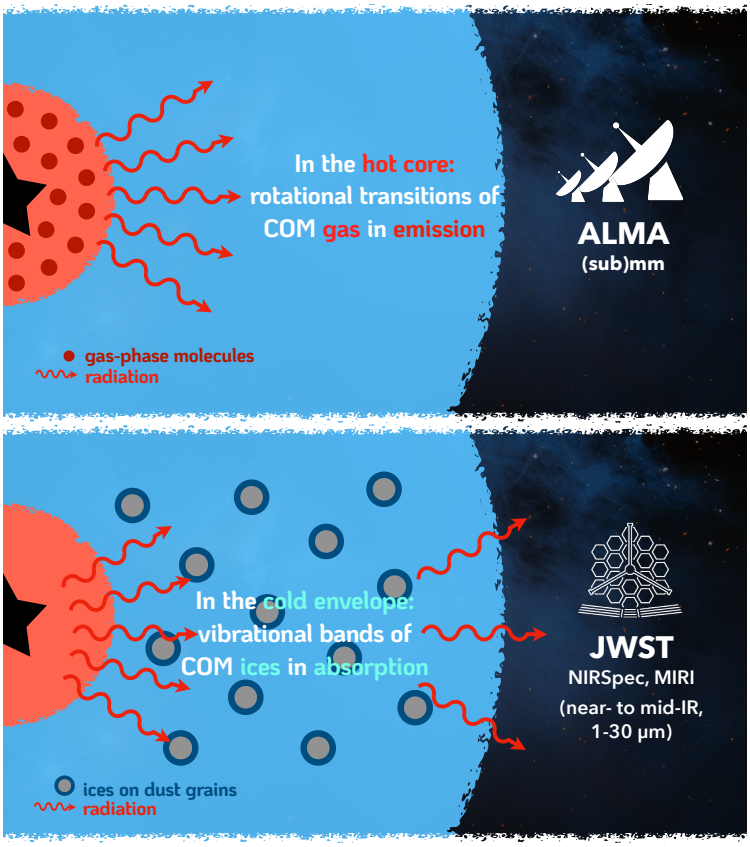


Figure 1.5: Cartoon of ALMA and JWST tracing the gas and ice reservoirs of COMs.

1.4 This thesis

This thesis presents a series of observational studies of COMs in the gas and solid phases using ALMA and JWST, respectively. The highlight as well as the ultimate goal is to make direct comparisons between the gas and ice reservoirs of COMs, which have only become feasible with the combination of ALMA and JWST observations (Fig. 1.5). Thanks to the successful operation of JWST, we are now able to dive deeper into the icy world of interstellar molecules and link the origin of COM formation with the subsequent gas-phase evolution during early stages of star formation.

1.4.1 Overview of chapters

This thesis consists of equal numbers of chapters relevant to ALMA and JWST, well reflecting its topic of combining these two powerful facilities in COM studies. The first two chapters (*2 and 3*) present systematic studies of gas-phase O-COMs in 12–14 high-mass hot cores observed by the CoCCoA program (PI: Brett A. McGuire). The following two chapters (*4 and 5*) shift the focus to searching for COM ice fingerprints in the MIRI-MRS spectra of (mostly) low-mass protostars in the JWST Observations of

Young protoStar (JOYS) program (PI: Ewine van Dishoeck) and several collaborating programs. In particular, *Chapter 4* directly engages with the topic of this thesis by making the first gas-to-ice comparisons of COMs in the *same* sources. The summary of each chapter is provided below:

Chapter 2 investigates six two-carbon O-COMs: acetaldehyde (CH_3CHO), ethanol (C_2H_5OH), dimethyl ether (CH_3OCH_3), methyl formate (CH_3OCHO), glycolaldehyde (CH_2OHCHO), and ethylene glycol ($(CH_2OH)_2$) in the gas phase toward 14 high-mass hot cores observed by the CoCCoA program, motivated by the scarcity of ALMA-based large-sample studies of O-COMs in high-mass protostars. The column density ratios between the six O-COMs and methanol (CH_3OH), the simplest and most abundant COM, are measured and compared to literature values of five low-mass protostars. No clear trend of O-COM/CH_3OH ratios as a function of stellar luminosity is observed between the low- and high-mass samples, suggesting common conditions of COM formation which is most likely to occur on icy dust grains in the cold prestellar phase. The observed gas-phase reservoirs then originate from the following ice sublimation as the protostar warms up the envelope. Among the studied O-COMs, CH_3OCH_3 and CH_3OCHO exhibit the highest and most consistent ratios with respect to CH_3OH. In contrast, the other four O-COMs display greater scatters in their ratios (typically around one order of magnitude), which may result from gas-phase chemistry or differences in emitting areas. Comparisons with astrochemical experiments and simulations show that CH_3OCH_3 and CH_3OCHO tend to be underproduced by experiments, whereas C_2H_5OH and CH_2OHCHO are overproduced by models.

Chapter 4 extends the systematic observational study from two-carbon O-COMs to two three-carbon ones, acetone (CH_3COCH_3) and propanal (C_2H_5CHO), which lack large-sample measurements of their gas-phase abundances. Physical properties of CH_3COCH_3, C_2H_5CHO, and two potentially chemically relevant species, ketene (CH_2CO) and propyne (CH_3CCH), are measured toward 12 out of the 14 CoCCoA sources studied in Chapter 2. Morphologies and kinematics indicate that CH_3COCH_3, C_2H_5CHO, and CH_2CO arise from compact hot cores like other O-COMs, whereas CH_3CCH is more likely to trace extended outflows. The statistics of gas-phase column density ratios between eight O-COMs (six studied in Chapter 2 and two in this chapter) and CH_3OH show that CH_3COCH_3 has quite high gas-phase abundances comparable to some two-carbon O-COMs, whereas aldehydes (those with –CHO group) are systematically less abundant than alcohols, ethers, esters, and ketones with the same amount of carbon atoms. Possible formation routes of acetone are searched in the literature and examined by comparing the observed column density ratios between CH_3COCH_3 and its chemical neighbors with modeling results. All discussed pathways are plausible from an observational point of view, but more evidence (e.g., from theoretical calculations and experiments) is needed to determine which pathway(s) is more likely to dominate.

Chapter 4 presents the first gas-to-ice comparative study of COMs in the same sources—two chemically-rich low-mass protostars, NGC 1333 IRAS 2A and B1-c. Column densities and temperatures of O-COMs in the gas and solid phases are measured from ALMA and JWST/MIRI-MRS spectra, respectively. Dedicated decomposition of the COM ice fingerprint range (6.8–8.8 μm) in the JWST spectrum of B1-c with laboratory spectra of ice mixtures suggests that most O-COM ices are likely present

in a H_2O- or CH_3OH-rich environment. Temperature, however, is difficult to distinguish, as COM ices usually have similar band profiles at low temperatures below their crystallization points. Four O-COMs were selected to make gas-to-ice comparisons in their column density ratios with respect to CH_3OH, and two cases are observed: CH_3OCH_3 and CH_3OCHO have similar ratios in both phases, while CH_3CHO and C_2H_5OH have higher ice ratios by 1–2 orders of magnitude. This implies that both inheritance and gas-phase reprocessing can play a role during the evolution of COMs.

Chapter 5 expands the investigation of ice features from the case study of two low-mass protostars (*Chapter 4*) to a systematic study of more than 20 (mostly low-mass) sources. An interactive workflow for isolating the COM ice fingerprint range is designed for large-sample analysis. As a preparatory step for isolating the weak COM ice bands, broad absorption bands of silicates (at ∼9.8 and 18 µm) and H_2O (at ∼13 µm) are fit. A strong correlation is found between the silicate optical depth and the H_2O ice column densities, confirming their roles of probing dust and ice reservoirs in protostellar envelopes. Regarding the COM ice fingerprint range, absorption bands at ∼7.02, 7.24, 7.40, and 7.67 µm (CH_4 band) are detected in most of the Class 0 and 0/I borderline sources; Class I sources generally have fewer dust and ices probed. Band properties including peak wavelength, FWHM, and integrated area of the four aforementioned bands are measured and compared with the same values of laboratory spectra of attributable species. Comparisons show that the co-occurrence of the 7.24 and the 7.42 µm bands can reasonably be attributed to $HCOO^-$ mixed with H_2O, but other COMs such as C_2H_5OH and CH_3CHO are also likely to contribute, especially to the 7.24 µm band. The 7.67 µm band is solely contributed by CH_4 ice mixed with H_2O and possibly other common species like CO_2 and CH_3OH. This first systematic study of the COM ice fingerprint range provides more observational evidence for solid-phase formation of salts and COMs during early stages of star formation.

Taken together, the four chapters presented in this thesis aim to provide a coherent picture of the origin and early evolution of COMs in star-forming regions. By combining systematic observations of gas-phase COMs in both low- and high-mass protostars with the first direct comparisons between gas and ice in the same sources, this thesis not only establishes an observational framework that links the gas and ice reservoirs of COMs, but also deepens our understanding of their origins in the cold prestellar stage and their subsequent release and reprocessing during protostellar evolution.

1.4.2 Outlook

Looking ahead, the study of COMs in both gas and ice will be continuously advanced by the high-quality, large-sample data obtained from ALMA and JWST observations. There are several approved but yet unpublished programs that focus specifically on COM chemistry in protostars. The ALMA Cycle 9 Large Program Complex Organic Molecules in Protostars with ALMA Spectral Surveys (COMPASS; PI: Jes Jørgensen) will extend the 30-GHz-wide line survey of IRAS 16293 to a sample of 11 low-mass protostars in different star-forming regions and at different evolutionary stages. IR spectra of the same sample will also be obtained by a complementary JWST program. The Cycle 3 JWST Large Program High angular resolution observations of stellar

Emergence in Filamentary Environments (HEFE; PI: Tom Megeath) will deliver IFU spectra in the full 2.9–28 µm range for 13 Class 0 protostars in the OMC2/3 region. Recently, a JWST Cycle 4 program proposed by the author has been approved, in which 15 high-mass protostars with rich detections of gas-phase COM will be observed by MIRI-MRS in search of COM ice fingerprints. Combined with the existing ALMA data for the same sources, the gas-to-ice comparisons presented in Chapter 4 will hopefully be extended to a larger sample and a different mass regime.

As observers look farther and deeper into the universe, more sophisticated models and experiments have also been developed by modelers and experimentalists. The evaluation by models and experiments not only deepens the interpretation of existing data, but also provides guidance for new observations. In particular, the laboratory-measured mm and IR spectra are essential for spectral fitting and directly affect the observational results. Although this thesis focuses exclusively on observational studies, a substantial part of the discussion is devoted to the comparisons with simulation and experimental results. The collaborations among observers, modelers, and experimentalists set the foundation of astrochemistry, and will only become tighter in the future.

1

2 CoCCoA: Complex Chemistry in hot Cores with ALMA

Selected oxygen-bearing species

Y. Chen, M. L. van Gelder, P. Nazari, C. L. Brogan, E. F. van Dishoeck, H. Linnartz, J. K. Jørgensen, T. R. Hunter, O. H. Wilkins, G. A. Blake, P. Caselli, K.-J. Chuang, C. Codella, I. Cooke, M. N. Drozdovskaya, R. T. Garrod, S. Ioppolo, M. Jin, B. M. Kulterer, N. F. W. Ligterink, A. Lipnicky, R. Loomis, M. G. Rachid, S. Spezzano, and B. A. McGuire

Astronomy & Astrophysics, *678*, A137 (2023)

Abstract

Context. Complex organic molecules (COMs), especially the oxygen-bearing species, have been observed to be abundant in the gas phase toward low-mass and high-mass protostars. Deep line surveys have been carried out only for a limited number of well-known star-forming regions using the Atacama Large Millimeter/submillimeter Array (ALMA), which has unprecedented resolution and sensitivity, and statistical studies on oxygen-bearing COMs (O-COMs) in a large sample of high-mass protostars using ALMA are still lacking.

Aims. We aim to determine the column density ratios of six O-COMs with respect to methanol (CH_3OH) in a sample of 14 high-mass protostellar sources in order to investigate their origin through ice and/or gas-phase chemistry. The selected species are: acetaldehyde (CH_3CHO), ethanol (C_2H_5OH), dimethyl ether (DME; CH_3OCH_3), methyl formate (MF; CH_3OCHO), glycolaldehyde (GA; CH_2OHCHO), and ethylene glycol (EG; $(CH_2OH)_2$).

Methods. We fit the spectra of 14 high-mass sources observed as part of the Complex Chemistry in hot Cores with ALMA (CoCCoA) survey and derived the column densities and excitation temperatures of the six selected O-COMs. We used the minor isotopologue of methanol $CH_3^{18}OH$ to infer the column density of the main isotopologue CH_3OH, of which the lines are generally optically thick. We compared our O-COM ratios with those of five low-mass protostars studied with ALMA from the available literature as well as with the results from experiments and simulations.

Results. Although the CoCCoA sources have different morphologies and brightness in their continuum and methanol emission, the O-COM ratios with respect to methanol have very similar values in the high-mass and low-mass samples. The DME and MF have the highest and most constant ratios within one order of magnitude, while the other four species have lower ratios and exhibit larger scatter by one to two orders of magnitude. The ratio between DME and MF is close to one, which agrees well with previous observational findings. Current simulations and experiments can reproduce most observational trends with a few exceptions. For example, they tend to overestimate the abundance of ethanol and GA with respect to methanol.

Conclusions. The constant column density ratios of the selected O-COMs among the low- and high-mass sources suggest that these species are formed in similar environments during star formation, probably on icy dust grains in the pre-stellar stages. Where deviations are found, hypotheses exist to explain the differences between observations, simulations, and experiments, such as the involvement of gas-phase chemistry and different emitting areas of molecules.

2.1 Introduction

Complex organic molecules (COMs), typically defined as carbon-bearing molecules with at least six atoms (Herbst & van Dishoeck 2009), have been intensively studied over the past several decades due to their importance of linking atoms and simple molecules with prebiotic species (Caselli & Ceccarelli 2012; Jørgensen et al. 2020; Ceccarelli et al. 2023). Up to now, more than 80 COMs have been detected in various environments (McGuire 2022). Nearly or fully saturated COMs containing oxygen and nitrogen atoms have been widely observed in line surveys of protostars with different masses by radio telescopes (e.g., Belloche et al. 2016; Jørgensen et al. 2016; El-Abd et al. 2019; Lee et al. 2019a; Csengeri et al. 2019; Belloche et al. 2020; Bianchi et al. 2020; van Gelder et al. 2020; Ligterink et al. 2020; Mininni et al. 2020; Colzi et al. 2021; Nazari et al. 2021, 2022a; Hsu et al. 2022; Imai et al. 2022; Codella et al. 2022). In particular, oxygen-bearing COMs (O-COMs), including methanol (CH_3OH), acetaldehyde (CH_3CHO), dimethyl ether (DME; CH_3OCH_3), and methyl formate (MF; CH_3OCHO), were first observed in massive star-forming regions more than two decades ago (Cummins et al. 1986; Blake et al. 1987; Schilke et al. 1997). They were also detected in subsequent observations of the low-mass protostars IRAS 16293–2422 (Cazaux et al. 2003) and NGC 1333 IRAS 4A/2A (Bottinelli et al. 2004; Jørgensen et al. 2005). Larger O-COMs, such as glycolaldehyde (GA; CH_2OHCHO) and ethylene glycol (EG; $(CH_2OH)_2$), were first detected in the high-mass star-forming cluster Sgr B2(N) near the Galactic center (Hollis et al. 2000, 2002) and then in other protostellar sources (Beltrán et al. 2009; Maury et al. 2014). Three abundant COMs, CH_3OH, CH_3CN, and DME, were even detected in protoplanetary disks around more evolved young stellar objects (Öberg et al. 2015; Walsh et al. 2016; Brunken et al. 2022).

However, the formation mechanisms of COMs are still under debate. The observed O-COMs were initially thought to exist exclusively in the gas phase in hot ($T \gtrsim 100$ K) environments around protostars (i.e., hot cores and hot corinos), primarily from thermal desorption of ices (Herbst & van Dishoeck 2009; Jørgensen et al. 2020). On the other hand, observations in the past decade started detecting them in cold ($T \sim$10 K) pre-stellar cores, albeit in low abundances (Bacmann et al. 2012; Vastel et al. 2014; Jiménez-Serra et al. 2016b; Soma et al. 2018; Scibelli et al. 2021). These detections indicate that COMs may already be formed on the surfaces of dust grains in the early pre-stellar stages before any ice desorption or subsequent gas-phase chemistry occurs. Simulations and experiments have found that solid-phase methanol can be formed in the ice mantles of dust grains. The formation begins with a series of CO hydrogenation under both energetic (e.g., UV radiation) and non-energetic (e.g., atom addition) conditions (Hiraoka et al. 1998; Shalabiea & Greenberg 1994; Watanabe & Kouchi 2002; Fuchs et al. 2009; Cuppen et al. 2009; Simons et al. 2020). Formation of bigger COMs on cold dust grains is also possible. Experiments by Fedoseev et al. (2015) show that GA and EG can be formed through surface hydrogenation of CO under cold, dense cloud conditions ($T = 12$ K). A follow-up experimental study by Chuang et al. (2016) mixed CO, H_2CO, and CH_3OH ices at $T = 15$ K and produced not only GA and EG but also some MF in the solid phase. However, compared with observational results in the gas phase, MF was underproduced in their experiments, while GA and EG were overproduced. The underproduction of MF was alleviated in Chuang et al. (2017) by introducing UV irradiation, which may photodissociate the

initial ingredients in the mixed ice into reactive free radicals and therefore boost the formation of larger species. Based on these laboratory works, Simons et al. (2020) explored the dependence of the final grain mantle composition on the initial gas-phase composition and the dust temperature by chemical simulations. Their simulations were still able to produce MF, GA, and EG at temperatures as low as 8 K, but MF was again underestimated (without UV). A much higher abundance of MF has been reproduced by Garrod et al. (2022), who introduced a set of non-diffusive mechanisms to their simulations (see detailed discussions in Sect. 2.5). These studies show that the abundances of some COMs, such as MF, may depend on environmental conditions, and a debate still exists whether and to what extent gas-phase chemistry is involved in their formation (Balucani et al. 2015; Ceccarelli et al. 2023).

With the development of new observational techniques, it is now possible to make a more complete inventory of COMs in star-forming regions. The Protostellar Interferometric Line Survey (PILS; Jørgensen et al. 2016) using the Atacama Large Millimeter/submillimeter Array (ALMA) detected more than 20 COMs, including all the O-bearing species mentioned above plus ethanol (C_2H_5OH) and some of their isotopologues in the low-mass protostellar binary IRAS 16293–2422 A and B (Jørgensen et al. 2018; Manigand et al. 2020). Recent ALMA surveys of a larger sample of sources have further confirmed the ubiquity of COMs in both low- and high-mass protostars (van Gelder et al. 2020; Yang et al. 2021; Nazari et al. 2022a). The next step in COM studies is therefore to investigate similarities and differences in COM abundances and their ratios among different types of sources in order to constrain their formation routes. Indeed, surveys have revealed an interesting consistency of O-COM ratios with respect to methanol among different sources. Coletta et al. (2020) summarized single-dish observations of DME and MF in various objects, including pre-stellar cores, star-forming regions, a protostellar shock, and galactic center cores. They found that DME and MF abundances are strongly correlated with a ratio of about one. As another example, van Gelder et al. (2020) report constant ratios of O-COMs with respect to methanol in five low-mass protostars. For N-bearing species, Nazari et al. (2022a) also found rather constant ratios with respect to methyl cyanide (CH_3CN) in more than 30 high-mass sources, though some species, such as formamide (NH_2CHO), show a larger scatter in their ratios. The constant abundance ratios revealed by observations are interesting since different objects have different physical environments, which are expected to influence the chemical evolution and alter the ratios among COMs. These similarities suggest that COMs are mainly formed under similar physical conditions, probably on the surfaces of dust grains during the cold pre-stellar phase.

Previous observational studies mostly focused on a small number of well-known sources, and many of them did not have enough spatial resolution and sensitivity to overcome beam dilution and optical depth issues. So far, a few ALMA surveys with enhanced sensitivities and resolutions on O-COMs in low-mass protostellar regions have been published (van Gelder et al. 2020; Hsu et al. 2022), but results for the high-mass counterparts, especially from ALMA surveys that cover minor isotopologues of methanol, are still lacking. We emphasize the importance of detecting optically thin lines of minor isotopologues of methanol, which requires higher sensitivity, in order to better constrain the column density of methanol and hence the ratios with respect to other O-COMs. There are several ALMA studies on O-COMs in the high-mass counterparts (e.g., Csengeri et al. 2019; El-Abd et al. 2019; Mininni et al. 2020),

but all of them are case studies that do not use minor isotopologues of methanol in their analyses and/or only focus on a small set of O-COMs, such as the $C_2H_4O_2$ isomers. High-mass protostellar regions are important to the investigation of COM chemistry under different physical conditions, and given the higher temperatures and the possibly enhanced UV radiation along outflow cavity walls, it is expected that the gas-phase data from high-mass sources will give a better representation of the solid-phase abundances, as icy COMs are expected to be mostly thermally desorbed from dust grains. It is also timely to present the results of ALMA observations on high-mass sources after the latest development of a comprehensive model on COM chemistry in hot cores by Garrod et al. (2022).

In this paper, we present the analysis of O-COMs observed in 14 high-mass star-forming regions in the Complex Chemistry in hot Cores with ALMA (CoCCoA) survey. The analysis is focused on six O-bearing COMs, C_2H_5OH, CH_3CHO, DME, MF, GA, and EG (including two conformers, anti and gauche, *a*-EG and *g*-EG), as well as two methanol isotopologues, $^{13}CH_3OH$ and $CH_3^{18}OH$. Covering minor isotopologues of methanol is essential because the lines of the main isotopologue, $^{12}CH_3^{16}OH$, are likely to suffer from a high optical depth, in which case its column density needs to be inferred from optically thin lines of a minor isotopologue. To this sample, we added C_2H_5OH and CH_3CHO for their potential detections in ices (Schutte et al. 1999; Yang et al. 2022; McClure et al. 2023). We selected MF, GA, and EG in order to study their relative abundances, based on laboratory and modeling findings in Chuang et al. (2017) and Simons et al. (2020). Finally, we included DME because of its high abundance observed in star-forming regions.

This work will provide a base for comparison with future observations of icy COMs by the *James Webb* Space Telescope (JWST). To verify if COMs are formed in ices, a direct approach is to observe their vibrational absorption features in the infrared. However, this requires very high sensitivity and spectral resolution in the fingerprint wavelength range at $\sim$2–15 μm, which has only become feasible with the successful operation of JWST. So far, methanol is the only COM that has confirmed detections in interstellar ices, whereas several other solid-phase COMs, such as C_2H_5OH and CH_3CHO, have only been tentatively identified (see review by Boogert et al. 2015, and references therein). Recently, several JWST teams have started hunting for solid-phase COM features (Yang et al. 2022; McClure et al. 2023). Important laboratory measurements of COM infrared spectra are now available for JWST data analyses (Terwisscha van Scheltinga et al. 2018, 2021; Rachid et al. 2020, 2021, 2022; Hudson & Gerakines 2019; Hudson & Ferrante 2020; Hudson et al. 2021; Hudson & Yarnall 2022; Gerakines et al. 2022; Rocha et al. 2022). The ultimate goal is to bridge the gas and grain chemistries and directly relate gas-phase spectra as presented here to solid-phase infrared features, from which we can gain a better understanding about the formation history of COMs.

2.2 Observations

2.2.1 Overview

This paper presents early results from a first look at data from the CoCCoA survey (PI: B. A. McGuire). A complete overview of the survey will be provided in a separate

publication once all observations are complete. This will be followed by the public release of fully reduced spectral-line image cubes. When complete, CoCCoA will comprise a dataset targeting 23 high-mass star-forming regions at a spatial resolution of $\sim$0.3″ using multiple configurations of the ALMA 12-m array as well as the ALMA 7-m array to capture both extended and compact emission. The CoCCoA survey will ultimately include data from ALMA projects 2019.1.00246.S, 2019.2.00112.S, and 2022.1.00499.S.

The sources were chosen to be within $\sim$1–4 kpc, corresponding to a linear-scaled spatial resolution of $\sim$300 to 1200 au at the nominal 0.3″ angular resolution. Other than this distance requirement, the only other criterion used for selection was that the sources have a prior literature report of "hot core chemistry," either described as such or evidenced by the presence of a rich array of emission lines of methanol and other COMs. The goal is to provide a diverse dataset in order to better sample the phase space of chemical complexity and avoid biases toward the chemistry seen in the commonly observed extraordinary sources such as Sgr B2 (see, e.g., El-Abd et al. 2019).

The observations cover two spectral tunings per source in Band 6. The lower tuning covers 238.0–241.7 GHz, while the upper tuning covers 258.0–261.7 GHz. All observations are taken with 0.488 MHz spectral resolution (twice the channel width due to online Hanning smoothing), corresponding to $\sim$0.6 km s^{-1}. These frequency ranges were chosen to maximize both the number of transitions and the range of upper-state rotational energy levels of key target molecules, including those highlighted here. These first-look results use data from the upper tuning (258.0–261.7 GHz), for which observations have been completed for 14 total sources, listed in Table 2.1.

2.2.2 Data reduction

The 12m array CoCCoA data presented here from project 2019.1.00246.S were observed between March 24, 2021, and April 2, 2021, using configuration C-5. The data were calibrated using the Cycle 8 version of the ALMA Pipeline (CASA version 6.2), including corrections for renormalization issues larger than 2%. Due to the copious line emission emanating from the massive protostars within each targeted cluster, the default parameters of the `findContinuum.py` procedure used by the pipeline to identify line-free channels did not yield optimal results (i.e., significant line contamination was present). Therefore, `findContinuum.py` (available in the `extern` directory of the pipeline distribution) was run manually with a few key parameters (primarily `sigmaFindContinuum`[1]) adjusted in order to minimize the line contamination.

After optimization of the channels used for continuum subtraction, the line-free channels were used to image the continuum and perform an iterative self-calibration (Brogan et al. 2018). The self-calibration solutions were also applied to the continuum-subtracted line datasets. The imaging of both the continuum and data cubes employed the `multiscale` functionality of CASA's `tclean` with scales of 0, 5, and 15 (the `cellsize` was chosen to oversample the beam on the smallest axis by a factor of $\sim$5). The `robust` parameter employed is 0.5, which yielded an angular resolution near 0.3″;

[1]The `sigmaFindContinuum` parameter sets how far above the corrected median of the baseline channels to place the initial threshold (the baseline channels, `nBaselineChannels`, were drawn from the lowest 10% of channels after excluding outliers). The `narrow` parameter was also set to two, instead of the default of four, in order to preserve as many narrow windows as possible.

Table 2.1: Names, phase center coordinates, coordinates where spectra were extracted, and properties of sources analyzed.

Source Name‡	Phase Center† R.A.	Phase Center† Dec.	Extraction Location† R.A.	Extraction Location† Dec.	L (10^4 $L_\odot$)	D (kpc)	D_{GC} (kpc)	$^{16}O/^{18}O$	Refs.
G19.01-0.03	18:25:44.80	–12:22:45.8	18:25:44.78	–12:22:45.95	1	4.0(3)	4.4	298±52	L: 1, D: 2
G19.88-0.53	18:29:14.57	–11:50:23.0	18:29:14.36	–11:50:22.50	0.47	3.31	5.1	335±60	L: 3, D: 4
G22.04+0.22	18:30:34.70	–09:34:47.0	18:30:34.70	–09:34:47.00	0.497(26)	3.4(5)	5.0	333±60	L: 5, D: 5
G23.21-0.37	18:34:55.26	–08:49:15.3	18:34:55.20	–08:49:14.70	1.3	4.6	4.2	286±50	L: 6, D: 6
G34.30+0.20	18:53:18.54	+01:14:57.9	18:53:18.62	+01:14:58.35	4.6	1.6	6.8	436±80	L: 7, D: 7
G34.41+0.24	18:53:17.90	+01:25:25.0	18:53:18.02	+01:25:25.15	0.48	1.6	6.8	436±80	L: 8, D: 8
G35.03+0.35	18:54:00.50	+02:01:18.0	18:54:00.66	+02:01:19.30	0.63	2.32	6.3	407±74	L: 9, D: 10
G35.20-0.74N	18:58:13.00	+01:40:36.5	18:58:12.96	+01:40:37.35	3	2.2(2)	6.4	411±75	L: 11, D: 12
G345.5+1.5	16:59:41.63	–40:03:43.6	16:59:41.59	–40:03:42.90	4.8	1.5	6.6	426±78	L: 13, D: 13
IRAS 18151-1208	18:17:58.00	–12:07:27.0	18:17:58.22	–12:07:24.90	2.2	2.9	5.4	353±63	L: 14, D: 14
IRAS 16547-4247	16:58:17.20	–42:52:07.0	16:58:17.22	–42:52:07.35	6.3	2.9	5.3	353±63	L: 13, D: 13
NGC 6334-38	17:20:18.00	–35:54:55.0	17:20:17.80	–35:54:43.85	<20	1.7	6.4	412±75	L: 15, D: 15
NGC 6334-43	17:20:23.00	–35:54:55.0	17:20:23.86	–35:54:56.90	<20	1.7	6.4	412±75	L: 15, D: 15
NGC 6334 I(N)-SM2	17:20:55.00	–35:45:40.0	17:20:55.64	–35:45:32.60	0.07	1.3(1)	6.8	434±79	L: 16, D: 17

†Coordinates given in J2000; Right Ascension (R.A) in units of °:′:″; Declination (Dec.) in units of hh:mm:ss.
‡Other names exist for many of these, including different designations for the larger star-forming complex and for individual sub-sources within the complex.

References – [1] Cyganowski et al. 2011, [2] Williams et al. 2022, [3] Issac et al. 2020, [4] Ge et al. 2014, [5] Towner et al. 2021, [6] Tang et al. 2018, [7] Csengeri et al. 2022, [8] König et al. 2017, [9] Beltrán et al. 2014, [10] Wu et al. 2014, [11] Sánchez-Monge et al. 2013, [12] Zhang et al. 2009, [13] Faúndez et al. 2004, [14] Maud et al. 2015, [15] Widicus Weaver et al. 2017, [16] Sandell 2000, [17] Chibueze et al. 2014.

both the continuum and line images were subsequently convolved to exactly 0.3″. The maximum recoverable scale of these data is ∼2.5″.

Considering that the data in the lower tuning are still under reduction, only the upper tuning was used for further analysis. The quality of the ALMA pipeline data of the lower tuning is good enough to see how many COM lines are included in each tuning. Among the 14 sources, we took G19.88-0.53 as an example for line identification, since its spectrum shows less line blending despite the large line intensities. This ensured a sufficiently large number of detected lines without suffering too much from blending issues. The upper tuning covers about 70% of the COM lines found in both tunings (not including CH_3OH), while the lower one is more abundant in strong lines of CH_3OH and CH_3CN, which are not directly relevant to our project. Therefore, we consider our results to be robust using only the data in the upper tuning.

2.3 Methods

2.3.1 Spectral analyses

Figure 2.1 shows the continuum maps and the integrated intensity (moment 0) maps of the CH_3OH $19_{3,17}$–$19_{2,18}$ transition with upper energy E_{up} = 490 K and Einstein A coefficient $A_{ij} = 9.27 \times 10^{-5}$ s^{-1}. This particular methanol line was chosen because it is unblended and has the lowest E_{up} and the highest A_{ij} in the upper tuning, which is expected to show the most extended methanol emission. Figure 2.2 presents an overview of the spectra in the upper tuning, with several representative line features of MF and DME indicated by shaded areas. The spectra were extracted from the peak pixel of the methanol intensity maps. The moment 0 maps of other selected O-COMs of one example source G19.88-0.53 are shown in Fig. 2.A.1, which confirmed that their emission peaks at similar regions to the methanol emission. In most sources, the methanol emission peaks at the same location as the continuum. However, in some bright sources, such as G34.30+0.20 (Fig. 2.1), the methanol emission shows a ring shape around the continuum peak, which is likely due to the high optical depth of dust (De Simone et al. 2020; van Gelder et al. 2022). In this case, we picked the brightest pixel on the ring to extract the spectra. For the two bright sources G34.41+0.24 and G23.21-0.37, the methanol emission peaks at the same location as the continuum, but the spectrum extracted from the central pixel has too much line blending, so we deliberately chose a pixel offset from the actual peak (∼9 pixels offset for G34.41 and ∼3 pixels offset for G23.21).

After extracting all the spectra, we performed line identification and spectral fitting using the spectral analysis software CASSIS[2] (Vastel et al. 2015). The spectroscopic data in CASSIS are taken from two databases: the Jet Propulsion Laboratory database (JPL; Pickett et al. 1998) and the Cologne Database for Molecular Spectroscopy (CDMS; Müller et al. 2001, 2005; Endres et al. 2016). Detailed references for each species can be found in Appendix A of van Gelder et al. (2020). Some species only have data available in one database, while others are included in both databases. In the latter case, we used either the database with smaller uncertainties in the central frequency or that has been used more frequently in previous work. We first went

[2] http://cassis.irap.omp.eu/

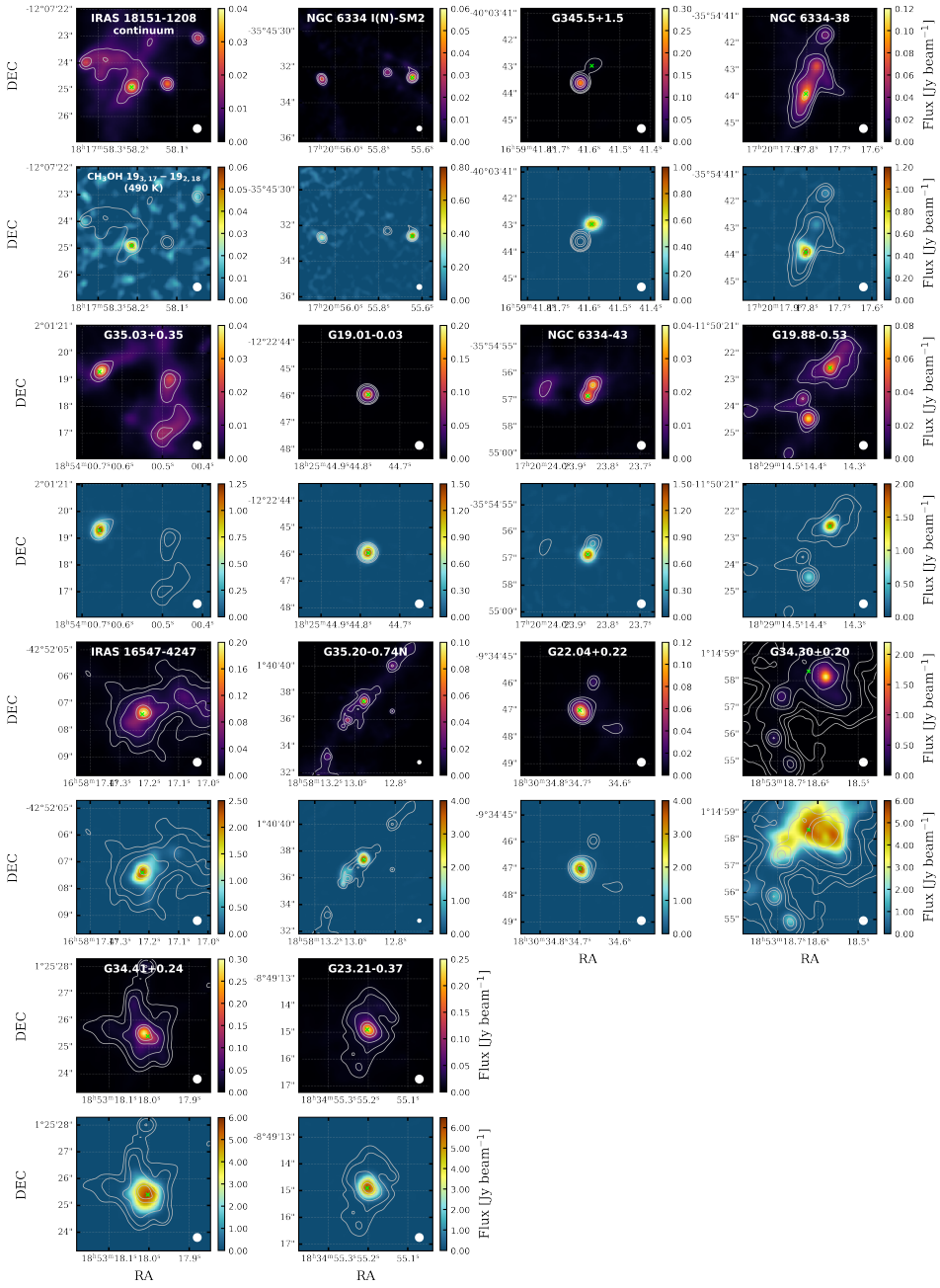


Figure 2.1: Continuum maps (odd rows, black-yellow) and moment 0 maps of the methanol line at 258.78 GHz ($19_{3,17}$–$19_{2,18}$) with $E_{\rm up}$ = 490 K (even rows, blue-red) for the 14 high-mass CoCCoA sources. The white contours indicate the continuum emission at the 3, 5, 10, 30, 50, and 100σ level (σ = 3 mJy beam^{-1}). The peak pixels from where the spectra were extracted are marked by crosses in light green. The sources are sorted by peak flux of methanol emission. The beam size (0.33″) is denoted by the white ellipse in the lower right corner of each panel. The field of view for NGC 6334 I(N)-SM2 and G35.20-0.74 N is set to be twice as large as that of others in order to include more emission peaks.

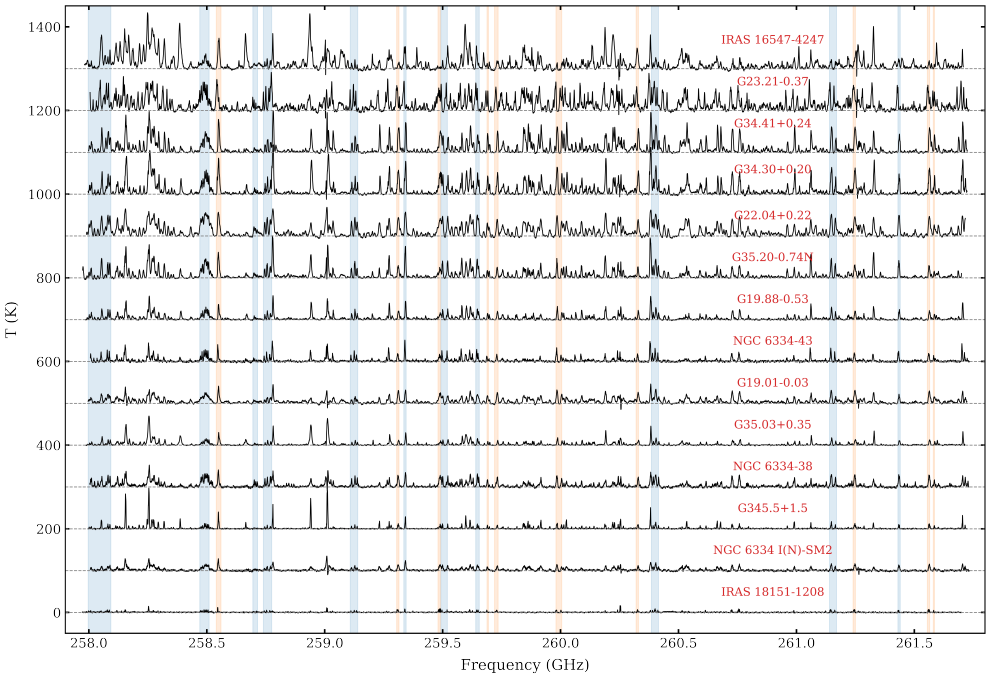


Figure 2.2: An overview of the spectra in the upper tuning (258.0–261.7 GHz) of the 14 high-mass sources observed in the CoCCoA survey. The line intensity has been converted from flux in Jansky per beam to brightness temperature in Kelvin using the Rayleigh-Jean approximation. The spectra are sorted by the average line intensities from bottom to top, and each one is offset by 100 K. The shaded regions in blue and orange indicate the emission lines of MF and DME, respectively.

through the detection inventory of the PILS survey (Jørgensen et al. 2016) and checked whether all the transitions in the databases have corresponding line features in the observed spectra for each species. The detection results are presented in Sect. 2.4.2.

In the next step, we chose the six O-COMs and the two methanol isotopologues ($^{13}CH_3OH$ and $CH_3^{18}OH$) mentioned in Sect. 2.1 for detailed spectral fitting using CASSIS. The two conformers of EG (*a*-EG and *g*-EG) were fit separately due to their different sets of transitions. The column density (N), excitation temperature (T_{ex}), and full width half maximum (FWHM) of each molecule were fit for each source. We assumed one T_{ex} for each species; that is, the populations of all levels can be characterized by a single T_{ex}, which is often called "local thermodynamic equilibrium (LTE)" in radio astronomy. However, the LTE here does not necessarily refer to its strict definition that T_{ex} approaches the kinetic temperature T_{kin} under high-density conditions. In fact, COMs are likely to be subthermally excited ($T_{ex} < T_{kin}$) in hot cores, while the observed lines can still be well characterized by one T_{ex}. Examples provided in Fig. 6 of Johnstone et al. (2003) and Fig. 8 of Jørgensen et al. (2016) show that for the case of methanol, densities of 10^9–10^{10} cm^{-3} are needed for thermalization. Correspondingly, physical models of the envelopes of high-mass protostars (e.g., van der Tak et al. 2013) indicate densities of 10^7–10^9 cm^{-3} at temperatures of 100–300 K

on scales of a few hundred au. At such densities, the level populations of COMs may not yet be fully in LTE, while the fitting results in Sect. 2.4 show that the single-$T_{\rm ex}$ assumption is reasonable and works well.

We adopted two methods for fitting the spectra: χ^2 minimization and visual inspection, when the former was not applicable. For each source, the radial velocity $v_{\rm lsr}$ and FWHM were determined for each species based on strong unblended lines. The uncertainties of $v_{\rm lsr}$ and FWHM are smaller than 0.5 km s^{-1}. The difference of $v_{\rm lsr}$ among the O-COMs is within 1 km s^{-1} in most cases. For sources where the lines are narrow and unblended, grid-fitting was used to determine the best-fit values as well as the uncertainties by calculating χ^2 in the parameter space and making the contour plot on the N–$T_{\rm ex}$ plane (also see Sect. 3.1 in van Gelder et al. 2020). For each species, χ^2 was calculated from the difference between the LTE model and the observed spectrum around unblended lines. We started with a sparse grid with broad ranges of N, $T_{\rm ex}$, and FWHM and large intervals between the grid points. We then gradually narrowed down to smaller ranges and smaller intervals. Finally, we ended up with a fine grid around the best-fit grid point, from which we could make the contour plot and estimate the 2σ uncertainties. This grid-fitting method works well with weaker sources where most of the lines are unblended. However, for bright sources where the lines are very broad and blended, grid-fitting does not converge to a solution and the results need to be visually inspected.

When it came to fitting by visual inspection, we started with an initial guess of the parameters and adjusted them to a better fit until no improvement could be made. This is more efficient and reliable for complex spectra with blended lines since we can monitor the change intuitively and interactively. This method was also adopted by van Gelder et al. (2020) and Nazari et al. (2021, 2022a). Uncertainties were estimated by comparisons shown in Fig. 2.A.2, in which the three panels correspond to the lower limit, the best fit, and the upper limit. Panel (a) shows that with an underestimated $T_{\rm ex}$, low-$E_{\rm up}$ transitions tend to overestimate the observation, while in panel (c), high-$E_{\rm up}$ lines are overestimating the observation with an overestimated $T_{\rm ex}$.

In some cases, grid-fitting did not work well for $T_{\rm ex}$, even when there are enough unblended lines available for a certain species. This is because a robust estimation on $T_{\rm ex}$ requires the unblended lines to cover a wide range of $E_{\rm up}$ so that the models can be sensitive to temperature changes. For species such as $CH_3^{18}OH$, GA, and g-EG, there are few unblended lines covering a wide range of $E_{\rm up}$. These lines are either too weak to be detected in faint sources or severely blended by strong lines in bright sources. Under these circumstances, the results of grid-fitting will not be able to constrain the $T_{\rm ex}$, as it is often accompanied by huge uncertainties. As a solution, we fixed $T_{\rm ex}$ to be equal to that of MF, which has the most identified transitions and therefore the best constraint on $T_{\rm ex}$, and only N was fit and estimated for uncertainties. Nevertheless, the $T_{\rm ex}$ of g-EG is always set to be the same as that of a-EG. We consider the fitting results of N to still be representative since the difference would be within a factor of two if we changed the fixed value of $T_{\rm ex}$ by 20–50 K in the range of 100–250 K, which is a typical temperature range for hot cores (e.g., Fig. 2 in Ligterink et al. 2015). Similarly, if the FWHM could not be constrained to better than 0.5 km s^{-1} uncertainty due to the blending of lines, it was also fixed to a value (e.g., if the possible range is 3.5–4.0 km s^{-1}, we used 3.8 km s^{-1}).

Additionally, there are two special cases that we encountered during the fitting.

One is that the spectra of several sources (i.e., G19.01-0.03, G34.30+0.20, G34.41+0.24, and G35.20-0.74N) show double-peaked line profiles in some O-COMs, which cannot be fit with only one component. These double-peaked features appear in nearly all the transitions and are thus unlikely to be due to self-absorption. An example of two-component fitting is given by Fig. 2.A.3, which shows that the spectra can be well fit by two components with different $v_{\rm lsr}$ and N, whereas $T_{\rm ex}$ remains the same (see Tables 2.B.1–2.B.5 for relevant sources and species). It is likely that these sources are not spatially resolved (e.g., maps of G19.01-0.03 show a perfect beam shape in Fig. 2.1), and there is more than one physical component contained in the beam. The second case is that in the two sources G23.21-0.37 and NGC 6334-38, we found the MF lines to be highly optically thick. All strong lines with $A_{\rm ij} \gtrsim 10^{-4}$ s^{-1} are saturated and therefore do not show correct intensity ratios against weaker lines. In this case, the fitting was only based on weak lines with $A_{\rm ij} < 10^{-4}$ s^{-1}.

2.3.2 Isotope ratio calibration for methanol

As mentioned in Sect. 2.3.1, the column density of CH_3OH needs to be inferred from its minor isotopologues since the main isotopologue itself is likely to be optically thick. The isotope ratios of $^{16}O/^{18}O$ and $^{12}C/^{13}C$ can be calculated from the distance to the Galactic center ($D_{\rm GC}$) using the equations in Wilson & Rood (1994) and Milam et al. (2005):

$$(^{16}O/^{18}O) = (58.8 \pm 11.8) D_{\rm GC} + (37.1 \pm 82.6) \quad (2.1)$$

$$(^{12}C/^{13}C) = (6.21 \pm 1.00) D_{\rm GC} + (18.71 \pm 7.37), \quad (2.2)$$

where $D_{\rm GC}$ can be derived from the coordinates of the sources and their distances to Earth. The two values before and after $D_{\rm GC}$ (with uncertainties) are the slope and intercept of these linear relationships, respectively. In the vicinity of the solar system, $D_{\rm GC}$ is 8.05 kpc, which gives $^{16}O/^{18}O \sim 510$ and $^{12}C/^{13}C \sim 69$. In the literature, 560 and 70 are commonly used for nearby low-mass sources. For the high-mass sources that are farther away from the solar system, isotope ratios can be calculated from $D_{\rm GC}$ before applying them to infer the column density of $^{12}CH_3^{16}OH$ (Table 2.1). If both $^{13}CH_3OH$ and $CH_3^{18}OH$ are optically thin, we can expect their column density ratio to be 7–8. However, our fitting results show a ratio of 2–5, which indicates that $^{13}CH_3OH$ is also (marginally) optically thick. Therefore, only $CH_3^{18}OH$ is used to calculate the column density of methanol.

If the errors of both the slope and the intercept are considered in error propagation, the uncertainties of $^{16}O/^{18}O$ would be around 30% of the ratios themselves (e.g., if $^{16}O/^{18}O = 300$, the error will be ∼90). However, a large portion of the total uncertainty comes from the intercept error (82.6), which is even two times larger than the intercept itself (37.1). Considering that only the slope in Eq. (2.1) contains the information of the trend between the $^{16}O/^{18}O$ ratio and $D_{\rm GC}$ and that the intercept error only represents the scatter of the sources from which the equation was originally fit, we did not include the intercept error in the error propagation. This yielded a decrease in the uncertainties of $^{16}O/^{18}O$ from ∼30% to ∼18%. Results that include the intercept error are shown in Fig. 2.A.4.

2.4 Results

2.4.1 Morphology

The continuum maps of the 14 CoCCoA sources show different morphologies (see odd rows in Fig. 2.1). Six out of the 14 sources show a single peak in both continuum and methanol emission. Others have multiple peaks in continuum but only one or two peaks in methanol emission, usually corresponding to the brightest continuum peak(s). The source G345.5+1.5 is an exception in that its methanol peak is offset from the continuum peak, and the corresponding region in continuum does not show a peak feature. The source G35.20-0.74N is an interesting source, as the four continuum peaks are located along a line and the methanol emission tends to follow the same alignment. Except for several sources that are not well resolved, most sources show extended weak continuum emission aside the flux peaks, implying the existence of dusty envelopes or parent cores. More detailed studies on the morphology of CoCCoA sources will be presented in a future paper.

2.4.2 Spectra and detection

As expected from the selection of sources and frequency range, the extracted spectra are rich in COM lines for all the 14 sources in the upper tuning. An overview of the full spectra of each source is presented in Fig. 2.2. We observed that the spectral appearance of our sources is diverse in intensities and line widths. Bright sources such as G34.41+0.24 and G23.21-0.37 have very strong and broad lines ($\gtrsim$5 km s^{-1}) that are severely blended and can only be fit by visual inspection. There are also faint sources, such as IRAS 18151-1208 and NGC 6334 I(N)-SM2, where the lines are much weaker and narrower (<2 km s^{-1}). Their spectra are clean enough for χ^2 fitting, but some less abundant species and some weak lines (with lower A_{ij}) may remain undetected. The spectra of the other sources have intermediate intensities and line widths (3–4 km s^{-1}), which are easiest to fit.

The 14 sources provide a rich inventory of detections. More than ten O-bearing COMs and about five N-bearing species are detected in the upper tuning. These detections include the originally targeted species as well as other O- and N-bearing species, covering alcohols, aldehydes, esters, ethers, ketones, acids, and the simplest sugar (GA). We did not detect methoxymethanol (CH_3OCH_2OH) since our frequency range only covers its weak lines with $A_{ij} < 10^{-5}$ s^{-1}, and no corresponding line features were spotted in the spectra. Besides the COMs mentioned above, two simple O-bearing molecules, ketene (H_2CCO) and formic acid (t-HCOOH), have one and three transitions covered in the upper tuning, respectively. The N-bearing species include HNCO, CH_3CN, C_2H_5CN, and NH_2CHO. Some abundant species such as HNCO and CH_3CN also have their ^{13}C isotopogues detected. A hydrocarbon molecule, propyne (CH_3CCH), is also detected but with all its transitions covered in the lower tuning. Simpler molecules such as SiO, SO, and SO_2 only have very few strong lines covered in the upper tuning and do not affect our analysis on O-COMs. Species detected in the CoCCoA sources but not the focus of this paper will be studied in future works.

Figure 2.3 shows the best-fit model of G19.88-0.53 in a line-rich frequency range (259.4–260.6 GHz). A version for the full upper tuning is given by Fig. 2.A.5, and Figs. 2.A.6–2.A.8 show zoom-in panels for selected unblended lines of each species.

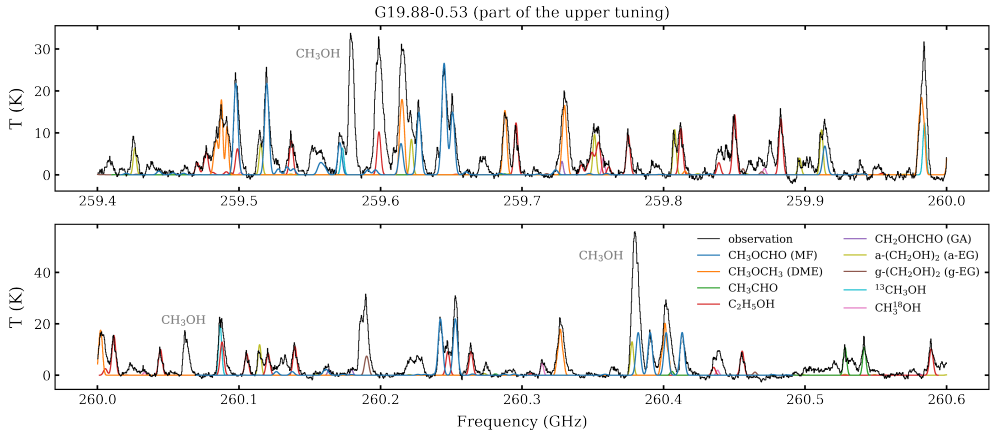


Figure 2.3: Best fit of the observed spectrum of G19.88-0.53 in a line-rich frequency range (259.4–260.6 GHz). The black and colorful spectra correspond to the observed spectrum and the modeled spectra of each species, respectively. Only the O-bearing COMs targeted in this paper (excluding CH_3OH) are shown here. Two strong lines of methanol are labeled in gray text in the panels. The same figure but for the full upper tuning is given in Fig. 2.A.5.

The model contains the minor isotopologues of methanol and the six O-COMs that we focus on in this paper. More than 70% of the line features can be fit quite well with uncertainties of less than 30%, assuming a single excitation temperature. Apart from several strong features that are attributed to the main isotopologue of methanol and some simple molecules (e.g., SO and SO_2), weaker features that are not well fit by the model likely originate from other COMs that are not included in the model, such as acetone (CH_3COCH_3), N-bearing COMs, and some minor isotopologues of the detected species. The identified transitions of the selected O-COMs are listed in Table 2.B.6, where transitions that are above 3σ and not fully blended with other strong lines are considered as "identified." We observed that most of the identified transitions have an upper energy level of 100–300 K.

2.4.3 Column density and excitation temperature

The column densities and excitation temperatures of all sources are presented in Tables 2.B.1–2.B.4. Vibrational corrections have been applied to the column densities of GA and EG (more information can be found in Sect. 5 of Jørgensen et al. 2016), and the column densities without vibrational corrections are provided in Table 2.B.5. With the uniform beam size of 0.33″, most sources have a methanol column density of 10^{18}–10^{19} cm^{-2}, except the weakest source IRAS 18151-1208, which has $N = 7 \times 10^{17}$cm^{-2}. Among the six O-COMs that we focus on (excluding methanol isotopologues), DME and MF are the most abundant, with column densities about one order of magnitude higher (10^{16}–10^{17} cm^{-2}) than those of the other four species. These two species always have a considerable number of strong lines available for fitting in the spectra of all sources (as shown by the shaded areas in Fig. 2.2). C_2H_5OH has somewhat fewer distinct transitions (unblended and large A_{ij}), and it is not always abundant enough to produce strong lines. In some sources, the column density of C_2H_5OH is of the

same order of magnitude as DME and MF, while in some other sources the difference can reach up to two orders of magnitude. This big variation is also present for a-EG, which has many strong transitions covered in our data. In some sources, the number of detected lines of a-EG are as high as those of DME, while in others there are only a few obvious line features. The fitting of GA and g-EG is more difficult since they only have two to three distinct transitions, which are often subject to blending issues. In summary, the average column densities of the six O-COMs rank as DME $\sim$ MF > C_2H_5OH > (a+g)EG > CH_3CHO > GA. Since the absolute values of N are related to the physical environments of the parent sources, it is more useful to look at the relative abundances of O-COMs, that is, their column density ratios with respect to methanol (see Sect. 2.5 for details).

As for the excitation temperature, C_2H_5OH, MF, and a-EG tend to have a warm T_{ex} of $\gtrsim$150 K, while CH_3CHO and DME have a relatively lower T_{ex} of 100–130 K. This may be because different species have a different emitting area of the hot core. For example, some species are emitting from a slightly colder and more extended region, which is not well resolved in our sample according to Fig. 2.A.1). Hot cores are known to have temperature gradients (e.g., van der Tak et al. 1999, 2000; Beltrán et al. 2018; Gieser et al. 2019), but they are only probed on scales larger than the observing beams. Since our analyses are based on the spectra at the peak pixels, the temperature structure on larger scales is not expected to affect our results. The lower T_{ex} of CH_3CHO and DME is also consistent with earlier single-dish findings of other high-mass sources (Bisschop et al. 2007; Isokoski et al. 2013) and the results of low-mass sources in van Gelder et al. (2020). As mentioned in Sect. 2.3.1, $CH_3^{18}OH$, GA, and g-EG usually have fixed T_{ex} due to a lack of unblended lines covering a wide range of E_{up}.

2.5 Discussion

2.5.1 Column density ratios

To further investigate the COM chemistry in our sample, we calculated the column density ratios of O-COMs with respect to methanol, and we summarize these column density ratios in Fig. 2.4. The column densities of the two conformers of EG (a-EG and g-EG) are summed up in the last subplot, although they were fit separately. The uncertainty of the column density ratios was propagated from the uncertainties of the O-COM column densities and the $^{16}O/^{18}O$ ratio (calculated in Sect. 2.3.2) of each source. In Fig. 2.4 we also compare our observational results of high-mass protostellar sources with the results of low-mass sources in literature (Jørgensen et al. 2018; Manigand et al. 2020; van Gelder et al. 2020) and the simulation results of COM chemistry in Garrod et al. (2022). For consistency, this figure does not include results of other COM observations (e.g., Csengeri et al. 2019; Yang et al. 2021), as they did not use minor isotopologues to derive the column density of methanol as we did.

Fig. 2.4 shows that there is no obvious difference in the O-COM ratios between low-mass and high-mass sources. However, molecules in the two groups should have experienced different physical conditions, such as temperature and fluence of energetic particles (UV photons, X-rays, cosmic rays) before the hot cores and hot corinos formed. This implies that these species are likely formed under similar conditions,

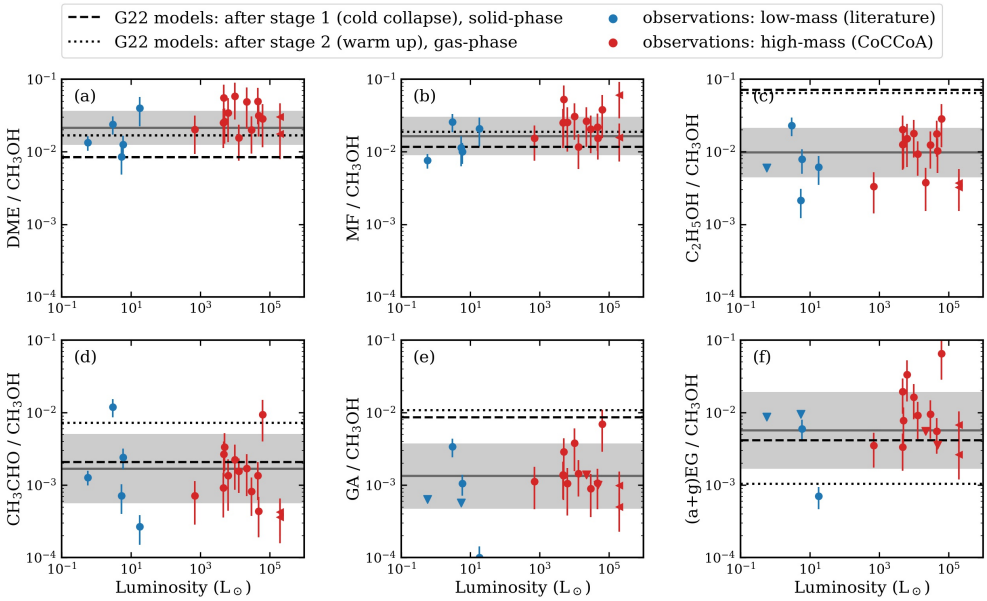


Figure 2.4: Column density ratios of the six selected O-COMs with respect to methanol versus bolometric luminosity. Panels from left to right and top to bottom: dimethyl ether (DME; CH_3OCH_3), methyl formate (MF; CH_3OCHO), ethanol (C_2H_5OH), acetaldehyde (CH_3CHO), glycolaldehyde (GA; CH_2OHCHO), and ethylene glycol (EG; $(CH_2OH)_2$). The red points show the CoCCoA data of 14 high-mass sources derived in this work, and the blue points show the literature data of five low-mass sources: IRAS 16293-2242 A & B (Manigand et al. 2020; Jørgensen et al. 2018), S68N, B1c, and B1-bS (van Gelder et al. 2020). Upper limits are denoted by downward triangles. The data points of NGC 6334-38 and NGC 6334-43 are denoted by leftward triangles due to their upper limits on luminosity. The solid line in gray corresponds to the mean column density ratio of each species, and the gray shaded area shows the standard deviation weighted by the uncertainty on log scales. The mean and standard deviation were calculated from both the high-mass sources and the low-mass ones. The dashed and dotted lines in black correspond to the modeled COM ratios at the end of the cold collapse stage (stage 1) and the warm-up stage (stage 2) in G22, respectively.

which points to a common and early pre-stellar stage before the star formation processes began to differentiate.

The six selected O-COMs have different column density ratios with respect to methanol, varying by over two orders of magnitude between 0.01% and 10%. Among the six species, DME and MF have the highest and similar ratios with respect to methanol (2–3%); C_2H_5OH and EG have intermediate ratios but with a much larger scatter; and CH_3CHO and GA have the lowest ratios, at around 0.1%.

Figure 2.5 shows the spread factor, that is, $\log_{10}$ of the standard deviation of O-COM ratios in log scales, of our combined sample of low-mass and high-mass protostars. It is clear that DME and MF have smaller scatter than the other four species. The scatter may result from an observational effect that different species have different gas-phase emitting areas depending on their sublimation temperatures from dust grains (Nazari et al. 2021, 2022a). The column densities that we derived from spectral

fitting represent the abundances averaged over the observational beam. If the actual emitting area of a species is smaller than the area that we can resolve, then we suffer from a beam dilution issue. This is likely the case in the CoCCoA observations since the moment 0 maps of selected O-COMs given in Fig. 2.A.1 show that the emission of these molecules are barely or not spatially resolved. The actual resolved area associated with a specific beam size depends not only on the angular resolution (which is constant across the CoCCoA sample) but also on the distance to a source (which varies by a factor of three to four across the sample). Moreover, the size of the hot core, as defined by the radius where $T = 100$ K, depends on the square root of the luminosity of the source (Bisschop et al. 2007):

$$R_{T=100\ \mathrm{K}} \approx 15.4\sqrt{\frac{L}{L_\odot}}\ \mathrm{au}. \tag{2.3}$$

Therefore, there is a beam dilution factor between the observationally inferred column density and the actual one. This factor can differ from source to source and from species to species, and hence, it can lead to the scatter in COM ratios. These effects are discussed and quantified in more detail in Nazari et al. (2022b, 2023a).

More generally, we note that surveys of large samples of low- and high-mass protostars have found that some fraction of sources do not show any methanol or COM emission (e.g., Yang et al. 2021; van Gelder et al. 2022, paper I). The reasons for this absence of COM emission are varied but include the possible presence of a disk that lowers the overall temperature structure (Nazari et al. 2022b, paper II) as well as different evolutionary stages, such as the presence of an H II region (Nazari et al. 2023a, modeling paper).

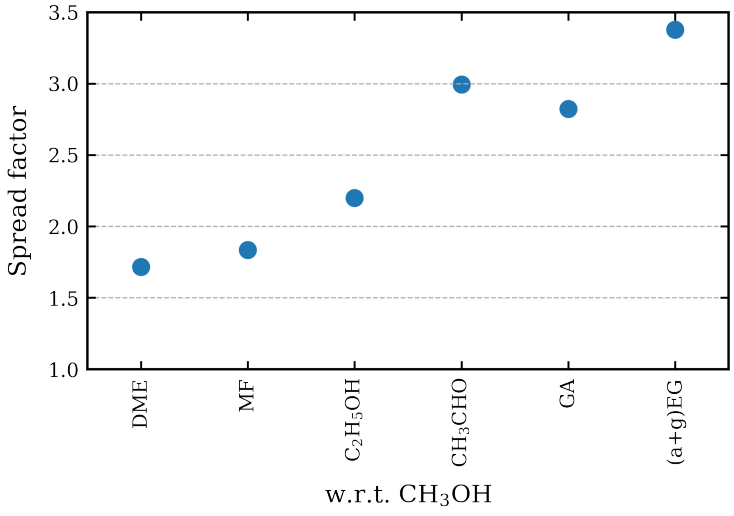


Figure 2.5: Spread factor ($\log_{10}$ of the 1σ scatter) for the ratios plotted in Fig. 2.4 of different O-COMs with respect to methanol.

2.5.2 Observations versus simulations

The simulations used to compare with our observational results are mainly based on the state-of-the-art models in Garrod et al. (2022), hereafter G22. The G22 models simulate the chemistry coupled in three phases: gas phase, ice surface, and bulk ice mantle. The last two phases are collectively known as the solid phase. The evolution of hot cores is treated as two stages, a cold collapse stage followed by a static (i.e., fixed density) warm-up stage once the central protostar has formed. There are three warm-up timescales used in the models: 5×10^4 yr (fast), 2×10^5 yr (medium), and 1×10^6 yr (slow). A major update in the G22 models involves including non-diffusive chemistry on surfaces and in bulk ices, which is proposed to be important in interstellar ices based on laboratory work (Fedoseev et al. 2015; Linnartz et al. 2015). The G22 test the effect of different non-diffusive mechanisms along with other parameters in about 20 models (see Table 1 in G22). The "final" model includes all the discussed non-diffusive mechanisms, and we used it as the fiducial model for further discussion. In Fig. 2.4, the horizontal dashed lines correspond to the COM ratios (w.r.t. methanol) in the solid phase at the end of the collapse stage, and the horizontal dotted lines correspond to the gas-phase ratios after the warm-up stage with the medium warm-up speed. In addition to the O-COM ratios (w.r.t. methanol) shown in Fig. 2.4, we also present a number of ratios between two O-COMs in Fig. 2.6. In the following subsections, we compare our observational results of each species to the "final" model with the "medium" warm-up speed in G22.

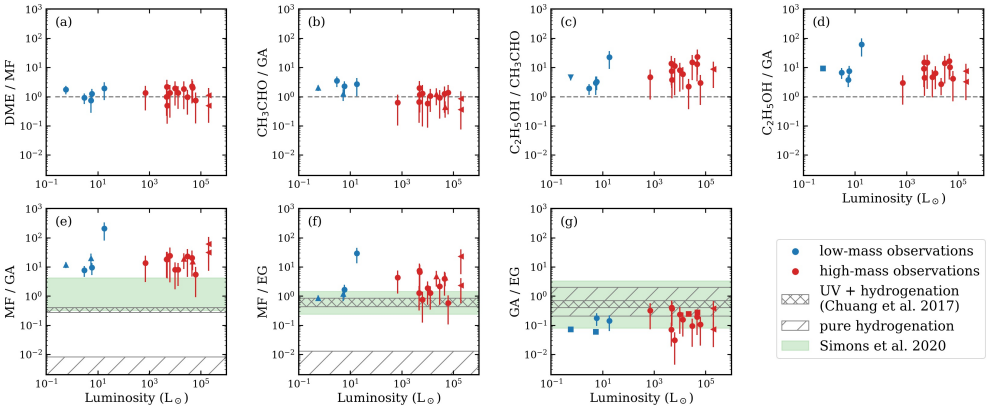


Figure 2.6: Observed abundance ratios between seven pairs of O-COMs. Data points in blue and red correspond to the five low-mass sources from the literature and the 14 high-mass sources from the CoCCoA survey presented in this work, respectively (the same as Fig. 2.4). If the column densities of one species in the pair are upper limits, the ratios would be upper (lower) limits, and the data points are then shown as downward (upward) triangles instead of circles. Square data points without error bars indicate that the column densities of both species in the pair are upper limits. Dashed lines in panels (a)–(d) indicate a ratio of one. In panels (e)–(g), the gray shaded areas with slashes or crosses indicate the ranges of the experimental results in Chuang et al. (2017), and the green shaded areas indicate the simulation predictions by Simons et al. (2020).

2.5.2.1 CH_3OCH_3 (DME) and CH_3OCHO (MF)

In the observations, DME and MF have the most stable column density ratios, and they have the best match with the G22 simulations (Fig. 2.4a,b). Their high abundances are often underproduced by experiments and simulations (Fedoseev et al. 2015; Chuang et al. 2016; Simons et al. 2020; Jin & Garrod 2020). However, the inclusion of the new formation routes in G22 enabled their models to reproduce our observational results on the gas-phase ratios of DME and MF with respect to methanol.

In the G22 models, more than 60% of the DME is formed in ices, through the reaction between methylene (CH_2) and methanol via

$$CH_2 + CH_3OH \rightarrow CH_3OCH_2, \text{ and} \tag{2.4}$$

$$H + CH_3OCH_2 \rightarrow CH_3OCH_3. \tag{2.5}$$

Nearly 90% of CH_3OCH_2 radicals are formed by the combination of CH_2 and CH_3O, in which CH_3O comes from the hydrogenation of formaldehyde (H_2CO) or H abstraction from methanol. As the reactant of two reactions, CH_2 directly affects the formation of DME ices. G22 set the activation energy barrier of the grain-surface reaction

$$C + H_2 \rightarrow CH_2 \tag{2.6}$$

to zero (Krasnokutski et al. 2016; Henning & Krasnokutski 2019) and added about 20 grain-surface CH_2-related reactions to the network (see Table 4 in G22). This greatly enhances the efficiency of forming CH_2 and its subsequent contribution to COMs formation. Nevertheless, this assumption is not fully supported by the combined experimental and theoretical work of Lamberts et al. (2022), who argued that the reaction between C and H_2 is unlikely to be fully barrierless on water ices. However, the effects to the full kinetic model may not be significant with a modest non-zero barrier of reaction (2.6). G22 also introduced a set of methylidyne (CH) reactions (Table 5 in G22) that can form CH_2 and larger hydrocarbons barrierlessly. These reactions can make up the CH_2 formation when reaction (2.6) has a barrier. According to G22, the inclusion of CH and CH_2 chemistry in the solid phase enhances the abundance of solid-phase DME by more than a factor of two. This emphasizes the importance of including the carbon hydrogenation to the chemical network (Qasim et al. 2020).

Besides the bottom-up formation of DME from CH_2, the photodissociation of CH_4 may also provide important ingredients. Several experimental studies have been able to produce DME from UV-irradiated ices of CH_3OH and CH_4 (Öberg et al. 2009b; Paardekooper et al. 2016; Yocum et al. 2021). However, Fedoseev et al. (2015) and Chuang et al. (2016, 2017) did not observe DME formation in their experiments with CH_4 or H_2CO not included in the deposition, even when UV was introduced in Chuang et al. (2017).

According to G22, the remaining 40% DME is formed in the gas phase through two reactions:

$$CH_3OH + CH_3OH_2^+ \rightarrow CH_3OCH_4^+ + H_2O, \text{ and} \tag{2.7}$$

$$CH_3OCH_4^+ + NH_3 \rightarrow CH_3OCH_3 + NH_4^+. \tag{2.8}$$

In the first step, protonated DME is formed via the reaction between methanol and protonated methanol. The second step is the proton transfer to ammonia (Charnley

et al. 1995; Rodgers & Charnley 2001; Taquet et al. 2016), where ammonia comes from ice sublimation, and therefore the reaction sequence would not be efficient under cold conditions (Skouteris et al. 2019).

In the solid phase, MF forms on grain surfaces mainly through the non-diffusive reaction (Chuang et al. 2016)

$$\mathrm{HCO} + \mathrm{CH_3O} \rightarrow \mathrm{CH_3OCHO}. \tag{2.9}$$

A small contribution is made by the newly introduced three-body excited formation (3-BEF) reactions in the bulk ice

$$\mathrm{CH_3O} + \mathrm{CO} \rightarrow \mathrm{CH_3OCO}, \tag{2.10}$$

whereby the hydrogenation of $\mathrm{H_2CO}$ would produce excited $\mathrm{CH_3O}$ that can overcome the reaction barrier (Jin & Garrod 2020). The $\mathrm{CH_3OCO}$ radicals can then recombine with H atoms to form MF. Reactions (2.9) and (2.10) occur in the cold collapse stage and contribute about 70% of the total MF. In the gas phase, the newly added reaction (Balucani et al. 2015)

$$\mathrm{O} + \mathrm{CH_3OCH_2} \rightarrow \mathrm{CH_3OCHO} + \mathrm{H} \tag{2.11}$$

becomes the main production route when $T > 100$ K. The gaseous $\mathrm{CH_3OCH_2}$ radicals can be released from ice mantles or converted from DME by OH abstraction (Shannon et al. 2014):

$$\mathrm{CH_3OCH_3} + \mathrm{OH} \rightarrow \mathrm{CH_3OCH_2} + \mathrm{H_2O}. \tag{2.12}$$

The formation of both DME and MF is strongly related to methanol and its precursor $\mathrm{CH_3O}$. Simulations with different input $n(\mathrm{H})/n(\mathrm{CO})$ ratios by Simons et al. (2020) and experiments by Santos et al. (2022) proposed that the final step of methanol formation is dominated by

$$\mathrm{CH_3O} + \mathrm{H_2CO} \rightarrow \mathrm{CH_3OH} + \mathrm{HCO}. \tag{2.13}$$

Chuang et al. (2016) noticed that $\mathrm{H_2CO}$ is a prerequisite for the formation of MF. In other words, without the input of $\mathrm{H_2CO}$, there would be no MF detected in the outcome, which was also the case in Fedoseev et al. (2015), where only H and CO were used. Since $\mathrm{CH_3O}$ is necessary to form MF, it is inferred that the H abstraction from $\mathrm{CH_3OH}$ yields primarily $\mathrm{CH_2OH}$, while $\mathrm{CH_3O}$ mainly comes from the hydrogenation of $\mathrm{H_2CO}$. This may explain why DME and MF can retain relatively stable abundance ratios with respect to methanol among a large sample of sources, since the formation of all the three species tends to be strongly related with the same precursor, $\mathrm{CH_3O}$. However, it is not clear whether the DME/MF ratio $\sim$1 (Fig. 2.6a) is a pure coincidence or there is some chemical balance between the two COMs.

2.5.2.2 $\mathrm{C_2H_5OH}$

The observed abundance ratio of ethanol is on average one order of magnitude lower than the simulation results in G22. The scatter is also relatively large compared to DME and MF. The G22 models give nearly the same abundance after the collapse stage

and the warm-up stage. This is because, in their models, ethanol is formed almost entirely in the early cold collapse stage on dust grains. The dominant reactions are:

$$C_2H_5 + OH \rightarrow C_2H_5OH \tag{2.14}$$

in the bulk ice and

$$CH_2 + CH_3OH \rightarrow C_2H_5OH \tag{2.15}$$

on the surface. The C_2H_5 radicals in reaction (2.14) are formed by diffusive reactions of atomic H with ethane (C_2H_6) and ethylene (C_2H_4). These routes may be supported by the experiments in Chuang et al. (2020) showing that ethanol can be formed through non-energetic processing of C_2H_2 ices. However, the efficiency of these reactions is not well constrained.

The OH radicals in reaction (2.14) come from photodissociation of water ice by external UV radiation. The G22 models used the same initial visual extinction ($A_{V,\ \mathrm{init}} = 3$) throughout the molecular cloud, while in reality, A_V is higher in the inner part of the cloud. This difference may result in an overestimation of the UV intensity in the cold collapse stage and hence the overproduction of ethanol on dust grains.

Apart from the production routes, there is also a recently proposed gas-phase destruction mechanism, the "ethanol tree" network (Skouteris et al. 2018), that is not included in the G22 models. This new network starts with the H abstraction from ethanol by halogen atoms or OH radicals, leading to two reactive radicals: CH_3CHOH and CH_2CH_2OH. These radicals are further converted into formic acid (HCOOH) and formaldehyde (H_2CO) by reacting with O atoms, along with other minor products (see branching ratios in Fig. 1 of Skouteris et al.). As a result, the abundance of gas-phase ethanol decreases over time. Skouteris et al. (2018) predicts the ethanol/GA ratio to fall from ∼200 to ∼10 in about 1000 years. In our observations, the C_2H_5OH/GA ratios are around ten (see Fig. 2.6d). This suggests that the gas-phase destruction of ethanol in a later stage may play a role in explaining its overproduction in chemical models.

2.5.2.3 CH_3CHO

In the G22 simulations, the abundance ratio of CH_3CHO with respect to methanol in the solid phase after the collapse stage agrees well with our observational data. However, there are discrepancies among the warm-up stages with different timescales. The longer the warm-up stage is, the more CH_3CHO is produced, indicating substantial gas-phase formation. In the G22 models, only 25% of the total amount of CH_3CHO is formed in ices during the cold collapse stage ($T \sim 10$ K) through the hydrogenation of ketene

$$CH_2CO + 2H \rightarrow CH_3CHO. \tag{2.16}$$

About 35% is produced in the early warm-up stage ($T < 100$ K) by

$$CH_2 + H_2CO \rightarrow CH_3CHO \tag{2.17}$$

on the grain surface and

$$CH_3 + CO + H \rightarrow CH_3CHO \tag{2.18}$$

in the bulk ice. Nearly 40% is formed in the gas phase through the reaction

$$C_2H_5 + O \rightarrow CH_3CHO + H \tag{2.19}$$

after the desorption from dust grains at $T > 200$ K in the warm-up stage. Vazart et al. (2020) give a brief overview on the recorded gas-phase formation routes of CH_3CHO and explore some new reactions by theoretical computations. They confirm that reaction (2.19) is efficient in the temperature range of 7–300 K.

The observed ratios of CH_3CHO only match the modeled values after the cold collapse stage and the fast warm-up stage (the latter of which is not shown in Fig. 2.4d), which implies that the formation of CH_3CHO in the early stage may be more dominant than suggested by the G22 models. The experimental studies by Fedoseev et al. (2022) proposed a similar formation route for CH_3CHO, that is, the hydrogenation of ketene on cold (10 K) surfaces of mixed C, H, CO, and H_2O ices. They proposed that H_2CCO can be formed through hydrogenation of CCO radicals, which are the product of the barrierless association reaction between CO molecules and C atoms. Unfortunately, the intermediate products, CCO and HCCO, are very difficult to observe in experiments due to their high reactivity, and the relevant reaction rates remain undetermined. Chuang et al. (2020, 2021) also proposed formation routes from C_2H_2 ices to CH_3CHO at $T = 10$ K under both non-energetic and energetic conditions. To better understand the difference between the chemical modeling results in G22 (especially the gas-phase ratios after the warm-up stage) and our observational results, more knowledge about the relative importance of CH_3CHO formation in the solid and gas phase is needed.

We find it interesting to note the relation between CH_3CHO and the two more hydrogenated O-COMs, GA and C_2H_5OH. The ratios of CH_3CHO/GA and C_2H_5OH/CH_3CHO shown in Fig. 2.6(b) and (c) behave differently: CH_3CHO/GA is ∼1 with a small scatter comparable to DMF/MF, while C_2H_5OH/CH_3CHO exhibits a larger scatter. These results are counterintuitive since C_2H_5OH is suggested to be a direct hydrogenation product of CH_3CHO (e.g., experiments by Fedoseev et al. 2022), while GA and CH_3CHO are not usually simultaneously present in experiments. More investigation is needed to verify if GA and CH_3CHO are chemically linked.

2.5.2.4 CH_2OHCHO (GA) and $(CH_2OH)_2$ (EG)

In contrast to MF, simulations and experiments usually produce more GA than observations (e.g., Chuang et al. 2016, 2017; Simons et al. 2020), and the same occurs with the G22 models. Fedoseev et al. (2015, 2017) have confirmed experimentally that GA and EG can be formed in the solid phase through surface hydrogenation of CO molecules at $T = 13$ K. They suggest that GA and EG are the outcome of successive hydrogenation of glyoxal (HC(O)CHO), which is formed from two HCO radicals. In the G22 models, EG is mainly formed through the addition of two CH_2OH radicals in the solid phase:

$$CH_2OH + CH_2OH \rightarrow (CH_2OH)_2. \tag{2.20}$$

Most of the GA is formed through repetitive H abstraction from EG:

$$H + (CH_2OH)_2 \rightarrow CH_2OH_2CH_2O + H_2, \text{ and} \tag{2.21}$$

$$H + CH_2OHCH_2O \rightarrow CH_2OHCHO + H_2. \tag{2.22}$$

About 60% of the conversion from EG to GA through reaction (2.21)–(2.22) is finished in the cold collapse stage when $T \lesssim 10$ K, while about 30% occurs in the middle of the warm-up stage when T ∼100–200 K. A small portion of solid-phase GA is formed via the route proposed by Chuang et al. (2016):

$$CH_2OH + HCO \rightarrow CH_2OHCHO. \quad (2.23)$$

Despite the large scatter in the observed abundance ratios of GA and EG, it is obvious that EG is overall more abundant than GA in observations. However, the G22 models give the opposite result. This implies that the interconversion between EG and GA may not be well modeled in G22.

Another simulation work by Simons et al. (2020) computed their O-COM network with four H/CO input ratios (5–60%) at six low temperatures (8–20 K). They summarized the flux distribution of the network for the fiducial model with $n(\mathrm{CO}) = 10.0$ cm^{-3}, $n(\mathrm{H}) = 2.5$ cm^{-3}, and $T = 10$ K (see Fig. 8 of Simons et al.). They found that the hydrogenation of glyoxal is more important to the GA formation than the H abstraction from EG. Their results also show that the relative abundance of GA to EG is very sensitive to the $n(\mathrm{H})/n(\mathrm{CO})$ ratio. The observed abundance ratios of GA over EG are ∼0.1–1 (Fig. 2.6g), which corresponds to $n_{\mathrm{initial}}(\mathrm{H})/n_{\mathrm{initial}}(\mathrm{CO}) > 0.25$ in their models. Chuang et al. (2017) were able to reproduce a similar GA/EG ratio through the hydrogenation of ices with $CO:CH_3OH = 4:1$ at $T = 14$ K and found it to be subject to the initial composition ratios. Besides pure hydrogenation, they tried introducing UV radiation, but the GA/EG ratio was not affected. The simulations and experiments mentioned above suggest that the formation of GA and EG is strongly regulated by the relative abundance of H atoms with respect to other ingredients such as CO. Parameters such as activation energy barriers and branching ratios of the related reactions of GA and EG formation may also attribute to the difference between simulations and observations.

The abundance ratio of GA with respect to methanol predicted by the G22 models is nearly one order of magnitude higher than our observational results. A possible reason for the overproduction by the chemical models is that GA has a higher desorption temperature than methanol (see Fig. 2 in Fedoseev et al. 2015). As a result, GA is expected to desorb and emit from a smaller region than methanol (for quantification, see the toy model described in Appendix B of Nazari et al. 2021). If our spatial resolution is not high enough to resolve the actual emitting region of GA (suggested by the moment 0 map in Fig. 2.A.1), there will be a beam dilution effect leading to an underestimated column density ratio of GA compared to the actual abundance ratio. An interesting fact is that EG has an even higher desorption temperature than GA (Fedoseev et al. 2015) and is expected to suffer more from beam dilution, but it is not overproduced by the G22 models. This means that if beam dilution actually accounts for the underestimation of GA in observations, EG would have been underproduced by the G22 models; otherwise, there should be other reasons for the observed GA depletion.

2.5.3 The influence of energetic processes

As mentioned in Sect. 2.5.2.4, Chuang et al. (2017) showed that the ratios among MF, GA, and EG can be altered by UV radiation. The ratios of MF with respect to GA and

EG both increase by nearly two orders of magnitude when introducing UV radiation to the experiments, while the ratio between GA and EG is not much affected. The second row of Fig. 2.6 shows a comparison between observations, simulations, and experiments (this work, Simons et al. 2020; Chuang et al. 2017, respectively). The UV intensity was not varied during the experiments, but the discrepancy between pure hydrogenation and UV irradiation implies a positive correlation between the UV irradiation and the relative abundance of MF. However, the observed ratios of MF/GA and MF/EG are still higher overall than the values produced by simulations and experiments.

In G22, one of the models tests the influence of cosmic ray-induced ionization and UV-induced photodissociation, which shows an obvious enhancement of the O-COM abundances. The "final" model includes these energetic processes but with a fixed efficiency, as do the simulations in Simons et al. (2020). The G22 models are able to reproduce enough MF by introducing a new set of non-diffusive reactions, but the abundances of GA and EG are not very consistent with our observations. Experiments by Öberg et al. (2009b) suggest that the final product composition after irradiating CH_3OH ices with UV lamps depends more on the UV fluence and temperature than the UV flux itself. There is a possible explanation that the short timescale offsets the high UV flux in high-mass sources, and the total UV fluence falls in the same order as low-mass sources. To figure out the influence of energetic processes on COM chemistry, especially in the solid phase, more experiments and simulations with varied parameters are needed. Observations of larger samples of protostellar objects with different masses and luminosities are also needed in order to provide more reliable statistics of COMs ratios.

2.5.4 O-COMs from clouds to comets

Figure 2.7 summarizes the observed O-COM ratios with respect to methanol in different astronomical objects. The data of protostars are represented by the average ratios of the low-mass and high-mass sources discussed in this work, and the uncertainties correspond to the standard deviation. In the figure, we compare the statistical data in this work with the literature data of the outbursting protostar V883(FU) Ori (Lee et al. 2019b) and the protoplanetary disk around Oph IRS 48 (Brunken et al. 2022). Two comets, 67P/Churyumov–Gerasimenko (67P/C–G) (Rubin et al. 2019a; Drozdovskaya et al. 2019) and 46P/Wirtanen (Biver et al. 2021), are also taken into account, as they reflect the pristine chemical composition in our solar system. All the sources except the two comets were observed by ALMA. The data of 67P/C-G were collected by the Rosetta Orbiter Spectrometer for Ion and Neutral Analysis (ROSINA), and the data of 46P/Wirtanen were taken by the Institut de Radio Astronomie Millimétrique (IRAM) 30-m telescope and the NOrthern Extended Millimeter Array (NOEMA). ROSINA is a mass spectrometer, and it cannot distinguish among isomers with the same mass (e.g., C_2H_5OH and DME have the same atomic mass unit of 46). Considering the data availability, we chose three groups of O-COMs for comparison:

amu = 44: CH_3CHO, CH_2CHOH (vinyl alcohol, VA), c-C_2H_4O (ethylene oxide, EO);

amu = 46: C_2H_5OH, DME; and

amu = 60: MF, GA, CH_3COOH (acetic acid, AA).

We mention EO, VA, and AA only for the potential degeneracy of the detection of 67P/C-G; they were not searched for nor detected in other sources except one low-mass protostellar object (IRAS 16293-2242 B) that we consider here. The O-COM ratios of IRS 48 may be overestimated since the column density of methanol was determined from lines that are likely to be optically thick (hence upper limit signs are used).

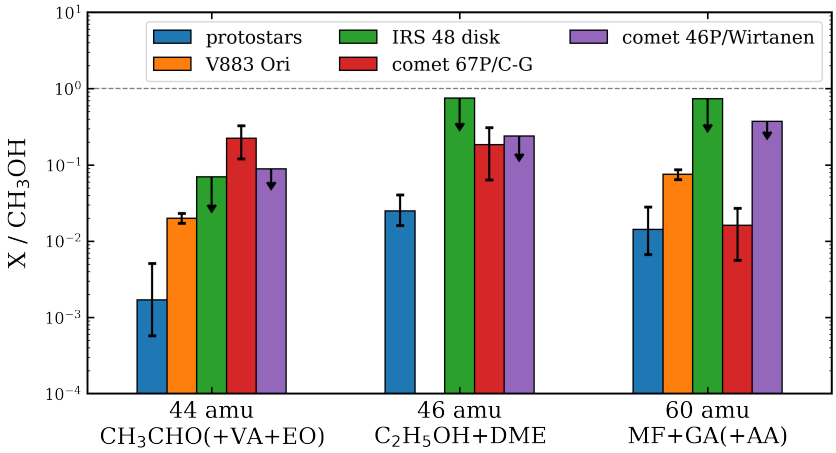


Figure 2.7: Column density ratios of selected O-COMs with respect to methanol in different astronomical objects (see references in Sect. 2.5.4). Species are divided into three groups with different atomic mass units (amu); those in parentheses are only shown for the degenerate detections in the comet 67P/C-G by the ROSINA mass spectrometer, and they have not been detected separately in other sources. The column densities of the species with the same atomic mass units have been added together. The data of IRS 48 are shown as upper limits due to the possible underestimation of the column density of methanol.

In general, protostellar sources have lower O-COM ratios than other sources, especially for the group of amu = 44. For the other two groups, considering that the data of IRS 48 may be overestimated, the O-COM ratios are more comparable among the samples. However, the sample here is too small to draw any robust conclusion. A larger sample of sources at different evolutionary stages are needed to study the chemical inheritance throughout star formation.

2.5.5 JWST ice observations

With JWST successfully operating, it is now becoming more possible to detect COMs other than methanol in ices. The absorption features for the identification of O-COM ices are mainly located in the mid-infrared between 2 and 15 μm (Boogert et al. 2015; Rocha et al. 2022). The absorption features at 7.24 and 7.41 μm observed by ISO (Schutte et al. 1999), *Spitzer*/IRS (Öberg et al. 2011), and the latest JWST/MIRI (Yang et al. 2022; McClure et al. 2023) are tentatively attributed to CH_3CHO and C_2H_5OH ices. Additionally, the feature around 11.3 μm in the JWST spectra may also have some contributions from CH_3CHO, C_2H_5OH, and MF (Terwisscha van Scheltinga et al. 2018, 2021). However, more detailed spectral modeling is needed to confirm the detection of these COMs. Once JWST observations with larger samples and higher

sensitivities are available, we will be able to estimate the ratios of CH_3CHO and C_2H_5OH with respect to methanol in ices based on the absorption features at 7.2 and 7.4 μm. By comparing their ice ratios to the gas-phase ones in the same sources, we can get an idea as to whether and to what extent these species participate in gas-phase chemistry after they sublimate from the ice mantles of dust grains.

2.6 Conclusions

We analyzed the spectra of 14 high-mass protostellar objects from the CoCCoA survey. We focused on six selected O-COMs: CH_3CHO, C_2H_5OH, DME, MF, GA, and EG, and we derived their column densities as well as the excitation temperatures for the 14 sources. We also performed various comparisons between the observed O-COM ratios with respect to methanol and the results from previous simulations and experiments. We summarize our conclusions as follows:

- The gas-phase column density ratios of the six selected O-COMs with respect to methanol show no clear difference between the five low-mass objects studied previously and the 14 high-mass protostellar objects observed with ALMA, suggesting that these species are mainly formed under similar conditions. Current astrochemical simulations and experiments support the possibility of early formation of COM ices on dust grains in the pre-stellar stage, before the environments in low- and high-mass star-forming regions begin to diverge. However, the possibility exists that other gas-phase formation routes also play an important role in shaping the COM ratios, which needs more investigations to pin down.
- DME and MF show smaller scatter in their ratios with respect to methanol than CH_3CHO, C_2H_5OH, GA, and EG. This may hint at some chemical links among DME, MF, and methanol, such as having the same precursor (e.g., CH_3O) in their formation routes.
- The ratios among pairs of O-COMs also show the same trends between the low-mass and the high-mass groups. In particular, the ratios of DME/MF and CH_3CHO/GA are quite consistently around one, while others show larger scatter.
- Previous experiments show that the ratios of MF/GA and MF/EG can be significantly enhanced by UV irradiation, but the observed values are even higher than the laboratory ones. The ratio of GA/EG is not affected by UV in experiments and match well with our observations.
- The comparison of our observational data with the state-of-the-art models shows consistency in the O-COM ratios for some species, such as DME and MF. The differences between models and observations may result from less constrained gas-phase chemistry (CH_3CHO and C_2H_5OH) and the different emitting areas under limited spatial resolutions (GA and EG). Further chemical simulations and laboratory experiments are important to testing and exploring possible explanations.
- The comparison of O-COM ratios among sources at different evolutionary stages may probe the chemical inheritance during star formation processes. However, observations toward larger samples are needed to enable statistical analyses.

ALMA line surveys toward large samples of star-forming regions are shedding light on the origin of COMs and the chemical evolution in the early stages of star formation. We look forward to linking our results to more ALMA observations on gas-phase COMs as well as to the upcoming JWST mid-infrared data on solid-phase COMs in order to probe their formation history.

Acknowledgements

The authors are grateful to the entire CoCCoA team for their interest in this study. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2019.1.00246.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. Astrochemistry in Leiden is supported by the Netherlands Research School for Astronomy (NOVA), by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 101019751 MOLDISK), by the Dutch Research Council (NWO) grants TOP-1 614.001.751 and 618.000.001, and by the Danish National Research Foundation through the Center of Excellence "InterCat" (Grant agreement no.: DNRF150). The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. Y.C. acknowledges Jiao He for useful discussion about astrochemical experiments. J.K.J. acknowledges support from the Independent Research Fund Denmark (grant number 0135-00123B). M.N.D. is supported by the Swiss National Science Foundation (SNSF) Ambizione grant 180079, the Center for Space and Habitability (CSH) Fellowship, and the IAU Gruber Foundation Fellowship. S.I. acknowledges the Danish National Research Foundation through the Center of Excellence "InterCat" (Grant agreement no.: DNRF150). B.M.K. is supported by the Swiss National Science Foundation (SNSF) Ambizione grant 180079. N.F.W.L. acknowledges funding by the Swiss National Science Foundation Ambizione grant 193453.

Appendix

2.A Additional figures

Figure 2.A.1 shows the moment 0 maps of selected O-COMs in G19.88-0.53. Figure 2.A.2 shows how the fitting with the best fit and the upper and lower limits of $T_{\rm ex}$ appear, taking MF and G19.88-0.53 as an example. Figure 2.A.3 shows the spectral fitting with two components, taking G19.01-0.03 as an example. Figure 2.A.4 shows the ratios of the six selected O-COMs with respect to methanol, which is the same as Fig. 2.4 but with the intercept error in Eq. (2.1) included. Figure 2.A.5 shows the full spectral fitting results of selected O-COMs for the representative source G19.88-0.53. Figures 2.A.6–2.A.8 show the zoom-in fitting of each O-COM.

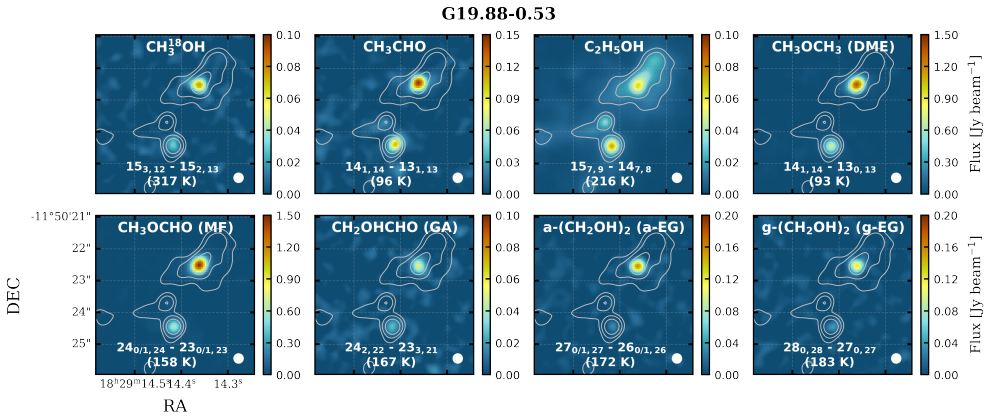


Figure 2.A.1: Moment 0 maps of selected O-COMs in the representative source G19.88-0.53. Contours and beams are the same as those in Fig. 2.1. We chose the strongest and unblended (or less blended) line in the upper tuning to make a moment 0 map of each species. The quantum numbers and the upper energy level are indicated in white bold text at the bottom of each panel. The images are integrated over the FWHM with respect to the line center.

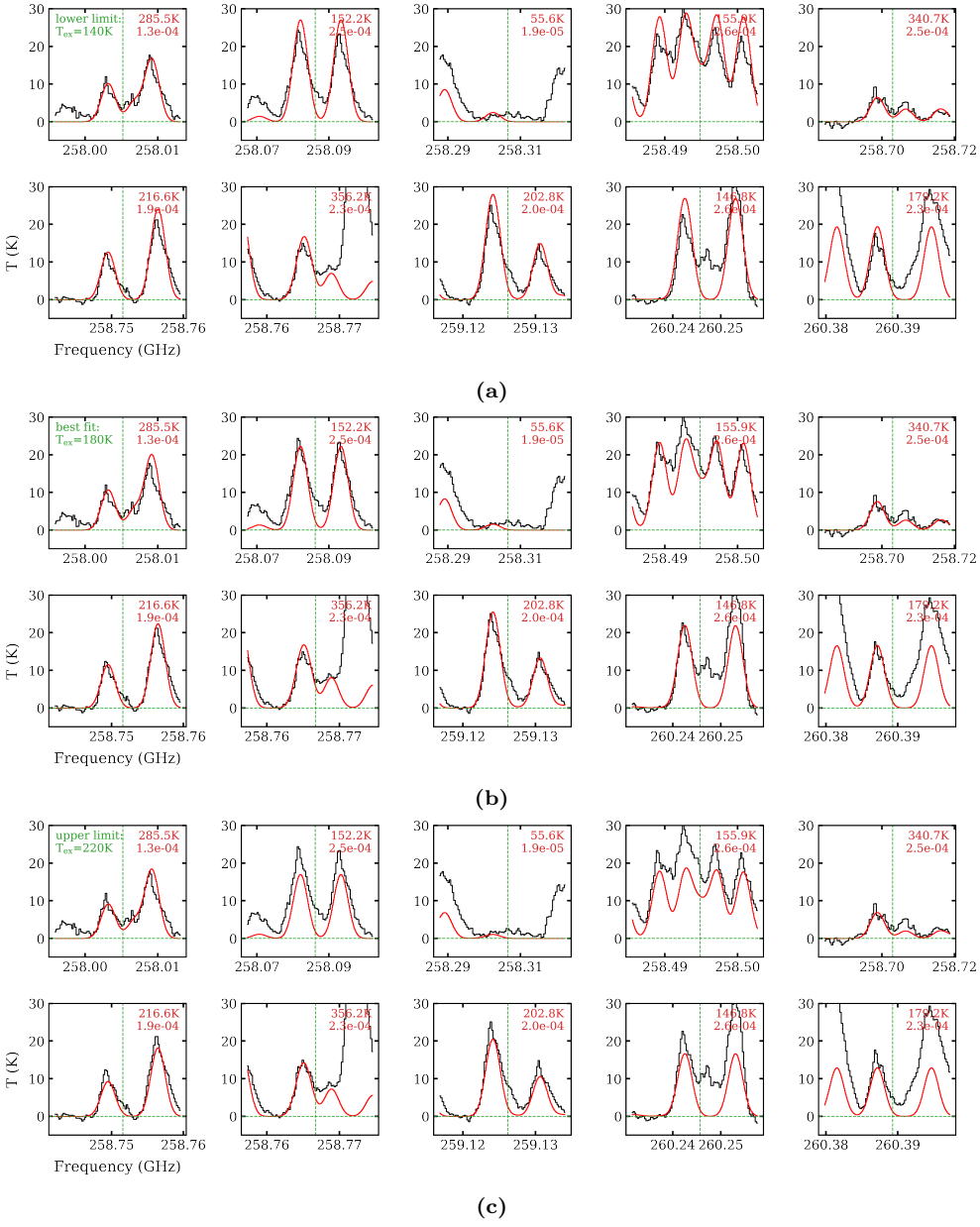


Figure 2.A.2: Example of spectral fitting with (a) the lower limit, (b) the best fit, and (c) the upper limit of the excitation temperature ($T_{\rm ex}$) of MF (CH_3OCHO) in the spectrum of G19.88-0.53. Ten transitions with different upper energy levels ($E_{\rm up}$) are shown for comparison.

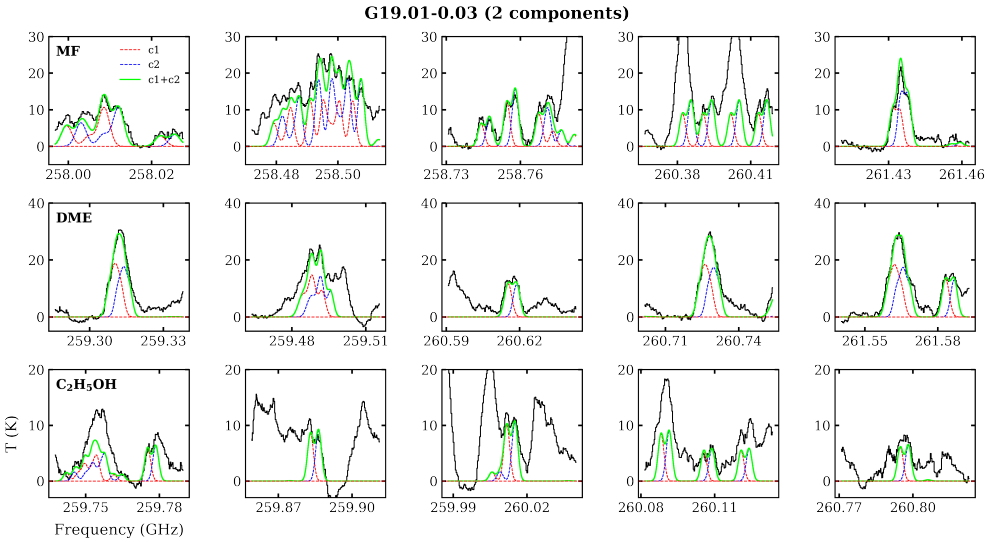


Figure 2.A.3: Example of spectral fitting that introduces two components, represented by three O-COMs of G19.01-0.03. The observed spectrum is shown in solid black. The two components (c1 and c2) are shown with dashed red and blue lines, respectively. The sum of the two components is shown in solid green. See Tables 2.B.1–2.B.4 for other sources and species that are fit by two components.

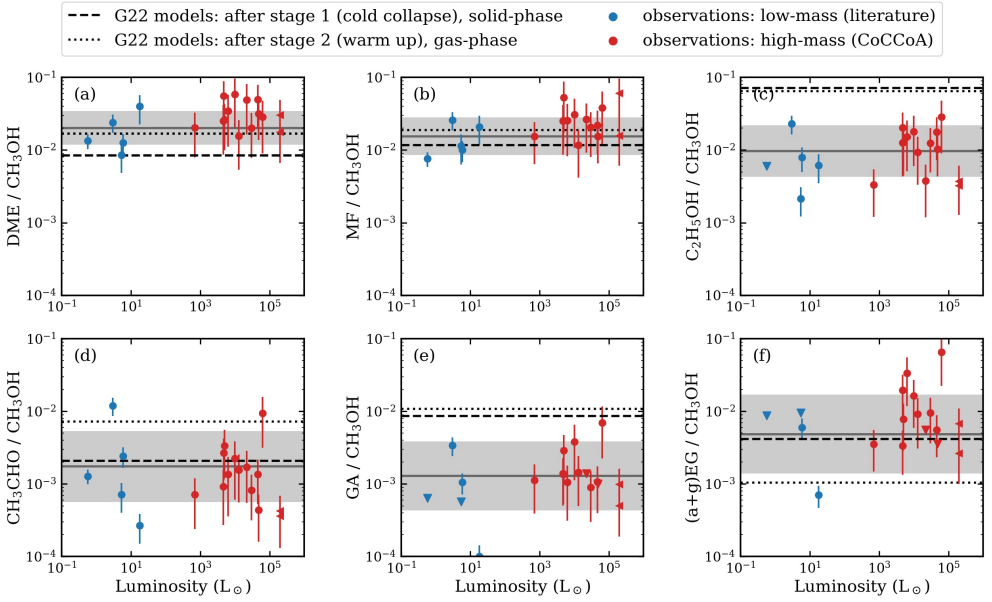


Figure 2.A.4: Same as Fig. 2.4 but includes the intercept error in Eq. (2.1).

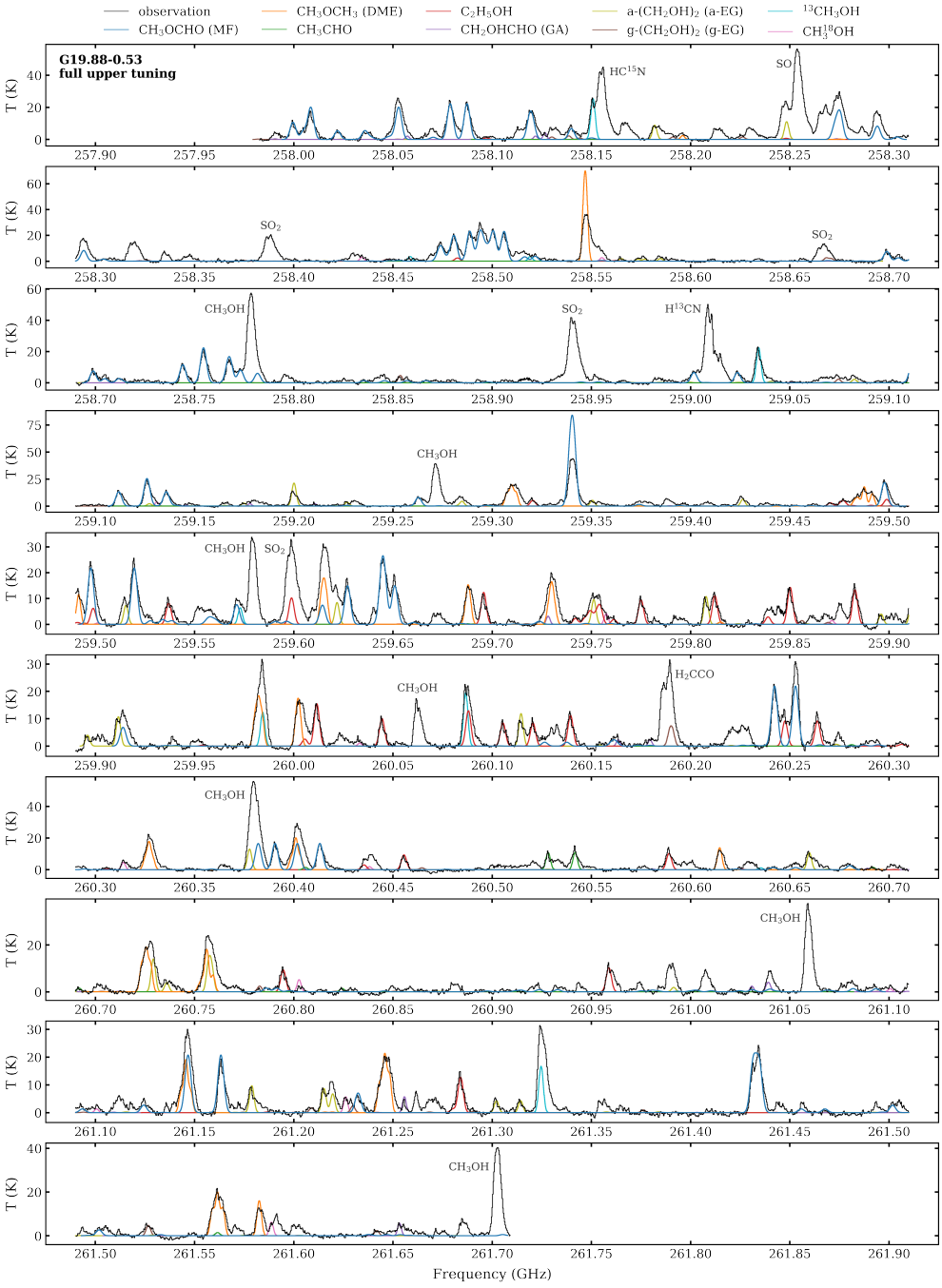
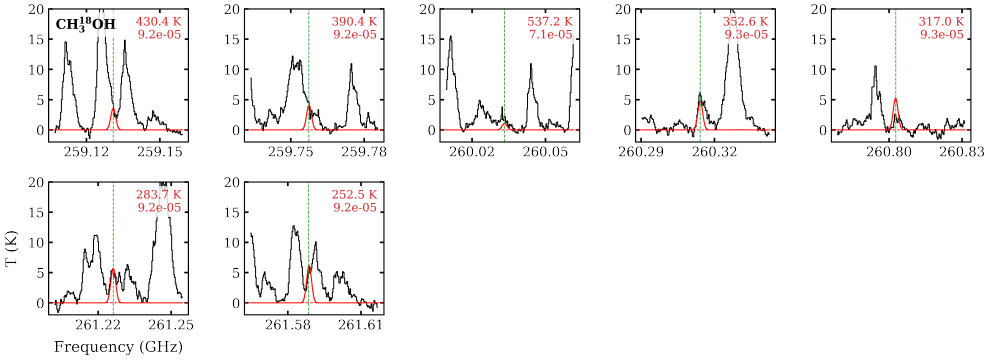
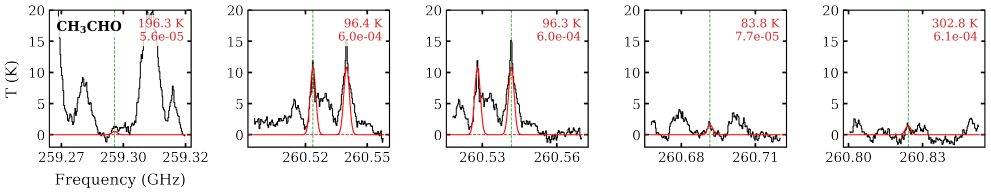


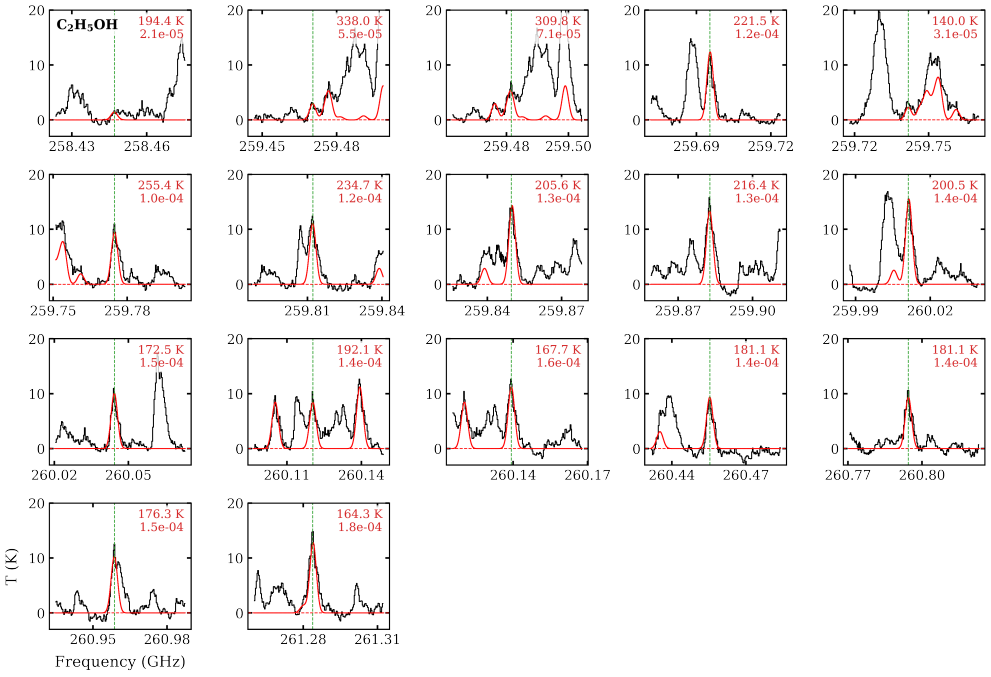
Figure 2.A.5: Same as Fig. 2.3 but for the full upper tuning. Some identified lines from species other than the selected O-COMs are labeled in gray text.



(a)



(b)



(c)

Figure 2.A.6: Best-fit model for $CH_3^{18}OH$, CH_3CHO, and C_2H_5OH, taking G19.88-0.53 as an example. Only selected unblended lines are shown. In each panel, the centered transition is indicated by the vertical dashed green line; the upper energy level (in K) and the Einstein A coefficient (in s^{-1}) are listed in red text in the upper right.

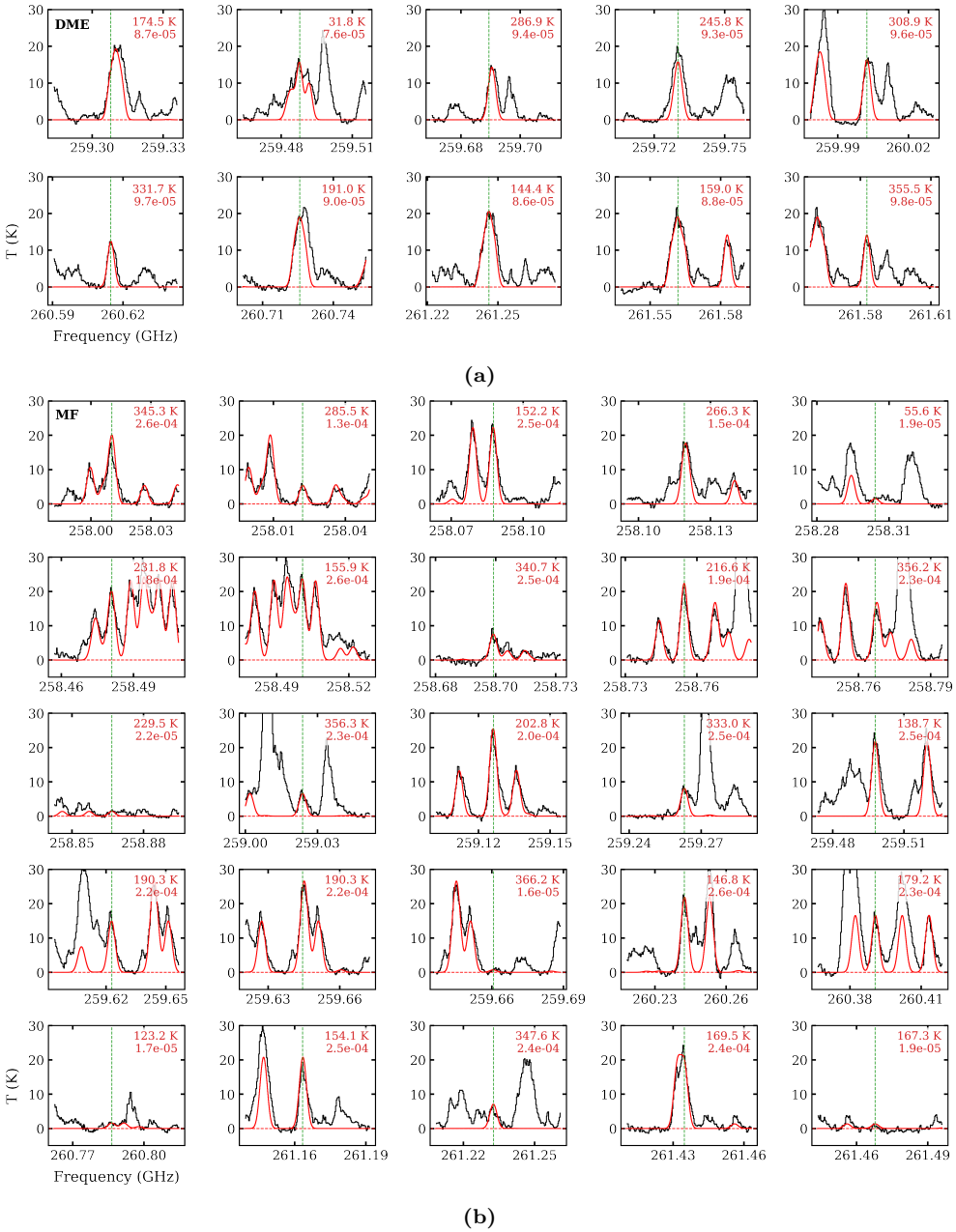
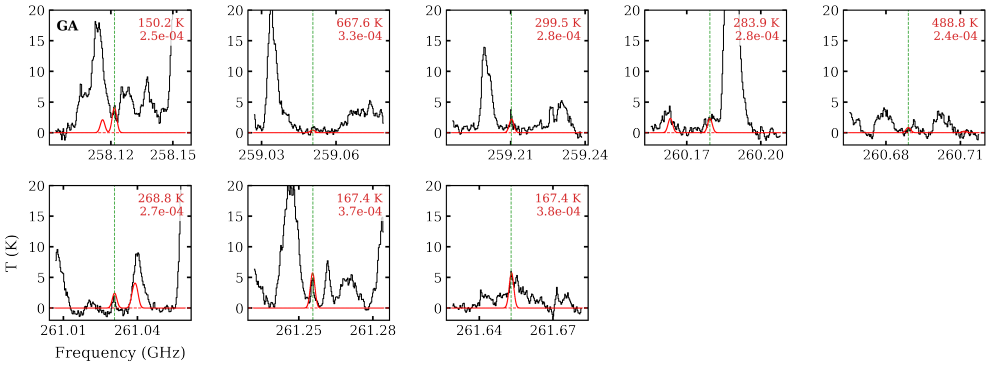
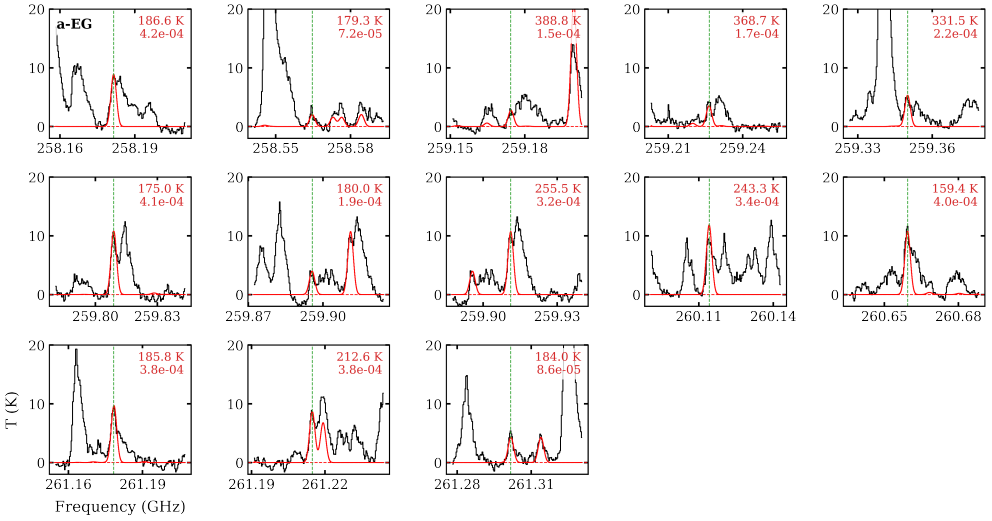


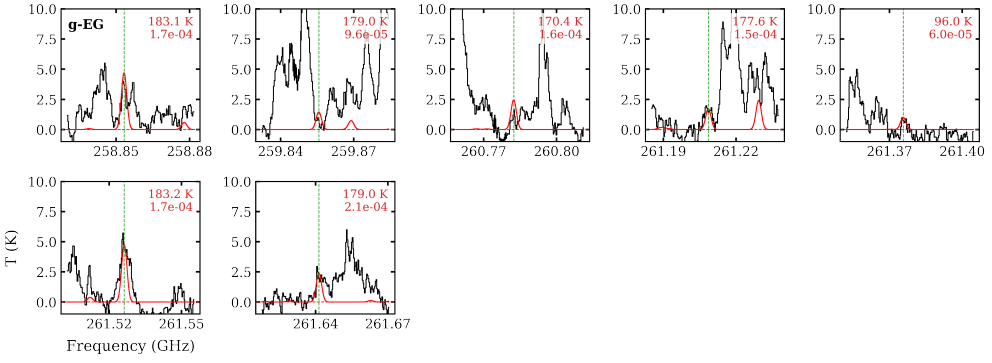
Figure 2.A.7: Same as Fig. 2.A.6 but for CH_3OCH_3 (DME) and CH_3OCHO (MF).



(a)



(b)



(c)

Figure 2.A.8: Same as Fig. 2.A.6 but for CH_2OHCHO (GA) and *a*- and *g*-$(CH_2OH)_2$ (EG).

2.B Additional tables

Tables 2.B.1–2.B.4 give the fitting results of the selected O-COMs of the 14 CoCCoA sources. Table 2.B.5 provides the column densities of GA and EG that were derived directly from the spectral fitting (i.e., without vibrational corrections). Table 2.B.6 lists the identified transitions of two minor isotopologues of CH_3OH and the six selected O-COMs focused on in this paper.

Table 2.B.1: Fitting results of $CH_3^{18}OH$ and CH_3CHO.

	CH_3OH^a	$CH_3^{18}OH$ (CDMS)^b				CH_3CHO (JPL)			
Source Name	N cm^{-2}	N (cm^{-2})	T_{ex} (K)	FWHM (km s^{-1})	v_{lsr}	N (cm^{-2})	T_{ex} (K)	FWHM (km s^{-1})	v_{lsr}
G19.01-0.03 (c1)^c	$(3.6\pm1.1)\times10^{18}$	$(1.2\pm0.3)\times10^{16}$	[180]^d	3.0	63.0				
G19.01-0.03 (c2)^c						$(8.0\pm3.0)\times10^{15}$	[150]	3.2	58.5
G19.88-0.53	$(4.4\pm1.3)\times10^{18}$	$(1.3\pm0.3)\times10^{16}$	[180]	2.7	46.2	$(4.0\pm1.5)\times10^{15}$	100±30	3.0	46.4
G22.04+0.22	$(1.3\pm0.4)\times10^{19}$	$(4.0\pm1.0)\times10^{16}$	[130]	4.5	53.5	$(4.5\pm1.0)\times10^{16}$	180±20	4.5	52.5
G23.21-0.37	$(2.6\pm0.7)\times10^{19}$	$(9.0\pm2.0)\times10^{16}$	[170]	3.0	76.8	$(4.0\pm0.6)\times10^{16}$	150±30	3.5	77.0
G34.30+0.20 (c1)	$(9.6\pm2.8)\times10^{18}$	$(2.2\pm0.5)\times10^{16}$	[140]	3.5	56.5	$(1.3\pm0.3)\times10^{16}$	130±20	3.5	56.3
G34.41+0.24 (c1)	$(7.9\pm2.3)\times10^{18}$	$(1.8\pm0.4)\times10^{16}$	[160]	3.5	60.0	$(1.2\pm0.3)\times10^{16}$	140±20	3.0	60.0
G34.41+0.24 (c2)						$(9.0\pm2.0)\times10^{15}$	[160]	3.0	56.8
G35.03+0.35	$(1.8\pm0.7)\times10^{18}$	$(4.5\pm1.5)\times10^{15}$	[150]	2.5	44.8	$(2.5\pm1.0)\times10^{15}$	110±30	4.0	44.8
G35.20-0.74N (c1)	$(8.2\pm2.5)\times10^{18}$	$(2.0\pm0.5)\times10^{16}$	[140]	3.5	31.0	$(5.0\pm1.0)\times10^{15}$	100±20	3.0	31.0
G35.20-0.74N (c2)						$(1.8\pm0.5)\times10^{15}$	100±20	1.7	33.5
G345.5+1.5	$(1.4\pm0.3)\times10^{18}$	$(3.2\pm0.5)\times10^{15}$	120±30	2.0	-15.8	$(6.0\pm2.0)\times10^{14}$	[120]	2.0	-16.3
IRAS 18151-1208	$(5.3\pm2.0)\times10^{17}$	$(1.5\pm0.5)\times10^{15}$	[180]	2.0	35.0	$(9.0\pm2.0)\times10^{14}$	150±30	2.0	35.0
IRAS 16547-4247	$(2.1\pm0.8)\times10^{18}$	$(6.0\pm2.0)\times10^{15}$	[200]	2.2	-35.5	$(2.0\pm0.4)\times10^{16}$	150±30	3.0	-35.5
NGC 6334-38	$(4.9\pm1.5)\times10^{18}$	$(1.2\pm0.3)\times10^{16}$	[120]	2.7	-5.0	$(1.8\pm0.5)\times10^{15}$	100±30	2.8	-5.1
NGC 6334-43	$(5.4\pm1.6)\times10^{18}$	$(1.3\pm0.3)\times10^{16}$	[160]	3.2	0.8	$(2.3\pm0.5)\times10^{15}$	100±50	3.0	1.0
NGC 6334 I(N)-SM2	$(3.9\pm1.1)\times10^{18}$	$(9.0\pm2.0)\times10^{15}$	[150]	3.0	-3.0	$(2.8\pm1.0)\times10^{15}$	80±30	3.0	-2.5

[a]The column density of the major isotopologue of methanol was inferred from the minor isotopologue $CH_3^{18}OH$ and the isotopic ratio between ^{16}O and ^{18}O.

[b]The spectroscopic databases used in the fitting.

[c]Component 1 and 2 (some sources exhibit two components in their spectra); see Fig. 2.A.3 for an example.

[d]The T_{ex} in square brackets means that it was fixed to that value during the fitting. In this case, the uncertainties of N were determined only for this particular T_{ex}.

Table 2.B.2: Fitting results of C_2H_5OH and CH_3OCH_3 (DME).

	C_2H_5OH (CDMS)				CH_3OCH_3 (CDMS)			
Source Name	N^a (cm^{-2})	T_{ex} (K)	FWHM (km s^{-1})	v_{lsr}	N^a (cm^{-2})	T_{ex} (K)	FWHM (km s^{-1})	v_{lsr}
G19.01-0.03 (c1)	$(3.1\pm0.8)\times10^{16}$	160±40	3.2	62.3	$(1.1\pm0.2)\times10^{17}$	130±20	3.8	61.8
G19.01-0.03 (c2)	$(3.3\pm0.6)\times10^{16}$	160±40	3.2	58.5	$(1.0\pm0.2)\times10^{17}$	130±20	3.5	57.8
G19.88-0.53	$(5.5\pm1.5)\times10^{16}$	170±30	3.5	46.4	$(1.1\pm0.3)\times10^{17}$	130±20	3.5	46.0
G22.04+0.22	$(1.7\pm0.4)\times10^{17}$	[160]	5.5	53.0	$(3.5\pm0.5)\times10^{17}$	120±20	5.5	52.3
G23.21-0.37	$(2.4\pm0.4)\times10^{17}$	160±40	3.3	77.2	$(4.0\pm0.8)\times10^{17}$	120±10	3.0	77.2
G34.30+0.20 (c1)	$(1.2\pm0.3)\times10^{17}$	170±30	3.3	56.5	$(3.7\pm0.6)\times10^{17}$	130±20	3.5	56.5
G34.30+0.20 (c2)	$(5.0\pm1.0)\times10^{16}$	170±30	3.0	60.0	$(1.1\pm0.3)\times10^{17}$	130±20	3.0	60.0
G34.41+0.24 (c1)	$(8.0\pm2.0)\times10^{16}$	180±20	3.0	60.0	$(2.2\pm0.3)\times10^{17}$	130±20	3.3	60.0
G34.41+0.24 (c2)	$(8.0\pm2.0)\times10^{16}$	180±20	3.0	56.8	$(2.2\pm0.3)\times10^{17}$	130±20	3.3	56.5
G35.03+0.35	$(2.8\pm0.7)\times10^{16}$	120±40	3.5	44.6	$(6.3\pm1.7)\times10^{16}$	130±20	3.7	45.0
G35.20-0.74N (c1)	$(8.5\pm1.5)\times10^{16}$	160±20	4.0	31.2	$(1.3\pm0.3)\times10^{17}$	130±20	4.0	31.0
G35.20-0.74N (c2)	$(1.8\pm0.4)\times10^{16}$	160±20	1.5	33.7	$(3.5\pm1.0)\times10^{16}$	130±20	1.7	33.7
G345.5+1.5	$(1.4\pm0.3)\times10^{16}$	140±40	2.0	-16.3	$(4.3\pm0.7)\times10^{16}$	120±20	2.0	-16.1
IRAS 18151-1208	$(2.0\pm0.5)\times10^{15}$	[180]	1.9	35.3	$(2.6\pm0.5)\times10^{16}$	110±20	2.0	35.1
IRAS 16547-4247	$(6.0\pm1.5)\times10^{16}$	140±40	3.0	-35.2	$(6.0\pm1.5)\times10^{16}$	180±20	2.7	-35.5
NGC 6334-38	$(1.6\pm0.3)\times10^{16}$	190±50	2.6	-5.2	$(1.5\pm0.3)\times10^{17}$	110±20	3.0	-5.2
NGC 6334-43	$(2.0\pm0.5)\times10^{16}$	180±30	3.0	0.8	$(9.5\pm2.5)\times10^{16}$	120±20	3.0	0.7
NGC 6334 I(N)-SM2	$(1.3\pm0.4)\times10^{16}$	180±20	3.0	-3.0	$(8.0\pm2.0)\times10^{16}$	100±20	4.5	-3.0

Table 2.B.3: Fitting results of CH_3OCHO (MF) and CH_2OHCHO (GA).

	CH_3OCHO (JPL)				CH_2OHCHO (CDMS)			
Source Name	N (cm^{-2})	T_{ex} (K)	FWHM (km s^{-1})	v_{lsr} (cm^{-2})	N^a (K)	T_{ex} (km s^{-1})	FWHM	v_{lsr}
G19.01-0.03 (c1)	$(5.0\pm1.0)\times10^{16}$	180±40	3.5	62.5	$(7.0\pm2.3)\times10^{15}$	<120	3.5	61.5
G19.01-0.03 (c2)	$(6.0\pm1.0)\times10^{16}$	180±40	3.5	58.5	$(7.0\pm2.3)\times10^{15}$	<120	3.5	58.5
G19.88-0.53	$(1.1\pm0.3)\times10^{17}$	180±40	4.0	46.0	$(6.0\pm1.3)\times10^{15}$	160±40	2.5	46.4
G22.04+0.22	$(7.0\pm2.0)\times10^{17}$	130±30	5.5	52.5	$(3.9\pm0.8)\times10^{16}$	150±50	5.0	52.5
G23.21-0.37	$(3.0\pm0.5)\times10^{17}$	170±30	3.3	77.2	$(3.7\pm0.8)\times10^{16}$	200±40	3.5	76.8
G34.30+0.20 (c1)	$(1.6\pm0.3)\times10^{17}$	140±40	3.2	56.5	$(1.0\pm0.3)\times10^{16}$	180±30	3.5	55.5
G34.30+0.20 (c2)	$(5.0\pm1.5)\times10^{16}$	140±40	2.5	60.0				
G34.41+0.24 (c1)	$(1.0\pm0.2)\times10^{17}$	160±20	3.0	60.0				
G34.41+0.24 (c2)	$(1.0\pm0.2)\times10^{17}$	160±20	3.0	56.8	$(1.1\pm0.3)\times10^{16}$	160±40	3.0	57.0
G35.03+0.35	$(4.7\pm1.3)\times10^{16}$	150±50	4.0	45.3	$(2.0\pm0.7\times10^{15}$	[150]	3.5	44.5
G35.20-0.74N (c1)	$(1.5\pm0.3)\times10^{17}$	140±30	4.0	31.2	$(7.4\pm2.5)\times10^{15}$	[140]	3.5	31.2
G35.20-0.74N (c2)	$(2.0\pm0.5)\times10^{16}$	140±30	1.2	33.7				
G345.5+1.5	$(2.1\pm0.4)\times10^{16}$	120±20	2.2	-15.8	$<1.4\times10^{15}$	[120]	2.2	-15.8
IRAS 18151-1208	$(1.4\pm0.2)\times10^{16}$	180±40	2.0	35.2	$<7.4\times10^{14}$	[180]	1.7	35.0
IRAS 16547-4247	$(8.0\pm2.0)\times10^{16}$	200±20	3.3	-35.5	$(1.5\pm0.3)\times10^{16}$	[200]	3.0	-35.5
NGC 6334-38	$(3.0\pm0.5)\times10^{17}$	120±30	3.0	-5.2	$(5.0\pm1.2)\times10^{15}$	140±70	2.8	-5.2
NGC 6334-43	$(8.5\pm2.0)\times10^{16}$	160±40	3.0	0.8	$(2.7\pm0.7)\times10^{15}$	[160]	3.0	0.8
NGC 6334 I(N)-SM2	$(6.0\pm1.0)\times10^{16}$	150±30	5.0	-2.5	$(4.4\pm1.5)\times10^{15}$	180±40	3.0	-2.5

[a]Vibrational corrections applied; see Table 2.B.5 for the uncorrected N.

Table 2.B.4: Fitting results of *a*-$(CH_2OH)_2$ (*a*-EG) and *g*-$(CH_2OH)_2$ (*g*-EG).

	a-$(CH_2OH)_2$ (CDMS)				g-$(CH_2OH)_2$ (CDMS)			
Source Name	N^a (cm^{-2})	T_{ex} (K)	FWHM (km s^{-1})	v_{lsr}	N^a (cm^{-2})	T_{ex} (K)	FWHM (km s^{-1})	v_{lsr}
G19.01-0.03 (c1)	$(4.0\pm0.8)\times10^{16}$	[180]	4.0	62.0	$(2.0\pm0.4)\times10^{16}$	[180]	3.5	61.0
G19.88-0.53	$(5.7\pm1.2)\times10^{16}$	220±20	3.0	46.7	$(2.9\pm0.7)\times10^{16}$	[220]	3.0	46.7
G22.04+0.22	$(5.9\pm1.5)\times10^{16}$	160±40	4.5	52.0	$(4.6\pm0.9)\times10^{16}$	[160]	4.5	52.0
G23.21-0.37	$(1.4\pm0.4)\times10^{17}$	180±40	3.8	77.0	$(1.0\pm0.4)\times10^{17}$	[180]	3.5	76.8
G34.30+0.20 (c1)	$(3.1\pm0.7)\times10^{16}$	[140]	3.5	55.5	$(2.2\pm0.5)\times10^{16}$	[140]	3.5	55.5
G34.41+0.24 (c1)	$(1.8\pm0.5)\times10^{16}$	160±40	3.0	60.8	$(8.2\pm1.8)\times10^{15}$	[160]	3.0	60.0
G35.03+0.35	$(4.3\pm1.0)\times10^{16}$	220±40	3.5	43.5	$(1.9\pm0.7)\times10^{16}$	[220]	3.5	43.5
G35.20-0.74N (c1)	$(5.3\pm1.6)\times10^{16}$	180±40	4.0	30.6	$(2.6\pm0.8)\times10^{16}$	[180]	3.5	30.6
G345.5+1.5	$<3.3\times10^{15}$	[120]	2.0	-15.0	$<1.6\times10^{15}$	[120]	2.0	-15.0
IRAS 18151-1208	$<1.0\times10^{15}$	[180]	2.0	35.0	$<2.0\times10^{15}$	[180]	2.0	35.0
IRAS 16547-4247	$(9.0\pm2.0)\times10^{16}$	180±40	3.2	-35.5	$(5.0\pm1.0)\times10^{16}$	[180]	3.0	-35.7
NGC 6334-38	$(6.5\pm1.6)\times10^{15}$	[120]	2.8	-5.2	$(6.5\pm2.4)\times10^{15}$	[120]	2.8	-5.2
NGC 6334-43	$(2.1\pm0.6)\times10^{16}$	200±40	3.0	0.8	$(1.5\pm0.4)\times10^{16}$	[200]	3.0	0.7
NGC 6334 I(N)-SM2	$(9.0\pm2.0)\times10^{15}$	180±50	4.0	-3.0	$<5.0\times10^{15}$	[180]	3.0	-3.0

[a] Vibrational corrections applied; see Table 2.B.5 for the uncorrected N.

Table 2.B.5: Column densities of CH_2OHCHO (GA) and *a*- and *g*-$(CH_2OH)_2$ (EG) without vibrational corrections.

Source Name	CH_2OHCHO N (cm^{-2})	a-$(CH_2OH)_2$ N (cm^{-2})	g-$(CH_2OH)_2$ N (cm^{-2})
G19.01-0.03 (c1, c2)	$(6.0\pm2.0)\times10^{15}$	$(2.0\pm0.4)\times10^{16}$	$(1.0\pm0.2)\times10^{16}$
G19.88-0.53	$(4.5\pm1.0)\times10^{15}$	$(2.4\pm0.5)\times10^{16}$	$(1.2\pm0.3)\times10^{16}$
G22.04+0.22	$(3.0\pm0.6)\times10^{16}$	$(3.2\pm0.8)\times10^{16}$	$(2.5\pm0.5)\times10^{16}$
G23.21-0.37	$(2.3\pm0.5)\times10^{16}$	$(7.0\pm2.0)\times10^{16}$	$(5.0\pm2.0)\times10^{16}$
G34.30+0.20	$(7.0\pm2.0)\times10^{15}$	$(1.8\pm0.4)\times10^{16}$	$(1.3\pm0.3)\times10^{16}$
G34.41+0.24	$(8.0\pm2.0)\times10^{15}$	$(1.0\pm0.3)\times10^{16}$	$(4.5\pm1.0)\times10^{15}$
G35.03+0.35	$(1.5\pm0.5)\times10^{15}$	$(1.8\pm0.4)\times10^{16}$	$(8.0\pm3.0)\times10^{15}$
G35.20-0.74N	$(6.0\pm2.0)\times10^{15}$	$(2.7\pm0.8)\times10^{16}$	$(1.3\pm0.4)\times10^{16}$
G345.5+1.5	$<1.2\times10^{15}$	$<2.0\times10^{15}$	$<1.0\times10^{15}$
IRAS 18151-1208	$<5.0\times10^{14}$	$<5.0\times10^{14}$	$<1.0\times10^{15}$
IRAS 16547-4247	$(9.0\pm2.0)\times10^{15}$	$(4.5\pm1.0)\times10^{16}$	$(2.5\pm0.5)\times10^{16}$
NGC 6334-38	$(4.0\pm1.0)\times10^{15}$	$(4.0\pm1.0)\times10^{15}$	$(4.0\pm1.5)\times10^{15}$
NGC 6334-43	$(2.0\pm0.5)\times10^{15}$	$(1.0\pm0.3)\times10^{16}$	$(7.0\pm2.0)\times10^{15}$
NGC 6334 I(N)-SM2	$(3.0\pm1.0)\times10^{15}$	$(4.5\pm1.0)\times10^{15}$	$<2.5\times10^{15}$

Table 2.B.6: Identified transitions of O-COMs.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$
$^{13}CH_3OH$	16 3 14 +0 – 16 2 15 -0	258153.004	358.01	9.03(-5)
(vt=0, 1)	19 5 15 +0 – 20 4 16 +0	258460.830	568.65	2.95(-5)
CDMS	19 5 14 -0 – 20 4 17 -0	258521.376	568.65	2.95(-5)
	17 3 15 +0 – 17 2 16 -0	259036.489	396.48	9.11(-5)
	12 -2 10 1 – 13 -3 10 1	259575.198	540.96	7.21(-5)
	2 1 1 0 – 1 0 1 0	259986.530	27.85	5.46(-5)
	18 3 16 +0 – 18 2 17 -0	260088.839	437.21	9.21(-5)
	24 -3 21 1 – 23 -2 21 1	260637.531	999.80	6.76(-5)
	19 3 17 +0 – 19 2 18 -0	261326.838	480.20	9.31(-5)
$CH_3^{18}OH$	19 3 16 0 – 19 2 17 0	258436.476	472.65	9.22(-5)
(v=0–2)	15 1 14 0 – 14 2 13 0	258557.479	283.36	4.05(-5)
CDMS	18 3 15 0 – 18 2 16 0	259133.097	430.41	9.24(-5)
	5 4 2 1 – 6 3 4 1	259391.755	121.80	8.35(-6)
	17 3 14 0 – 17 2 15 0	259759.660	390.39	9.25(-5)
	17 2 15 2 – 16 3 13 2	259873.428	362.98	2.95(-5)
	12 2 10 4 – 13 3 10 4	260035.462	537.16	7.08(-5)
	16 3 13 0 – 16 2 14 0	260316.321	352.59	9.26(-5)
	15 1 15 1 – 14 2 13 1	260440.344	272.16	2.76(-5)
	15 3 12 0 – 15 2 13 0	260804.867	317.02	9.26(-5)
	15 4 11 2 – 16 3 13 2	261102.783	363.04	3.00(-5)
	14 3 11 0 – 14 2 12 0	261228.369	283.66	9.25(-5)
	13 3 10 0 – 13 2 11 0	261590.879	252.53	9.24(-5)
CH_3CHO	19 3 17 1 – 19 2 18 1	259298.775	196.31	5.60(-5)
JPL	13 1 12 6 – 12 1 11 6	259375.784	459.31	5.70(-4)
	13 2 12 7 – 12 2 11 7	259761.410	474.77	5.34(-4)
	19 3 17 0 – 19 2 18 0	260283.405	196.35	5.69(-5)
	13 1 13 1 – 12 0 12 2	260408.016	83.89	7.68(-5)
	14 1 14 1 – 13 1 13 1	260530.403	96.39	6.02(-4)
	14 1 14 0 – 13 1 13 0	260544.019	96.32	6.01(-4)
	13 1 13 0 – 12 0 12 0	260694.002	83.82	7.73(-5)
	14 1 14 3 – 13 1 13 3	260826.516	302.82	6.12(-4)
C_2H_5OH	19 2 18 1 – 18 3 16 0	258099.207	222.51	1.42(-5)
(v=0)	16 4 13 1 – 16 3 13 0	258449.274	194.44	2.12(-5)
CDMS	25 1 24 2 – 25 0 25 2	258484.664	272.70	3.90(-5)
	14 3 11 0 – 13 2 11 1	259322.639	155.72	7.25(-5)
	15 12 3 1 – 14 12 2 1	259472.683	337.99	5.52(-5)
	15 12 4 1 – 14 12 3 1	259472.683	337.99	5.52(-5)
	15 13 2 1 – 14 13 1 1	259477.660	368.56	3.81(-5)
	15 13 3 1 – 14 13 2 1	259477.660	368.56	3.81(-5)
	15 11 4 1 – 14 11 3 1	259479.315	309.83	7.09(-5)
	15 11 5 1 – 14 11 4 1	259479.315	309.83	7.09(-5)
	15 10 5 1 – 14 10 4 1	259501.052	284.11	8.53(-5)
	15 10 6 1 – 14 10 5 1	259501.052	284.11	8.53(-5)

Table 2.B.6: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$
	15 9 6 1 – 14 9 5 1	259539.131	260.82	9.82(-5)
	15 9 7 1 – 14 9 6 1	259539.131	260.82	9.82(-5)
	15 8 8 1 – 14 8 7 1	259601.059	239.96	1.10(-4)
	15 8 7 1 – 14 8 6 1	259601.060	239.96	1.10(-4)
	15 7 9 1 – 14 7 8 1	259697.897	221.55	1.20(-4)
	15 7 8 1 – 14 7 7 1	259697.903	221.55	1.20(-4)
	13 1 12 1 – 12 2 10 0	259744.133	140.01	3.11(-5)
	18 4 15 1 – 18 3 15 0	259748.793	223.63	3.48(-5)
	15 11 4 0 – 14 11 3 0	259751.561	304.31	7.53(-5)
	15 11 5 0 – 14 11 4 0	259751.561	304.31	7.53(-5)
	15 12 3 0 – 14 12 2 0	259754.443	332.42	5.88(-5)
	15 12 4 0 – 14 12 3 0	259754.443	332.42	5.88(-5)
	15 10 5 0 – 14 10 4 0	259756.535	278.65	9.03(-5)
	15 10 6 0 – 14 10 5 0	259756.535	278.65	9.03(-5)
	15 13 2 0 – 14 13 1 0	259763.493	362.98	4.08(-5)
	15 13 3 0 – 14 13 2 0	259763.493	362.98	4.08(-5)
	15 9 6 0 – 14 9 5 0	259777.143	255.43	1.04(-4)
	15 9 7 0 – 14 9 6 0	259777.143	255.43	1.04(-4)
	15 14 1 0 – 14 14 0 0	259778.002	395.98	2.13(-5)
	15 14 2 0 – 14 14 1 0	259778.002	395.98	2.13(-5)
	15 8 8 0 – 14 8 7 0	259814.437	234.67	1.16(-4)
	15 8 7 0 – 14 8 6 0	259814.438	234.67	1.16(-4)
	32 5 28 2 – 32 4 29 2	259841.221	477.18	1.17(-4)
	15 6 10 1 – 14 6 9 1	259852.185	205.59	1.29(-4)
	15 6 9 1 – 14 6 8 1	259852.573	205.59	1.29(-4)
	15 7 9 0 – 14 7 8 0	259885.072	216.37	1.27(-4)
	15 7 8 0 – 14 7 7 0	259885.079	216.37	1.27(-4)
	25 2 24 2 – 25 1 25 2	260007.736	272.77	3.96(-5)
	15 6 10 0 – 14 6 9 0	260013.579	200.53	1.36(-4)
	15 6 9 0 – 14 6 8 0	260014.001	200.53	1.36(-4)
	15 3 13 1 – 14 3 12 1	260046.644	172.46	1.46(-4)
	15 1 14 0 – 14 1 13 0	260090.201	159.58	1.75(-4)
	15 5 11 1 – 14 5 10 1	260107.609	192.10	1.37(-4)
	15 5 10 1 – 14 5 9 1	260122.761	192.10	1.37(-4)
	15 3 13 0 – 14 3 12 0	260141.678	167.72	1.59(-4)
	15 5 11 0 – 14 5 10 0	260249.761	187.15	1.45(-4)
	15 5 10 0 – 14 5 9 0	260266.127	187.15	1.45(-4)
	7 1 6 1 – 6 0 6 0	260437.554	86.70	5.46(-5)
	15 4 12 1 – 14 4 11 1	260457.726	181.10	1.43(-4)
	15 4 12 0 – 14 4 11 0	260591.330	176.26	1.52(-4)
	31 4 28 2 – 30 5 25 2	260627.301	437.93	1.86(-5)
	15 4 11 1 – 14 4 10 1	260796.868	181.14	1.43(-4)
	15 4 11 0 – 14 4 10 0	260960.989	176.30	1.53(-4)

Table 2.B.6: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$
	20 3 18 1 – 19 4 16 0	261282.323	247.21	2.17(-5)
	15 1 14 1 – 14 1 13 1	261286.306	164.30	1.79(-4)
CH_3OCH_3	30 5 26 3 – 29 6 23 3	258197.879	459.88	2.47(-5)
(v=0)	30 5 26 5 – 29 6 23 5	258197.880	459.88	2.47(-5)
CDMS	30 5 26 1 – 29 6 23 1	258198.172	459.88	2.47(-5)
	30 5 26 0 – 29 6 23 0	258198.464	459.88	2.47(-5)
	30 14 17 3 – 31 15 17 3	258278.183	738.73	1.24(-5)
	14 1 14 3 – 13 0 13 3	258548.819	93.33	1.31(-4)
	14 1 14 5 – 13 0 13 5	258548.819	93.33	1.31(-4)
	14 1 14 1 – 13 0 13 1	258549.063	93.33	1.31(-4)
	14 1 14 0 – 13 0 13 0	258549.308	93.33	1.31(-4)
	17 5 12 5 – 17 4 13 5	259309.472	174.54	8.77(-5)
	17 5 12 3 – 17 4 13 3	259309.758	174.54	8.74(-5)
	17 5 12 1 – 17 4 13 1	259311.947	174.54	8.76(-5)
	17 5 12 0 – 17 4 13 0	259314.279	174.54	8.77(-5)
	6 3 4 3 – 5 2 3 3	259484.856	31.77	7.40(-5)
	6 3 4 5 – 5 2 3 5	259486.616	31.77	7.65(-5)
	6 3 4 1 – 5 2 3 1	259489.733	31.77	7.58(-5)
	6 3 4 0 – 5 2 3 0	259493.749	31.77	7.65(-5)
	22 5 18 3 – 22 4 19 3	259615.889	265.87	9.34(-5)
	22 5 18 5 – 22 4 19 5	259615.896	265.87	9.34(-5)
	22 5 18 1 – 22 4 19 1	259617.339	265.87	9.34(-5)
	22 5 18 0 – 22 4 19 0	259618.785	265.87	9.34(-5)
	23 5 19 3 – 23 4 20 3	259688.831	286.90	9.44(-5)
	23 5 19 5 – 23 4 20 5	259688.835	286.90	9.44(-5)
	23 5 19 1 – 23 4 20 1	259690.065	286.90	9.44(-5)
	23 5 19 0 – 23 4 20 0	259691.297	286.90	9.44(-5)
	21 5 17 3 – 21 4 18 3	259730.487	245.76	9.26(-5)
	21 5 17 5 – 21 4 18 5	259730.502	245.76	9.26(-5)
	21 5 17 1 – 21 4 18 1	259732.149	245.76	9.26(-5)
	21 5 17 0 – 21 4 18 0	259733.804	245.76	9.26(-5)
	33 6 28 3 – 32 7 25 3	259816.557	562.74	2.54(-5)
	33 6 28 5 – 32 7 25 5	259816.583	562.74	2.54(-5)
	33 6 28 1 – 32 7 25 1	259817.273	562.74	2.54(-5)
	33 6 28 0 – 32 7 25 0	259817.976	562.74	2.54(-5)
	20 5 16 3 – 20 4 17 3	259982.535	226.58	9.18(-5)
	20 5 16 5 – 20 4 17 5	259982.568	226.58	9.18(-5)
	20 5 16 1 – 20 4 17 1	259984.408	226.58	9.18(-5)
	20 5 16 0 – 20 4 17 0	259986.265	226.58	9.18(-5)
	24 5 20 3 – 24 4 21 3	260003.375	308.85	9.55(-5)
	24 5 20 5 – 24 4 21 5	260003.377	308.85	9.55(-5)
	24 5 20 1 – 24 4 21 1	260004.389	308.85	9.55(-5)
	24 5 20 0 – 24 4 21 0	260005.402	308.85	9.55(-5)

Table 2.B.6: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$
	19 5 15 3 – 19 4 16 3	260327.133	208.31	9.10(-5)
	19 5 15 5 – 19 4 16 5	260327.201	208.31	9.10(-5)
	19 5 15 1 – 19 4 16 1	260329.221	208.31	9.10(-5)
	19 5 15 0 – 19 4 16 0	260331.275	208.31	9.10(-5)
	16 5 11 5 – 16 4 12 5	260400.539	159.03	8.73(-5)
	16 5 11 3 – 16 4 12 3	260401.135	159.03	8.64(-5)
	16 5 11 1 – 16 4 12 1	260403.244	159.03	8.71(-5)
	16 5 11 0 – 16 4 12 0	260405.649	159.03	8.73(-5)
	25 5 21 3 – 25 4 22 3	260616.059	331.72	9.68(-5)
	25 5 21 5 – 25 4 22 5	260616.059	331.72	9.68(-5)
	25 5 21 1 – 25 4 22 1	260616.851	331.72	9.68(-5)
	25 5 21 0 – 25 4 22 0	260617.643	331.72	9.68(-5)
	18 5 14 3 – 18 4 15 3	260725.448	190.97	9.02(-5)
	18 5 14 5 – 18 4 15 5	260725.587	190.97	9.02(-5)
	18 5 14 1 – 18 4 15 1	260727.767	190.97	9.02(-5)
	18 5 14 0 – 18 4 15 0	260730.017	190.97	9.02(-5)
	6 3 3 5 – 5 2 4 5	260754.380	31.77	7.71(-5)
	6 3 3 3 – 5 2 4 3	260756.140	31.77	7.47(-5)
	6 3 3 1 – 5 2 4 1	260758.402	31.77	7.65(-5)
	6 3 3 0 – 5 2 4 0	260761.524	31.77	7.72(-5)
	17 5 13 3 – 17 4 14 3	261145.206	174.54	8.91(-5)
	17 5 13 5 – 17 4 14 5	261145.492	174.54	8.94(-5)
	17 5 13 1 – 17 4 14 1	261147.803	174.54	8.93(-5)
	17 5 13 0 – 17 4 14 0	261150.257	174.54	8.94(-5)
	15 5 10 5 – 15 4 11 5	261245.097	144.44	8.66(-5)
	15 5 10 3 – 15 4 11 3	261246.331	144.44	8.30(-5)
	15 5 10 1 – 15 4 11 1	261248.113	144.44	8.56(-5)
	15 5 10 0 – 15 4 11 0	261250.488	144.44	8.66(-5)
	16 5 12 3 – 16 4 13 3	261560.797	159.03	8.74(-5)
	16 5 12 5 – 16 4 13 5	261561.393	159.03	8.84(-5)
	16 5 12 1 – 16 4 13 1	261563.781	159.03	8.81(-5)
	16 5 12 0 – 16 4 13 0	261566.471	159.03	8.84(-5)
	26 5 22 3 – 26 4 23 3	261584.210	355.51	9.84(-5)
	26 5 22 5 – 26 4 23 5	261584.210	355.51	9.84(-5)
	26 5 22 1 – 26 4 23 1	261584.781	355.51	9.84(-5)
	26 5 22 0 – 26 4 23 0	261585.353	355.51	9.84(-5)
CH_3OCHO	21 8 13 3 – 20 8 12 3	257906.128	365.98	2.22(-4)
JPL	21 16 5 0 – 20 16 4 0	257910.566	306.01	1.09(-4)
	21 16 6 0 – 20 16 5 0	257910.566	306.01	1.09(-4)
	21 16 5 2 – 20 16 4 2	257919.890	306.01	1.09(-4)
	21 16 6 1 – 20 16 5 1	257933.830	306.00	1.09(-4)
	21 15 6 0 – 20 15 5 0	258001.757	285.47	1.28(-4)
	21 15 7 0 – 20 15 6 0	258001.757	285.47	1.28(-4)

Table 2.B.6: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$
	21 15 6 2 – 20 15 5 2	258007.150	285.47	1.28(-4)
	24 0 24 3 – 23 1 23 3	258010.378	345.26	4.19(-5)
	24 1 24 3 – 23 1 23 3	258010.754	345.26	2.59(-4)
	24 0 24 3 – 23 0 23 3	258010.754	345.26	2.59(-4)
	24 1 24 3 – 23 0 23 3	258011.019	345.26	4.19(-5)
	21 15 7 1 – 20 15 6 1	258024.240	285.46	1.28(-4)
	21 9 13 4 – 20 9 12 4	258037.974	376.94	2.13(-4)
	24 10 15 1 – 24 9 16 1	258041.077	243.70	2.22(-5)
	24 10 14 2 – 24 9 15 2	258049.562	243.71	2.22(-5)
	24 10 14 0 – 24 9 15 0	258052.000	243.70	2.23(-5)
	24 0 24 5 – 23 1 23 4	258054.742	344.52	4.10(-5)
	24 1 24 4 – 23 1 23 4	258055.043	344.52	2.60(-4)
	24 0 24 5 – 23 0 23 5	258055.043	344.52	2.60(-4)
	24 1 24 4 – 23 0 23 5	258055.296	344.52	4.10(-5)
	24 10 15 0 – 24 9 16 0	258072.452	243.70	2.23(-5)
	22 2 20 2 – 21 2 19 2	258081.042	152.23	2.52(-4)
	22 2 20 0 – 21 2 19 0	258089.491	152.22	2.52(-4)
	21 14 7 0 – 20 14 6 0	258121.191	266.26	1.45(-4)
	21 14 8 0 – 20 14 7 0	258121.191	266.26	1.45(-4)
	21 14 7 2 – 20 14 6 2	258122.660	266.25	1.45(-4)
	21 14 8 1 – 20 14 7 1	258142.090	266.25	1.45(-4)
	21 13 8 2 – 20 13 7 2	258274.950	248.38	1.61(-4)
	21 13 8 0 – 20 13 7 0	258277.434	248.38	1.61(-4)
	21 13 9 0 – 20 13 8 0	258277.434	248.38	1.61(-4)
	21 13 9 1 – 20 13 8 1	258296.300	248.36	1.61(-4)
	11 5 7 0 – 10 4 6 0	258306.279	55.60	1.92(-5)
	27 10 17 3 – 27 9 18 3	258380.810	476.33	2.37(-5)
	27 10 18 4 – 27 9 19 4	258425.612	476.33	2.36(-5)
	18 4 15 4 – 17 3 14 5	258450.614	298.62	1.49(-5)
	23 1 22 2 – 22 2 21 1	258475.052	155.91	3.57(-5)
	21 12 9 2 – 20 12 8 2	258476.450	231.83	1.76(-4)
	23 1 22 0 – 22 2 21 0	258480.586	155.90	3.57(-5)
	21 12 9 0 – 20 12 8 0	258482.981	231.83	1.76(-4)
	21 12 10 0 – 20 12 9 0	258482.981	231.83	1.76(-4)
	23 2 22 1 – 22 2 21 1	258490.870	155.91	2.57(-4)
	27 10 18 3 – 27 9 19 3	258493.927	476.33	2.38(-5)
	23 2 22 0 – 22 2 21 0	258496.242	155.90	2.57(-4)
	21 12 10 1 – 20 12 9 1	258499.332	231.82	1.76(-4)
	23 1 22 2 – 22 1 21 2	258502.735	155.91	2.57(-4)
	23 1 22 0 – 22 1 21 0	258508.181	155.90	2.57(-4)
	23 2 22 1 – 22 1 21 2	258518.554	155.91	3.57(-5)
	23 2 22 0 – 22 1 21 0	258523.821	155.90	3.57(-5)
	41 8 34 0 – 41 7 35 0	258541.136	557.04	2.34(-5)

Table 2.B.6: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$
	21 5 17 3 – 20 5 16 3	258701.047	340.65	2.46(-4)
	22 3 20 1 – 21 2 19 2	258706.904	152.26	2.85(-5)
	22 3 20 0 – 21 2 19 0	258714.075	152.25	2.85(-5)
	21 11 10 2 – 20 11 9 2	258746.248	216.63	1.91(-4)
	21 11 11 0 – 20 11 10 0	258756.673	216.63	1.91(-4)
	21 11 10 0 – 20 11 9 0	258756.673	216.63	1.91(-4)
	21 7 15 3 – 20 7 14 3	258768.938	356.20	2.33(-4)
	21 11 11 1 – 20 11 10 1	258769.974	216.62	1.91(-4)
	21 3 18 3 – 20 3 17 3	258775.320	333.28	2.52(-4)
	21 8 14 4 – 20 8 13 4	258783.896	365.70	2.25(-4)
	23 10 14 1 – 23 9 15 1	258837.620	229.45	2.20(-5)
	23 10 13 2 – 23 9 14 2	258847.920	229.46	2.20(-5)
	23 10 13 0 – 23 9 14 0	258859.163	229.46	2.20(-5)
	23 10 14 0 – 23 9 15 0	258868.739	229.46	2.20(-5)
	11 5 7 4 – 10 4 7 4	258955.739	242.40	1.97(-5)
	21 7 14 3 – 20 7 13 3	259003.875	356.22	2.34(-4)
	21 7 14 5 – 20 7 13 5	259025.827	356.29	2.34(-4)
	21 10 11 2 – 20 10 10 2	259113.950	202.78	2.04(-4)
	21 10 12 0 – 20 10 11 0	259128.178	202.78	2.04(-4)
	21 10 11 0 – 20 10 10 0	259128.178	202.78	2.04(-4)
	21 10 12 1 – 20 10 11 1	259137.930	202.77	2.04(-4)
	21 3 18 5 – 20 3 17 5	259264.990	333.04	2.54(-4)
	34 4 30 0 – 34 3 31 0	259299.779	366.21	1.61(-5)
	24 0 24 2 – 23 1 23 1	259341.865	158.23	4.24(-5)
	24 1 24 1 – 23 1 23 1	259342.015	158.23	2.63(-4)
	24 0 24 2 – 23 0 23 2	259342.143	158.23	2.63(-4)
	24 1 24 1 – 23 0 23 2	259342.293	158.23	4.24(-5)
	24 0 24 0 – 23 1 23 0	259342.727	158.22	4.24(-5)
	24 1 24 0 – 23 1 23 0	259342.876	158.22	2.63(-4)
	24 0 24 0 – 23 0 23 0	259343.004	158.22	2.63(-4)
	24 1 24 0 – 23 0 23 0	259343.152	158.22	4.24(-5)
	11 5 6 0 – 10 4 7 0	259376.253	55.61	1.94(-5)
	42 7 35 2 – 42 6 36 2	259422.473	581.45	2.31(-5)
	11 5 6 2 – 10 4 7 1	259445.549	55.62	1.08(-5)
	26 10 17 3 – 26 9 18 3	259455.455	460.41	2.36(-5)
	42 7 35 0 – 42 6 36 0	259463.004	581.46	2.31(-5)
	20 4 16 2 – 19 4 15 2	259499.905	138.67	2.54(-4)
	20 4 16 0 – 19 4 15 0	259521.812	138.67	2.54(-4)
	22 10 13 1 – 22 9 14 1	259529.940	215.81	2.16(-5)
	22 3 19 2 – 21 4 18 1	259536.089	159.72	2.05(-5)
	22 10 12 2 – 22 9 13 2	259540.700	215.82	2.16(-5)
	22 10 12 0 – 22 9 13 0	259558.582	215.81	2.16(-5)
	22 3 19 0 – 21 4 18 0	259560.402	159.72	2.05(-5)

Table 2.B.6: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$
	22 10 13 0 – 22 9 14 0	259562.851	215.81	2.16(-5)
	21 5 17 4 – 20 5 16 4	259573.430	340.45	2.49(-4)
	34 5 30 1 – 34 4 31 1	259592.949	366.22	1.62(-5)
	26 10 17 4 – 26 9 18 4	259593.370	460.37	2.35(-5)
	21 6 16 3 – 20 6 15 3	259616.603	347.84	2.43(-4)
	21 9 12 2 – 20 9 11 2	259629.300	190.30	2.16(-4)
	21 9 13 0 – 20 9 12 0	259646.531	190.29	2.17(-4)
	21 9 12 0 – 20 9 11 0	259647.705	190.29	2.17(-4)
	21 9 13 1 – 20 9 12 1	259653.078	190.28	2.17(-4)
	34 5 30 0 – 34 4 31 0	259662.604	366.23	1.62(-5)
	21 7 15 4 – 20 7 14 4	259916.051	355.90	2.37(-4)
	21 10 12 1 – 21 9 13 1	260128.610	202.77	2.12(-5)
	21 10 11 2 – 21 9 12 2	260139.772	202.78	2.12(-5)
	21 10 11 0 – 21 9 12 0	260162.612	202.78	2.12(-5)
	21 10 12 0 – 21 9 13 0	260164.454	202.78	2.12(-5)
	21 3 18 2 – 20 3 17 2	260244.498	146.76	2.56(-4)
	21 3 18 0 – 20 3 17 0	260255.080	146.75	2.56(-4)
	25 10 16 3 – 25 9 17 3	260296.709	445.09	2.34(-5)
	21 8 13 2 – 20 8 12 2	260384.268	179.20	2.28(-4)
	21 8 14 0 – 20 8 13 0	260392.731	179.19	2.29(-4)
	21 8 14 1 – 20 8 13 1	260404.026	179.19	2.28(-4)
	21 8 13 0 – 20 8 12 0	260415.333	179.19	2.29(-4)
	20 10 11 1 – 20 9 12 1	260643.800	190.33	2.07(-5)
	20 10 10 2 – 20 9 11 2	260655.151	190.34	2.07(-5)
	20 10 10 0 – 20 9 11 0	260682.101	190.34	2.07(-5)
	20 10 11 0 – 20 9 12 0	260682.850	190.34	2.07(-5)
	19 4 16 0 – 18 3 15 0	260788.406	123.25	1.70(-5)
	19 4 16 1 – 18 3 15 2	260793.540	123.26	1.70(-5)
	11 5 6 5 – 10 4 6 5	260799.899	242.97	1.90(-5)
	32 3 29 2 – 32 2 30 2	260916.838	316.40	1.31(-5)
	32 3 29 0 – 32 2 30 0	261005.113	316.40	1.31(-5)
	32 4 29 0 – 32 3 30 0	261070.724	316.40	1.31(-5)
	19 10 10 1 – 19 9 11 1	261084.120	178.50	2.01(-5)
	19 10 9 2 – 19 9 10 2	261095.680	178.51	2.01(-5)
	19 10 9 0 – 19 9 10 0	261126.435	178.50	2.01(-5)
	19 10 10 0 – 19 9 11 0	261126.726	178.50	2.01(-5)
	21 5 17 1 – 20 5 16 1	261148.904	154.10	2.54(-4)
	21 5 17 0 – 20 5 16 0	261165.456	154.09	2.54(-4)
	21 6 16 4 – 20 6 15 4	261234.608	347.62	2.43(-4)
	21 7 15 0 – 20 7 14 0	261433.791	169.52	2.41(-4)
	21 7 15 1 – 20 7 14 1	261436.771	169.52	2.36(-4)
	18 10 9 1 – 18 9 10 1	261458.040	167.26	1.94(-5)
	18 10 8 2 – 18 9 9 2	261469.927	167.28	1.94(-5)

Table 2.B.6: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$
	18 10 8 0 – 18 9 9 0	261504.039	167.27	1.94(-5)
	18 10 9 0 – 18 9 10 0	261504.039	167.27	1.94(-5)
	21 7 14 2 – 20 7 13 2	261715.518	169.55	2.37(-4)
	21 6 15 5 – 20 6 14 5	261727.149	348.04	2.45(-4)
CH_2OHCHO	29 11 18 0 – 29 10 19 0	258059.438	315.77	2.78(-4)
(v=0)	29 11 19 0 – 29 10 20 0	258123.981	315.77	2.78(-4)
CDMS	22 4 19 0 – 21 3 18 0	258128.769	150.22	2.53(-4)
	40 6 34 0 – 40 5 35 0	259147.668	488.73	2.40(-4)
	28 11 17 0 – 28 10 18 0	259179.986	299.53	2.77(-4)
	28 11 18 0 – 28 10 19 0	259212.684	299.53	2.77(-4)
	27 11 16 0 – 27 10 17 0	260165.716	283.86	2.75(-4)
	27 11 17 0 – 27 10 18 0	260181.761	283.86	2.75(-4)
	40 7 34 0 – 40 6 35 0	260691.347	488.81	2.44(-4)
	26 11 15 0 – 26 10 16 0	261033.098	268.76	2.72(-4)
	26 11 16 0 – 26 10 17 0	261040.709	268.76	2.72(-4)
	24 4 20 0 – 23 5 19 0	261041.952	182.31	1.89(-4)
	24 2 22 0 – 23 3 21 0	261257.959	167.42	3.73(-4)
	24 3 22 0 – 23 2 21 0	261655.476	167.43	3.75(-4)
a-$(CH_2OH)_2$	27 3 25 0 – 26 3 24 1	258184.073	186.63	4.19(-4)
CDMS	27 2 25 0 – 26 2 24 1	258250.590	186.62	4.10(-4)
	27 1 26 0 – 26 2 25 0	258567.278	179.35	7.15(-5)
	27 1 26 1 – 26 2 25 1	258575.594	179.69	7.16(-5)
	27 2 26 0 – 26 1 25 0	258579.070	179.35	7.11(-5)
	27 2 26 1 – 26 1 25 1	258586.940	179.69	7.21(-5)
	26 6 21 0 – 25 6 19 0	259084.460	191.35	3.28(-5)
	26 23 3 0 – 25 23 2 1	259084.899	431.87	9.23(-5)
	26 23 4 0 – 25 23 3 1	259084.899	431.87	9.23(-5)
	26 21 5 0 – 25 21 4 1	259177.002	388.80	1.48(-4)
	26 21 6 0 – 25 21 5 1	259177.002	388.80	1.48(-4)
	27 1 27 1 – 26 1 26 0	259202.265	171.76	4.28(-4)
	27 0 27 1 – 26 0 26 0	259202.322	171.76	4.28(-4)
	11 8 3 0 – 10 7 3 1	259222.484	63.87	1.55(-5)
	11 8 4 0 – 10 7 4 1	259222.486	63.87	1.55(-5)
	26 20 6 0 – 25 20 5 1	259229.051	368.72	1.74(-4)
	26 20 7 0 – 25 20 6 1	259229.051	368.72	1.74(-4)
	26 19 7 0 – 25 19 6 1	259286.996	349.62	1.98(-4)
	26 19 8 0 – 25 19 7 1	259286.996	349.62	1.98(-4)
	26 18 8 0 – 25 18 7 1	259352.696	331.48	2.22(-4)
	26 18 9 0 – 25 18 8 1	259352.696	331.48	2.22(-4)
	26 17 9 0 – 25 17 8 1	259428.525	314.32	2.44(-4)
	26 17 10 0 – 25 17 9 1	259428.525	314.32	2.44(-4)
	26 16 10 0 – 25 16 9 1	259517.576	298.14	2.65(-4)
	26 16 11 0 – 25 16 10 1	259517.576	298.14	2.65(-4)

Table 2.B.6: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$
	26 15 11 0 – 25 15 10 1	259623.977	282.95	2.85(-4)
	26 15 12 0 – 25 15 11 1	259623.977	282.95	2.85(-4)
	26 14 12 0 – 25 14 11 1	259753.385	268.74	3.03(-4)
	26 14 13 0 – 25 14 12 1	259753.385	268.74	3.04(-4)
	25 5 20 0 – 24 5 19 1	259810.009	174.98	4.14(-4)
	26 3 23 0 – 25 4 22 0	259898.267	180.03	1.90(-4)
	26 13 13 0 – 25 13 12 1	259913.769	255.51	3.21(-4)
	26 13 14 0 – 25 13 13 1	259913.769	255.51	3.21(-4)
	26 12 15 0 – 25 12 14 1	260116.738	243.28	3.38(-4)
	26 12 14 0 – 25 12 13 1	260116.738	243.28	3.38(-4)
	26 11 16 0 – 25 11 15 1	260379.902	232.05	3.53(-4)
	26 11 15 0 – 25 11 14 1	260379.910	232.05	3.53(-4)
	24 4 20 1 – 23 4 19 0	260662.054	159.41	4.03(-4)
	26 10 17 0 – 25 10 16 1	260731.312	221.83	3.68(-4)
	26 10 16 0 – 25 10 15 1	260731.522	221.83	3.68(-4)
	25 4 22 1 – 24 4 21 0	260738.168	167.90	1.45(-4)
	28 2 27 0 – 27 2 26 1	260759.249	192.20	4.32(-4)
	28 1 27 0 – 27 1 26 1	260760.891	192.20	4.31(-4)
	27 4 23 1 – 26 5 22 1	260993.665	198.65	9.51(-5)
	26 5 22 0 – 25 5 21 1	261181.033	185.84	3.79(-4)
	26 9 18 0 – 25 9 17 1	261217.419	212.64	3.82(-4)
	26 9 17 0 – 25 9 16 1	261221.825	212.64	3.82(-4)
	28 0 28 0 – 27 1 27 0	261304.289	183.96	8.62(-5)
	28 1 28 0 – 27 0 27 0	261304.396	183.96	8.62(-5)
	28 0 28 1 – 27 1 27 1	261316.502	184.30	8.64(-5)
	28 1 28 1 – 27 0 27 1	261316.604	184.30	8.64(-5)
g-$(CH_2OH)_2$ CDMS	19 3 16 0 – 18 2 16 1	258132.256	99.64	8.77(-5)
	25 6 20 1 – 24 6 19 0	258140.799	177.14	1.49(-4)
	25 4 22 1 – 24 3 21 1	258146.236	166.51	5.54(-5)
	25 7 18 1 – 24 7 17 0	258193.293	183.22	1.40(-4)
	25 4 22 0 – 24 3 21 0	258228.389	166.46	1.24(-4)
	27 4 23 0 – 26 5 22 0	258328.455	196.89	5.30(-5)
	27 4 23 1 – 26 5 22 1	258376.787	196.94	6.77(-5)
	27 2 26 1 – 26 2 25 0	258670.537	178.47	1.62(-4)
	27 1 26 1 – 26 1 25 0	258673.113	178.47	1.65(-4)
	28 1 28 0 – 27 1 27 1	258855.869	183.15	1.66(-4)
	28 0 28 0 – 27 0 27 1	258855.890	183.15	1.66(-4)
	17 2 16 1 – 16 1 16 0	258880.434	74.76	3.25(-5)
	25 6 19 0 – 24 6 18 1	259076.676	177.52	1.52(-4)
	26 4 23 0 – 25 4 22 1	259858.291	178.98	9.58(-5)
	28 0 28 1 – 27 1 27 1	260191.593	183.21	1.70(-4)
	28 1 28 1 – 27 0 27 1	260191.662	183.21	1.70(-4)
	28 0 28 0 – 27 1 27 0	260193.105	183.15	1.70(-4)

Table 2.B.6: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$
	28 1 28 0 – 27 0 27 0	260193.175	183.15	1.70(-4)
	19 3 16 1 – 18 2 16 0	260466.788	99.70	9.11(-5)
	25 4 21 0 – 24 4 20 1	260784.969	170.44	1.63(-4)
	25 6 19 1 – 24 6 18 0	261211.160	177.57	1.54(-4)
	26 3 23 0 – 25 3 22 1	261231.929	178.81	2.07(-4)
	18 5 14 1 – 17 4 13 1	261293.071	96.03	4.57(-5)
	18 5 14 0 – 17 4 13 0	261378.330	95.98	6.03(-5)
	28 1 28 1 – 27 1 27 0	261528.877	183.21	1.71(-4)
	28 0 28 1 – 27 0 27 0	261528.899	183.21	1.72(-4)
	26 4 23 1 – 25 4 22 0	261644.038	179.02	2.06(-4)

3 CoCCoA: Complex Chemistry in hot Cores with ALMA

The chemical evolution of acetone from ice to gas

Y. Chen, R. T. Garrod, M. Rachid, E. F. van Dishoeck, C. L. Brogan, R. Loomis, A. Lipnicky, and B. A. McGuire

Astronomy & Astrophysics, 696, A198 (2025)

Abstract

Context. Acetone (CH_3COCH_3) is one of the most abundant three-carbon oxygen-bearing complex organic molecules (O-COMs) that have been detected in space. The previous detections were made in the gas phase toward star-forming regions that are chemically rich, mostly in protostellar systems. Recently, acetone ice has also been reported as (tentatively) detected toward two low-mass protostars, allowing comparisons in acetone abundances between gas and ice. The detection of acetone ice warrants a more systematic study of its gaseous abundances which is currently lacking.

Aims. We aim to measure the gas-phase abundances of acetone in a large sample obtained from the CoCCoA program, and investigate the chemical evolution of acetone from ice to gas in protostellar systems.

Methods. We fit the ALMA spectra to determine the column density, excitation temperature, and line width of acetone in 12 high-mass protostars as part of CoCCoA. We also constrained the physical properties of propanal (C_2H_5CHO), ketene (CH_2CO), and propyne (CH_3CCH), which might be chemically linked with acetone. We discuss the possible formation pathways of acetone by making comparisons in its abundances between gas and ice and between observations and simulations.

Results. We firmly detect acetone, ketene, and propyne in the 12 high-mass protostars. The observed gas-phase abundances of acetone are surprisingly high compared to those of two-carbon O-COMs (especially aldehydes). Propanal is considered as tentatively detected due to lack of unblended lines covered in our data. The derived physical properties suggest that acetone, propanal, and ketene have the same origin from hot cores as other O-COMs, while propyne tends to trace the more extended outflows. The acetone-to-methanol ratios are higher in the solid phase than in the gas phase by one order of magnitude, which suggests gas-phase reprocessing after sublimation. There are several suggested formation pathways of acetone (in both ice and gas) from acetaldehyde, ketene, and propylene. The observed ratios between acetone and these three species are rather constant across the sample, and can be well reproduced by astrochemical simulations.

Conclusions. On the one hand, the observed high gas-phase abundances of acetone along with dimethyl ether (CH_3OCH_3) and methyl formate (CH_3OCHO) may hint at specific chemical mechanisms that favor the production of ethers, esters, and ketones over alcohols and aldehydes. On the other hand, the overall low gas-phase abundances of aldehydes may result from destruction pathways that are overlooked or underestimated in previous studies. The discussed formation pathways of acetone from acetaldehyde, ketene, and propylene seem plausible from observations and simulations, but more investigations are needed to draw more solid conclusions. We emphasize the importance of studying acetone, which is an abundant COM that deserves more attention in the future.

3.1 Introduction

The formation of complex organic molecules (COMs), usually considered as carbon-bearing molecules with at least six atoms (Herbst & van Dishoeck 2009), has been a hot topic in astrochemistry since their first detections in the gas phase toward hot molecular cores (i.e., the compact, chemically rich regions around protostars) in the last century (e.g., Cummins et al. 1986; Blake et al. 1987). In the past decade or so, powerful radio telescopes, including both single dishes and interferometer arrays, have detected more than 100 gas-phase COMs in space (Jørgensen et al. 2020; McGuire 2022), and both the number and complexity keep increasing[1]. In particular, the Atacama Large Millimeter/submillimeter Array (ALMA), known for its unprecedented sensitivity and resolution, has significantly increased the number of the identified COM-rich sources. Thanks to ALMA's high efficiency of detecting weak lines and resolving the compact hot core regions, observations on gas-phase COMs are moving from case studies (e.g., toward Sgr B2, IRAS 16293-2243, and G31.41+0.31; Belloche et al. 2016; Jørgensen et al. 2016; Mininni et al. 2023) to large-sample surveys (e.g., van Gelder et al. 2020; Yang et al. 2021; Hsu et al. 2022; Nazari et al. 2022a; Chapter 2), where the COM abundances can be accurately measured and systematically analyzed.

Although gas-phase COMs have been intensively studied, the exploration of their icy counterparts has just been enabled by the *James Webb* Space Telescope (JWST), or specifically, its Mid-Infrared Instrument (MIRI) and the Medium Resolution Spectroscopy (MRS) mode, which cover the fingerprint range of COM ices around 7–8 µm. Before the era of JWST, only methanol (CH_3OH), the simplest and most abundant COM, was firmly detected in interstellar ice (Boogert et al. 2015), with some tentative attribution of the absorption bands at 7.24 and 7.41 µm to ethanol (C_2H_5OH) and acetaldehyde (CH_3CHO) ices (Schutte et al. 1999; Öberg et al. 2011). Recently, more quantitative detections of COM ices, including ethanol, acetaldehyde, dimethyl ether (CH_3OCH_3), methyl formate (CH_3OCHO), and acetone (CH_3COCH_3), have been claimed toward three protostars using JWST/MIRI-MRS (Rocha et al. 2024; Chapter 4). Methyl cyanide (CH_3CN), the most abundant nitrogen-bearing COM, has also been searched in ices and considered as tentatively detected in several other protostars using the Near Infrared Spectrograph (NIRSpec; Nazari et al. 2024a). With the new detections of COM ices, it is now feasible to make gas-to-ice comparisons in their abundances, which can help us probe into the chemical evolution of COMs from ice (on dust grains) to gas (in hot cores). These comparisons have been conducted for four O-COMs (CH_3CHO, C_2H_5OH, CH_3OCH_3, and CH_3OCHO) toward two low-mass protostars B1-c and NGC 1333 IRAS 2A (Chapter 4), and two O-COMs (CH_3CHO and C_2H_5OH) show signs of undergoing reprocessing in the gas phase. However, the conclusions remain to be strengthened by investigating more species and sources.

Besides the aforementioned four O-COMs, acetone is also considered as detected in ice in B1-c (Chapter 4), but is not included in the gas-to-ice comparisons, largely because of lacking statistics on the gas-phase abundances of acetone. In previous observational studies on gas-phase COMs, acetone was either overlooked or served as a "byproduct" of those more abundant COMs with one or two carbon atoms. However, previous measurements show that the relative abundances of acetone with respect to

[1] `https://cdms.astro.uni-koeln.de/classic/molecules`

methanol are 10^{-3}–10^{-2} (Isokoski et al. 2013; Fuente et al. 2014; Lykke et al. 2017; van Gelder et al. 2020; Chahine et al. 2022; Baek et al. 2022; Mininni et al. 2023), which is well above the detection limit of ALMA when covering the strong transitions. In Sgr B2 and IRAS 16293-2243 (hereafter IRAS 16293), acetone was found to be much more abundant than its isomer, propanal (C_2H_5CHO), by around one order of magnitude (Belloche et al. 2013; Lykke et al. 2017). This raises an intriguing question why acetone is such abundant as a three-carbon and ten-atom molecule.

To answer this question, we conduct a follow-up study of Chapter 2, where the physical properties of six selected O-COMs (i.e., CH_3CHO, C_2H_5OH, CH_3OCH_3, CH_3OCHO, CH_2OHCHO, and $(CH_2OH)_2$) were measured toward 14 high-mass protostellar sources in the Complex Chemistry in hot Cores with ALMA (CoCCoA) program. In this work, we extend our attention from two-carbon O-COMs to acetone and propanal, which will increase the sample size by more than ten sources for the detection of three-carbon O-COMs. We also check ketene (CH_2CO) and propyne (CH_3CCH) which may be relevant to the formation of acetone (see discussion in Sect. 3.5 for details; Sung & Tidwell 1998; Hudson et al. 2017). By making comparisons in acetone abundances between gas and ice and between observations and simulations, we can shed light on the chemical evolution of acetone from ice to gas in protostellar systems. This work is also helpful for future studies on the formation mechanisms of O-bearing COMs in general (see Sect. 3.5.1).

3.2 Observations

The CoCCoA program (PI: B. A. McGuire) includes data from four ALMA projects: 2019.1.00246.S, 2019.2.00112.S, 2022.1.00499.S, and 2023.1.00466.S. This paper makes use of the original 2019.1.00246.S project, which observed 24 high-mass star-forming regions that are known to be chemically rich. The observations aimed to cover two spectral tunings per source in Band 6, which are 238.0–241.7 GHz in the lower tuning, and 258.0–261.7 GHz in the upper one. In the 2019.1.00246.S project, both tunings were observed for 16 out of 24 sources at a uniform angular resolution of $\sim$0.3″, and one source (Orion KL) was observed at a lower resolution of $\sim$0.7″. The remaining seven sources only have the lower tuning observed but at a higher resolution of $\sim$0.17″. In the follow-up project 2022.1.00499.S, the upper-tuning observations of these seven sources were executed, and all the targeted sources except for Orion KL were observed again in both tunings at a lower resolution of $\sim$1.0″ to cover the molecular emission on a larger scale.

So far, only the upper-tuning data (258.0–261.7 GHz) of the 16 sources observed in project 2019.1.00246.S have been manually reduced, which already yield fruitful results on six selected O-COMs in Chapter 2. In this work, we switch our focus to a larger O-COM that was not studied in Chapter 2, acetone, of which the emission lines are mostly blended in the upper tuning. Therefore, we want to make use of the lower tuning as well, in order to cover as wide a frequency range as possible when fitting the spectra.

Although the lower-tuning data have not been manually reduced, we found that the pipeline data provided in the ALMA Science Archive are already of very good quality. In Chapter 2, the upper-tuning spectra were extracted from the peak pixels of the integrated intensity (moment 0) maps of a representative CH_3OH line ($19_{3,17}$–

$19_{2,18}$ at 258.78 GHz). This line is the strongest unblended CH_3OH line in the upper tuning. Since CH_3OH is the simplest and most abundant COM, its emission is likely to share a similar morphology with other O-COMs. We found that the pipeline-reduced datacubes can produce similar spectra if extracted from the same positions (i.e., the CH_3OH emission peaks). The only difference is that the manually reduced data have a larger beam (0.33″) than the pipeline data (0.24″–0.27″), which introduces 10–20% difference in the line intensities. To unify the beam size, we implemented the `imsmooth` task in CASA and adjusted the beam size of the pipeline data to 0.33″. In this way, the difference becomes very small ($\lesssim 5\%$) between the manually reduced and the pipeline data, which is smaller than the calibration uncertainties in ALMA Band 6 ($\sim$10%) according to our quality assurance stage 2 (QA2) report.

By concatenating the pipeline spectra in the lower tuning and the manually reduced spectra in the upper tuning extracted from the same coordinates, we now obtain "complete" spectra covering both tunings (7.5 GHz in total) for 12 sources: G19.01-0.01, G19.88-0.53, G22.04+0.22, G23.21-0.37, G34.30+0.20, G34.41+0.24, G345.5+1.5, G35.03+0.35, G35.20-0.74N, IRAS 16547-4247 (hereafter IRAS 16547), IRAS 18151-1208 (hereafter IRAS 18151), and NGC 6334-38. The source luminosities are listed in Table 3.1, and we refer to Table 2.1 for other information on coordinates and distances. Reusing the manually reduced data in the upper tuning also means that the fitting results of the six selected O-COMs in Chapter 2 are still valid.

3.3 Methods

We performed local thermodynamic equilibrium (LTE) fitting to the ALMA spectra using the spectral analysis software CASSIS[2] (Vastel et al. 2015). We derived the column densities (N), excitation temperatures ($T_{\rm ex}$), line widths (FWHM), and velocities ($v_{\rm lsr}$) of gas-phase acetone in 12 CoCCoA sources. Here LTE means assuming the populations of all levels can be characterized by a single temperature, that is, assuming one temperature component in the fitting, whereas this temperature does not need to be the same as the kinetic temperature. This assumption has been shown to work well in Chapter 2 and other observational studies (e.g., Jørgensen et al. 2016).

In addition to acetone, we also checked ketene, propyne, and propanal, which may be chemically relevant to acetone according to astrochemical simulations and experiments. For this section, we only introduce how we handled the spectral fitting, and we reserve the discussion about their relations to acetone in Sect. 3.5.

3.3.1 Acetone (CH_3COCH_3)

The concatenated spectra cover tens of acetone lines above the detection limit, but only a small portion is unblended or marginally blended in the fitting (i.e., unblended within the FWHM) due to the large amount of emission lines emerging from the hot cores. The line blending issue is more severe for bright sources in which the lines are stronger and wider. Figure 3.1 shows the LTE-modeled spectra of acetone in comparison with the observed spectra, taking three sources with different line widths as examples. The key transitions are the triplets around 239.99 GHz and the one at

[2] `http://cassis.irap.omp.eu/`

Table 3.1: Luminosities of the 12 CoCCoA sources and the best-fit parameters of acetone.

Source	L (10^4 $L_\odot$)	$N_{CH_3COCH_3}$ (cm^{-2})	T_{ex} (K)		FWHM (km s^{-1})		v_{lsr} (km s^{-1})	
			CH_3COCH_3	O-COMs	CH_3COCH_3	O-COMs	CH_3COCH_3	O-COMs
G19.01-0.03	1	$2.6^{+2.0}_{-1.0} \times 10^{16}$	160^{+45}_{-30}	120∼180	3.5	3.0∼4.0	62.3	61.8∼63.0
		$4.4^{+1.3}_{-1.2} \times 10^{16}$	200±25		3.5	3.2∼3.5	57.5	57.8∼58.5
G19.88-0.53	0.47	$7.2^{+2.8}_{-1.9} \times 10^{15}$	110±15	100∼220	3.0	2.5∼4.0	46.7	46.0∼46.7
		$4.5^{+1.6}_{-1.3} \times 10^{15}$	95±15	–	3.0	–	43.7	–
G22.04+0.22	0.497	$(1.5\pm0.2) \times 10^{17}$	145±10	130∼170	6.5	4.5∼5.5	52.5	52.0∼53.0
G23.21-0.37	1.3	$4.2^{+0.8}_{-0.7} \times 10^{16}$	105^{+15}_{-10}	120∼200	3.5	3.0∼3.8	77.2	76.8∼77.2
G34.30+0.20	4.6	$2.8^{+0.4}_{-0.4} \times 10^{16}$	120±10	130∼180	3.0	3.2∼3.5	56.2	55.5∼56.5
G34.41+0.24	0.48	$(3.7\pm0.5) \times 10^{16}$	138±8	130∼180	3.3	2.5∼3.0	59.8	60.0
		$9.5^{+3.0}_{-2.5} \times 10^{15}$	93±15		3.5	3.0∼3.3	56.8	56.3∼57.0
G345.5+1.5	4.8	$7.9^{+3.8}_{-2.8} \times 10^{15}$	160±30	120∼140	3.5	2.0∼2.2	-14.5	-16.3∼-15.0
G35.03+0.35	0.63	$5.5^{+2.5}_{-1.8} \times 10^{15}$	110±20	110∼220	4.4	2.5∼4.0	44.5	43.5∼45.3
G35.20-0.74N	3	$3.8^{+0.5}_{-0.3} \times 10^{16}$	118^{+8}_{-5}	100∼180	4.2	3.0∼4.0	31.5	30.6∼31.2
IRAS 16547-4247	6.3	$(6.5\pm\textit{2.1}) \times 10^{16}$	180±20	140∼200	3.5	2.2∼3.3	-35.5	-35.7∼-35.2
		$(8.0\pm\textit{2.0}) \times 10^{16}$	110±20	–	3.0	–	-38.7	–
IRAS 18151-1208	2.2	$1.5^{+0.5}_{-0.6} \times 10^{15}$	110∼180	110∼180	2.0	1.7∼2.0	35.0	35.0∼35.3
NGC 6334-38	<20	$2.7^{+2.0}_{-0.9} \times 10^{16}$	170^{+50}_{-30}	100∼190	3.5	2.6∼3.2	-5.4	-5.2∼-5.0

261.36 GHz (i.e., the second and the second to last panels in Fig. 3.1). The multiple lines within 238.03–238.35 GHz are useful in constraining the $T_{\rm ex}$, since they have high upper energy levels ($E_{\rm up}$) of ∼300 K, whereas the $E_{\rm up}$ of other strong transitions fall in 77–171 K. These high-$E_{\rm up}$ lines are weak (optically thin) and unblended, thus are also helpful for monitoring overestimation in the fitting. More information on the transitions and the LTE-modeled spectra can be found in Appendices 3.C–3.D

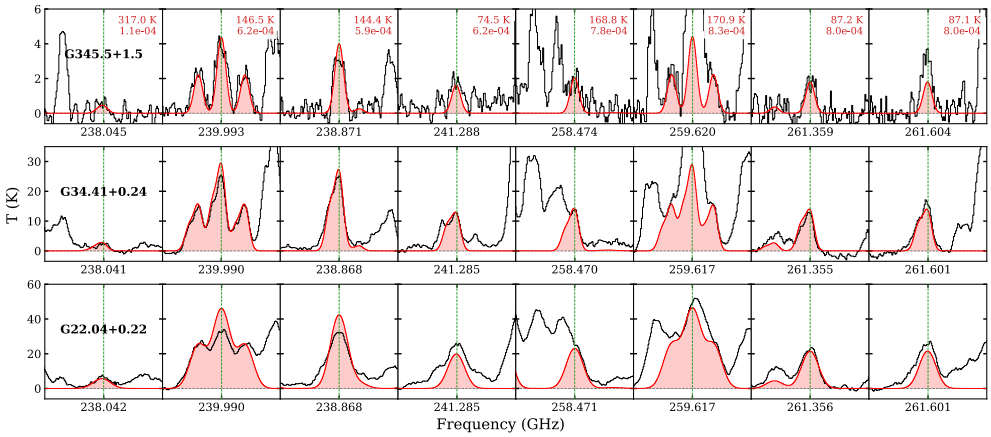


Figure 3.1: Best-fit LTE models of acetone spectra (red) at eight strong transitions overlaid on the ALMA spectrum of three example sources in CoCCoA (black). The upper energy levels and Einstein A coefficients are annotated in red text in the upper right. On the x axis, only the central frequency of each line is labeled, but the velocity span of each panel is fixed at [–15, +15] km s^{-1}.

Similar to Sect. 2.3.1, we tried both manual fitting (by visual inspection) and grid fitting (by running script of χ^2 minimization). The manual fitting was used for sources that have more than one velocity component (e.g., G19.01-0.03) or bright sources (e.g., IRAS 16547) that suffer too much from line blending, and the grid fitting was more suitable for weak and intermediate sources. In the latter case, a contour plot on the N–$T_{\rm ex}$ plane (as shown in Fig. 3.2) was plotted to determine the best-fit N and $T_{\rm ex}$, and in the meantime, estimate the uncertainties on a 2σ level.

For some sources such as IRAS 18151, there is a strong degeneracy between N and $T_{\rm ex}$ of acetone, that is, similar fitting results can be produced by increasing N and $T_{\rm ex}$ simultaneously (see the bottom panel of Fig. 3.2). As a consequence, the possible $T_{\rm ex}$ ranges from 100 K to 300 K, and the corresponding best-fit N varies across one order of magnitude, well above the normally estimated uncertainty of 30%. Interestingly, this problem was also encountered by Lykke et al. (2017), whose data are in a different band and cover a much wider frequency range (329.15–362.90 GHz, literally half of the ALMA Band 7). They chose to assign a fixed $T_{\rm ex}$ and determine the best-fit N. In our sample, most sources have a contour plot between the cases of G23.21-0.37 and IRAS 18151. For sources that suffer from degeneracy between N and $T_{\rm ex}$ like IRAS 18151, the range of $T_{\rm ex}$ set by all the other O-COMs derived in Chapter 2 were used as a constraint on the $T_{\rm ex}$ of acetone (see the fifth column in Table 3.1).

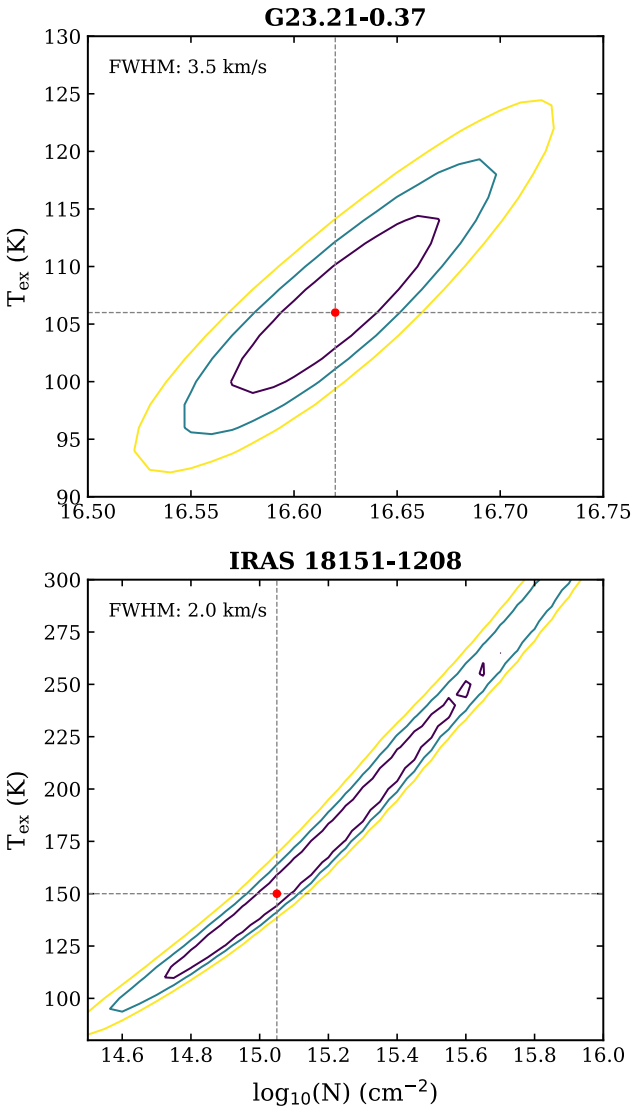


Figure 3.2: χ^2 contour plots on the $\log_{10}(N) - T_{\rm ex}$ plane for two example sources, G23.21-0.37 (top) and IRAS 18515-1208 (bottom). The contours in yellow, blue, and purple correspond to the 1, 2, and 3σ around the minimum of χ^2. The best-fit $\log_{10}(N)$ and $T_{\rm ex}$ are labeled with a red dot, and the best-fit FWHM is annotated in the top left corner.

3.3.2 Ketene (CH_2CO)

For CH_2CO, two strong lines, one in each tuning, are covered in our data. Unfortunately, these two lines are partially blended with strong lines of CH_2DOH and NH_2CHO, respectively. In our sample, the NH_2CHO line is optically thick and overfits the emission feature in most sources, and the CH_2DOH is optically thick in about half of the sources. In the cases that both CH_2DOH and NH_2CHO are optically thick, we could only estimate an upper limit of $N(CH_2CO)$ at a fixed $T_{\rm ex}$ of 100 K. If only NH_2CHO is optically thick, a column density was fit also at $T_{\rm ex}$=100 K, given that the $E_{\rm up}$ of the two CH_2CO lines are around 88 K and 100 K, respectively. In the cases that both CH_2DOH and NH_2CHO are optically thin, the $T_{\rm ex}$ was constrained to

be 50–120 K. Although $T_{\rm ex}$ could not be constrained very well by only two lines with similar $E_{\rm up}$, changing $T_{\rm ex}$ between 50 K and 150 K would only make a small difference (<20%) in the best-fit $N(\rm CH_2CO)$. A higher $T_{\rm ex}$ of 200–300 K would increase the best-fit $N(\rm CH_2CO)$ by 50–100%, but is also less plausible given that the $T_{\rm ex}$ of O-COMs are generally lower than 200 K (as given by the fifth column in Table 3.1).

The Einstein A coefficients ($A_{\rm ij}$) of the two $\rm CH_2CO$ lines are (1.5–2.0)$\times10^{-4}$ s^{-1}, which corresponds to a modeled optical depth of $\tau \sim 0.1$–1.1 under the best-fit conditions (specifically, six sources have $\tau \sim 0.1$–0.25, two $\tau \sim 0.5$, two $\tau \sim 0.7$, and two $\tau \sim 1.1$), which does not fully meet the criterion of optically thin lines ($\tau \ll 1$). To avoid underestimation of $N(\rm CH_2CO)$, we checked the minor isotopologs of $\rm CH_2CO$, and only $\rm CH_2^{13}CO$ has two lines covered in our data. Since the $\rm CH_2^{13}CO$ lines are next to the strong lines of the main isotopolog, only upper limits were estimated. We compared these upper limits, $N_{\rm uplim}(\rm CH_2^{13}CO)$, with the supposed column densities of $\rm CH_2^{13}CO$, which is the measured $N(\rm CH_2CO)$ divided by the ^{12}C/^{13}C ratio. The ^{12}C/^{13}C ratio empirically scales with the distance to the Galactic Center ($D_{\rm GC}$; Milam et al. 2005):

$$R(^{13}\rm C) \equiv {}^{12}\rm C/{}^{13}\rm C = (6.21 \pm 1.00)\,\mathit{D}_{GC} + (18.71 \pm 7.37). \tag{3.1}$$

According to the $D_{\rm GC}$ provided in Table 2.1, the ^{12}C/^{13}C ratios of the considered sources range from 45 to 60, which are smaller than the value of local ISM (∼70). If

$$N_{\rm uplim}(\rm CH_2^{13}CO) > \frac{\mathit{N}(\rm CH_2CO)}{\mathit{R}(^{13}\rm C)}, \tag{3.2}$$

then $\rm CH_2CO$ is likely to be optically thick and its column densities were underestimated; otherwise, $\rm CH_2CO$ is likely to be optically thin, or optically thick to a small degree (as $N(\rm CH_2^{13}CO)$ were not well constrained). For all of our sources, Eq. (3.2) is not satisfied, so we consider the measured $N(\rm CH_2CO)$ to be reliable.

3.3.3 Propyne ($\rm CH_3CCH$)

Propyne lines are only present in the lower tuning, and most of them are strong and unblended, enabling good constraints on the physical properties (N, $T_{\rm ex}$, $v_{\rm lsr}$, and FWHM). The line at 239.234 GHz is blended with a strong line of a-$(\rm CH_2OH)_2$; to crosscheck, we plugged in the fitting results of a-$(\rm CH_2OH)_2$ reported by Chapter 2, and the overall fitting looks decent when $(\rm CH_2OH)_2$ is not optically thick. We also checked the ^{13}C isotopolog of $\rm CH_3CCH$ as we did for $\rm CH_2CO$: only $\rm CH_3^{13}CCH$ is covered in our data, and its upper limits suggest that $\rm CH_3CCH$ is likely to be optically thin.

We noticed that some sources (e.g., G19.01-0.03, see Table 3.B.1 for details) clearly show more than one velocity components in the propyne lines. In some sources such as G345.5+1.5 (Fig. D.7) and G35.20-0.74N (Fig. D.9), the velocity separation between different components is larger than the FWHM of propyne lines, so that each velocity component can be fit individually by running the χ^2 minimization script. For G19.01-0.03, G35.03+0.35, and IRAS 16547, the two velocity components are not separated enough, and we fit them by visual inspection with the relative uncertainty on N estimated to be 25%.

3.3.4 Propanal (C_2H_5CHO) and other three-carbon O-COMs

Propanal drew attention as an isomer of acetone. The *syn* conformer (*s*-C_2H_5CHO) is the lowest energy conformer and has many transitions with E_{up} of 140–200 K and $A_{ij} > 10^{-4}$ s^{-1} covered in our data, while the *gauche* conformer only has weak transitions (i.e., $E_{up} > 600$ K and $A_{ij} < 10^{-5}$ s^{-1}). Although the *s*-propanal lines look "strong" in terms of E_{up} and A_{ij}, they are mostly clusters of hyperfine transitions and correspond to weak or (partially) blended emission features in the spectra. Therefore, we only estimated upper limits of N(*s*-C_2H_5CHO), mainly based on the strongest line at 261.29 GHz, along with other unblended weaker lines (see Fig. 3.3 for the fitting and Table C.1 for the information of transitions and references). The v_{lsr} of *s*-propanal was also determined by the strongest line. However, we noticed that the modeled spectra have slight offset from the observed spectra in the central frequencies of several lines, which was also spotted in the fitting of Lykke et al. (2017) (see Fig. A.2 therein). This may result from the inaccuracy of laboratory measurements, and may lead to some difference in the v_{lsr} between *s*-propanal and other O-COMs in our case, where too few strong and unblended lines are detected.

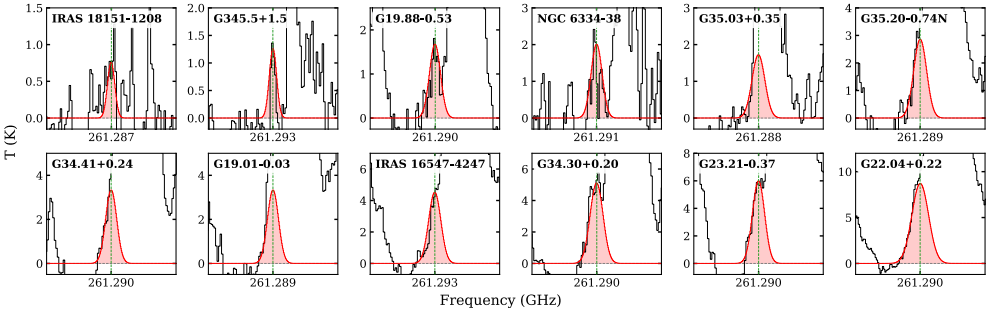


Figure 3.3: Best-fit LTE models of *s*-propanal spectra (red) overlaid on the ALMA spectrum of the 12 CoCCoA sources (black). All panels are centered on the strongest *s*-propanal line at 261.290 GHz. The velocity span of each panel is [–15, +15] km s^{-1}.

We used a fixed T_{ex} of 150 K for all sources, considering that this is a typical temperature of hot cores and also agrees well with the T_{ex} range set by other O-COMs (Table 3.1). Using a different T_{ex} between 80 and 220 K (i.e., 150±70 K) would only change the best-fit N by ∼20%, since the strong lines used in the fitting have similar E_{up} of 140–200 K (mostly around 180 K).

Other C_3H_6O isomers were also searched for, but no valid fitting could be performed. Propylene oxide (*c*-C_3H_6O) has no transition covered in our data. Propenol ($CH_3CHCHOH$) and isopropenol (CH_3COHCH_2) are not included in the databases. Oxetane ($CH_2CH_2CH_2O$) and methyl vinyl ether (CH_3OCHCH_2) are too rare in the astrochemical context and therefore not considered in this work.

We also tried fitting other three-carbon O-COMs such as propanol (C_3H_7OH) that do have transitions covered in our data, but their lines are either too weak or blended with others, making the constraints on column densities highly uncertain. We therefore did not report the results of these fitting attempts in this paper.

3.3.5 Methanol (CH_3OH)

In Chapter 2, the column densities of methanol were estimated from the minor isotopolog, $CH_3^{18}OH$, by multiplying its column densities with the $^{16}O/^{18}O$ ratios, which also scale with D_{GC} as follows (Wilson & Rood 1994):

$$R(^{18}O) \equiv {}^{16}O/{}^{18}O = (58.8 \pm 11.8)D_{GC} + (37.1 \pm 82.6). \quad (3.3)$$

This indirect determination of $N(CH_3OH)$ was applied because almost all (except one) detected methanol lines in the upper tuning are optically thick and cannot be used in the fitting. However, in some weak sources such as IRAS 18501-1208, the $CH_3^{18}OH$ lines are not clearly detected, which may lead to inaccurate estimation of $N(CH_3OH)$. Fortunately, there are many more methanol lines covered in the lower tuning, and several of them are likely to be optically thin given their high E_{up} ($\gtrsim$ 800–1000 K) and small A_{ij} ($\lesssim 3 \times 10^{-5}$ s^{-1}). We therefore revisited the fitting of methanol lines by only taking those optically thin lines into account. The newly derived $N(CH_3OH)$ do not differ too much from the values given in Chapter 2; the difference is within a factor of two for all sources but G34.41+0.24. The O-COM ratios with respect to methanol involved in this paper were recalculated using the new values of $N(CH_3OH)$.

3.4 Results

We claim firm detections of acetone, ketene, and propyne in all the 12 CoCCoA sources. The transitions with $E_{up} < 800$ K and $A_{ij} > 5 \times 10^{-5}$ s^{-1} covered in the dataset are listed in Table 3.C.1Figure 3.1 shows the best-fit spectra of acetone at key lines in three example sources; the fitting of more lines and all sources are provided by Figs. 3.D (Figs. 3.D.1–3.D.12). Overall, the LTE models fit the observations well, except that in some bright sources, some strong lines are optically thick and overestimated by the models (mostly methanol lines, and one or two acetone lines composed of hyperfine transitions with high A_{ij}).

3.4.1 Emission morphology

Figure 3.4 shows the integrated intensity maps of representative emission lines of methanol, acetone, propanal, ketene, and propyne. The selection criteria of these emission lines are: they should be unblended within their FWHM, and also bright enough to depict the emitting regions of the molecules. For methanol, we chose the same line as used in Chapter 2, which has $E_{up} = 490.59$ K and $A_{ij} = 9.01 \times 10^{-5}$ s^{-1}. Even though this line has an E_{up} that is considerably higher than the typical T_{ex} of hot cores (100–200 K), and many more methanol lines were included after adding the lower tuning, it is still the brightest among all the unblended methanol lines (largely because our data only cover methanol lines with $E_{up} < 90$ K and > 320 K).

Two sources are taken as examples: G19.88-0.53 with compact (i.e., not spatially resolved) methanol emission, and G34.30+0.20 with the most extended (i.e., spatially resolved) methanol emission in the CoCCoA sample. Most sources in the CoCCoA sample are more similar to G19.88-0.53 in that the methanol emission is not or only partially resolved by the beam of 0.33″.

In G19.88-0.53, the emission maps of acetone, propanal, and ketene show a similar compact morphology to methanol and other O-COMs around the continuum peak, indicating that they all come from the hot core region (see Fig. 2.A.1 for the integrated intensity maps of other O-COMs). As for propyne, the emission peak is slightly offset from the continuum peak, and the morphology shows an east-west elongation in G19.88-0.53. It seems that propyne is not tracing the hot core but some outflow-like structures. This is consistent with previous observational studies that this source has a bipolar outflow in the east–west direction revealed by H_2 and CO (isotopolog) emission (Varricatt et al. 2010; Issac et al. 2020). The emission morphology of propyne is also similar to that of C_2H shown in Saha et al. (2022), suggesting that they are probably tracing the outflow cavity walls where the UV intensity is higher (Tychoniec et al. 2021) and the abundances of hydrocarbons are enhanced. Propyne is also possibly present in the warm envelope, which is colder (<100 K) and more extended than the hot core. The supposed origins of propyne emission are also supported by its physical properties ($T_{\rm ex}$, FWHM, and $v_{\rm lsr}$; see discussion in Sect. 3.4.2).

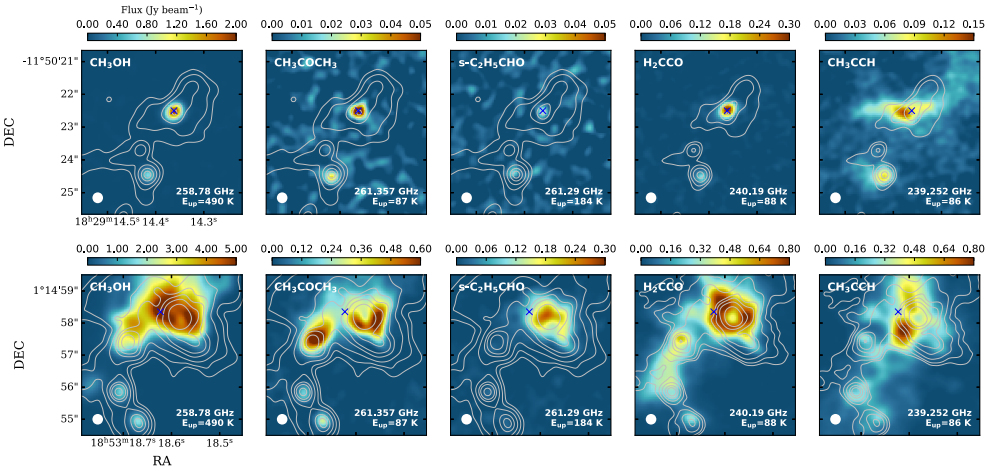


Figure 3.4: Moment 0 maps of the methanol, acetone, propanal, ketene, and propyne emission in G19.88-0.53 and G34.30+0.20. In each panel, the rest frequency and the upper energy level of the corresponding line is annotated in the bottom right corner, and the beam size is indicated in the bottom left corner. Contours in gray denote the 3, 5, 10, and 20σ levels of the continuum emission at 258.465 GHz (∼1.16 mm), where σ = 3 mJy beam^{-1}. The continuum maps are displayed in Fig. 2.1. The location where the ALMA spectrum was extracted is indicated by a blue cross in each panel.

In G34.30+0.20, there are four continuum peaks and the ALMA spectrum was extracted toward the brightest one on the upper right (the other three are much weaker on the left side). The hot core region is also around the continuum peak, but much more extended than that of G19.88-0.53. The molecular emission is attenuated by dust at the brightest continuum peak and shows an annular feature, which is particularly evident in the emission map of ketene. The emission of acetone, however, is asymmetric and brighter on the south side of the annulus. We noticed that the emission peak of methanol (also where the ALMA spectrum was extracted in Chapter 2) is actually different from that of acetone emission. This may cause underestimation of the column

density of acetone and its ratios with respect to other species, (e.g., methanol), which will be discussed in Sect. 3.5. However, this is not the case in other sources where the hot regions are not well spatially resolved and the emission of O-COMs all comes from a similar compact region. We think G34.30+0.20 is an interesting source for investigating the spatial distribution of COM emission, which we decided to leave for future studies.

Similar to the case of G19.88-0.53, the propyne emission in G34.30+0.20 again does not follow the morphology of other species and shows a bipolar feature instead of an annulus shape. According to the emission maps of these two sources, we consider ketene to be co-distributed with O-COMs in the hot core regions, while propyne tends to trace different physical components (e.g., outflow cavity walls) in the protostellar systems.

3.4.2 Physical properties

Table 3.1 lists the derived N, $T_{\rm ex}$, FWHM, and $v_{\rm lsr}$ of acetone along with the ranges of these parameters obtained from other O-COMs in Chapter 2. The same results for propanal, ketene, and propyne are provided in Table 3.B.1. In most sources, acetone has comparable $T_{\rm ex}$, FWHM, and $v_{\rm lsr}$ to other six O-COMs, suggesting similar origins from hot cores.

For ketene, as described in Sect. 3.3.2, the only two covered lines are blended with strong CH_2DOH and NH_2CHO lines. As a result, we could only fit N at a fixed $T_{\rm ex}$. The FWHM and $v_{\rm lsr}$ were estimated normally, but less accurately than for other species (the determination of FWHM also affects the best-fit N). The $v_{\rm lsr}$ of ketene agrees well with O-COMs, and the FWHM is comparable to or slightly larger, suggesting an origin from hot cores.

For *s*-propanal, we constrained the upper limits of its column densities at a fixed $T_{\rm ex}$ of 150 K by fitting the strongest line at 261.29 GHz, as shown in Fig. 3.3 ($T_{\rm ex}$ could not be constrained due to very limited detected lines). For some sources such as G345.5+1.5, G19.88-0.53, NGC 6334-38, G34.30+0.20, and G23.21-0.37, the profile of this line can still be recognized, but for others it cannot, and the constraints had to depend on other weaker lines (see figures in Appendix D). Although a full constraint of N and $T_{\rm ex}$ cannot be well achieved with the frequency coverage of our data, the upper limits are reliable at the order-of-magnitude level, or even better than a factor of two for those sources with a distinguishable line profile at 261.29 GHz.

For propyne, the physical properties were generally well constrained. The $T_{\rm ex}$ of propyne is overall consistent with those of O-COMs (∼100–200 K), except in some sources (e.g., IRAS 18151 and NGC 6334-38) lower than 100 K. The FWHM of propyne is generally comparable to that of O-COMs (except for propanal), with some exceptions where the propyne lines are wider or narrower, but this may due to the merge or split of different velocity components. For instance, in G345.5+1.5, G35.05+0.35, and G35.20-0.74N, only propyne shows more than one velocity component, and each of them has smaller FWHM. The match and mismatch in FWHM can correspond to origins from warm envelopes and outflow cavity walls, respectively. The $v_{\rm lsr}$ of propyne lines in the main velocity component tend to slightly deviate from the O-COM range; in most cases the difference is small (∼0.5 km s^{-1}), but sometimes can be as large as 1.5–2.0 km s^{-1}, and even larger in the secondary or third velocity components if

there is one. The offset in $v_{\rm lsr}$ may suggest a different emitting region of propyne from O-COMs, which is displayed more straightforwardly in the moment 0 maps.

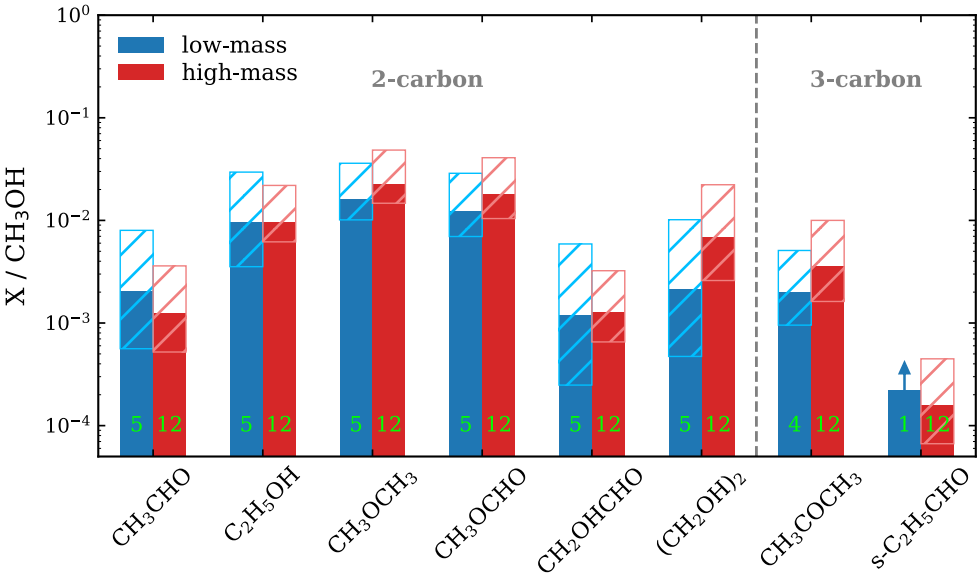


Figure 3.5: Column density ratios of six two-carbon and two three-carbon O-COMs with respect to methanol in low-mass (blue) and high-mass (red) protostars. The hatched regions indicate the standard deviation weighted by the uncertainties on log scales. The sample size (i.e., the number of sources) is labeled at the bottom of each bar in green. The column density of *s*-propanal was only measured in one low-mass source (IRAS 16293-2243 B), and its methanol column density was not measured at the same location as *s*-propanal but closer to the continuum peak, which generally has higher column densities; therefore the calculated ratio between propanal and methanol of this source is considered as lower limit.

3.5 Discussion

3.5.1 Acetone and other O-COMs

With the column densities of two three-carbon O-COMs (i.e., acetone and propanal) measured for the CoCCoA sample, we can now compare their relative abundances with other two-carbon O-COMs that have been studied in Chapter 2. Figure 3.5 summarizes the column density ratios of these eight O-COMs with respect to methanol. The column densities of the six two-carbon O-COMs in the CoCCoA sample were originally taken from Chapter 2, but their ratios with respect to methanol were recalculated using the updated methanol column densities (see Sect. 3.3.5). Here the high-mass sample size is reduced from 14 in Chapter 2 to 12, since the new measurements of methanol were only applied to the 12 sources with concatenated spectra (Sect. 3.2). We also do not add other literature values in Fig. 3.5 to ensure that the column densities are uniformly measured from the same high-quality dataset. The low-mass sample

remains the same as in Chapter 2, including B1-c, B1-bS, S68N, IRAS 16293 A and B (van Gelder et al. 2020; Manigand et al. 2020; Jørgensen et al. 2018). For acetone, there is no measurement toward IRAS 16293 A, and propanal was only quantified in IRAS 16293 B, so the low-mass sample size is smaller for these two species.

For the first time, we show the column density ratios of eight O-COMs (or seven if excluding the less constrained propanal) with respect to methanol in a uniform sample of a dozen sources. On the one hand, as a three-carbon O-COM, acetone is not much less abundant than the two-carbon ones (within one order of magnitude), and even tends to be more abundant than CH_3CHO and CH_2OHCHO. On the other hand, propanal, as an isomer of acetone, is significantly less abundant than other O-COMs. Actually, not only propanal, all the three species with aldehyde groups (i.e., CH_3CHO, CH_2OHCHO, and C_2H_5CHO) are the least abundant among COMs with the same amount of carbon atoms. There are two possible explanations: aldehydes are easily destroyed or converted into other species, or aldehydes are less produced than O-COMs of other types. However, it is not practical to prove these two cases only by observations; insights from theoretical and experimental studies are needed.

Besides the low abundances of aldehydes, it is also noticeable that alcohols are not the most abundant species among two-carbon O-COMs. Instead, those with CH_3O radicals (CH_3OCH_3 and CH_3OCHO) are even more abundant than alcohols (C_2H_5OH and $(CH_2OH)_2$). A similar trend may also hold for three-carbon O-COMs, where acetone can be more abundant than propanol (C_3H_7OH), or at least comparably abundant. Propanol has only been detected toward two high-mass star-forming regions G+0.693-0.027 and Sgr B2(N) (Jiménez-Serra et al. 2022; Belloche et al. 2022), and its column densities cannot be reliably constrained using our data. In Sgr B2(N), the column density ratio of propanol with respect to methanol is 10^{-3}–10^{-2}, which is comparable to the acetone ratios in our sample. While acetone (CH_3COCH_3) is not a CH_3O-bearing species, it has a similar structure to CH_3OCH_3 in that there is a CH_3 group on each side and an oxygen atom in between. According to the available measurements, it seems that O-COMs with CH_3O or similar components are systematically more abundant in the gas phase.

However, the high abundances of CH_3OCH_3 and CH_3OCHO are difficult to reproduce in simulations and experiments. Instead, they were often underpredicted in simulations before the introduction of the barrierless reaction

$$C + H_2 \rightarrow CH_2 \tag{3.4}$$

and the nondiffusive grain-surface and ice-mantle chemistry in Garrod et al. (2022), although there are different opinions on whether reaction 3.4 is barrierless or not (Krasnokutski et al. 2016; Henning & Krasnokutski 2019; Lamberts et al. 2022; Potapov & Garrod 2024). In many experimental studies on solid-phase formation of O-COMs under non-energetic (e.g., Fedoseev et al. 2015, 2022; Chuang et al. 2020) or energetic (Chuang et al. 2017, 2021) conditions, CH_3OCH_3 is the only two-carbon O-COM that was not reported in the products, and CH_3OCHO is often less produced than its isomer CH_2OHCHO. Although CH_3OCH_3 has a viable gas-phase production route through the reaction of methanol with protonated methanol, its constant ratio of $\sim$1 with respect to CH_3OCHO across a large number of star-forming regions suggest non-negligible formation in ices (Coletta et al. 2020). The production of CH_3OCH_3 was only seen in the experiment of UV-irradiated CH_3OH:CH_4/CO ice mixtures by Öberg

et al. (2009b). In comparison, all the experiments with non-detection of CH_3OCH_3 did not include CH_4 in their initial ingredients; instead, they used C atoms, CO, C_2H_2, and CH_3OH. It seems that irradiating CH_4 is necessary to produce CH_3 radicals and boost the formation of CH_3OCH_3 in experiments. A question then comes up why CH_3 radicals are hard to generate experimentally, but the relevant molecules, represented by CH_3OCH_3, are observed to be considerably abundant in space. This question is well-known among experimentalists, and becomes more evident after the systematic high-abundance of acetone being revealed in this work.

3.5.2 Acetone: Ice versus gas

The recent detection of COM ices with JWST observations in two low-mass protostars (IRAS 2A and B1-c) enables the gas-to-ice comparisons for O-COMs including CH_3CHO, C_2H_5OH, CH_3OCH_3, and CH_3OCHO (Rocha et al. 2024; Chapter 4). By conducting gas-to-ice comparisons in their abundances, we can probe into their chemical evolution during the phase transition from ice in cold envelopes to gas in hot cores.

In addition to these two-carbon O-COMs, acetone ice is also considered as detected in B1-c and tentatively detected in IRAS 2A as the only known candidate for the observed absorption feature at 7.3 µm (detailed discussion on this point can be found in Sect. 4.4.2.3). This conclusion may be modified in the future if new candidates are found to fit the observations better, but the currently measured ice abundances of acetone can at least serve as upper limits.

Since the quantification of acetone ice has only been available in two sources, direct gas-to-ice comparisons in the same sources are still limited. However, with gas-phase acetone quantified in a larger sample of hot cores, we now have better constraints on its gas-phase abundances, making the gas-to-ice comparisons more reliable. The observed column density ratios between acetone and methanol in gas and ice are summarized in Fig. 3.6. For the gas-phase ratios, in addition to the 12 high-mass CoCCoA sources, we added literature data of four low-mass (B1-c, B1-bS, S68N, and IRAS 16293 B; van Gelder et al. 2020; Jørgensen et al. 2018; Nazari et al. 2024b), one intermediate (HOPS-108; Chahine et al. 2022), and two additional high-mass sources (G31.41+0.31 and Sgr B2(N); Mininni et al. 2023; Belloche et al. 2013). We also considered the work of Baek et al. (2022) (hereafter B22) who report acetone column densities in 13 cores from eight high-mass protostellar sources. However, their methanol column densities $N(CH_3OH)$ were likely measured from optically thick lines, and the reported $N(CH_3OH)$ values are obviously underestimated, given that the ratios between the main and the minor isotopologs are significantly off from the isotope ratios. Since they also provide $N(CH_3^{18}OH)$ and the source distance, more reliable values of $N(CH_3OH)$ can be calculated by multiplying $N(CH_3^{18}OH)$ and $R(^{18}O)$ given in Eq. (3.3). Nevertheless, the reported $N(CH_3^{18}OH)$ in B22 are higher than $N(^{13}CH_3OH)$, which is not normal since $R(^{13}C)$ is higher than $R(^{18}O)$, and $N(^{13}CH_3OH)$ should be higher than $N(CH_3^{18}OH)$. A possible explanation is that their $N(CH_3^{18}OH)$ were overestimated for some reasons, for instance, the line blending issue was not correctly handled. As a result, our re-calculated $N(CH_3OH)$ from $N(CH_3^{18}OH)$ may be overestimated, and the acetone-to-methanol ratios plotted in Fig. 3.6 may be underestimated for the B22 sample.

Figure 3.6 shows that the gas-phase abundances of acetone with respect to methanol

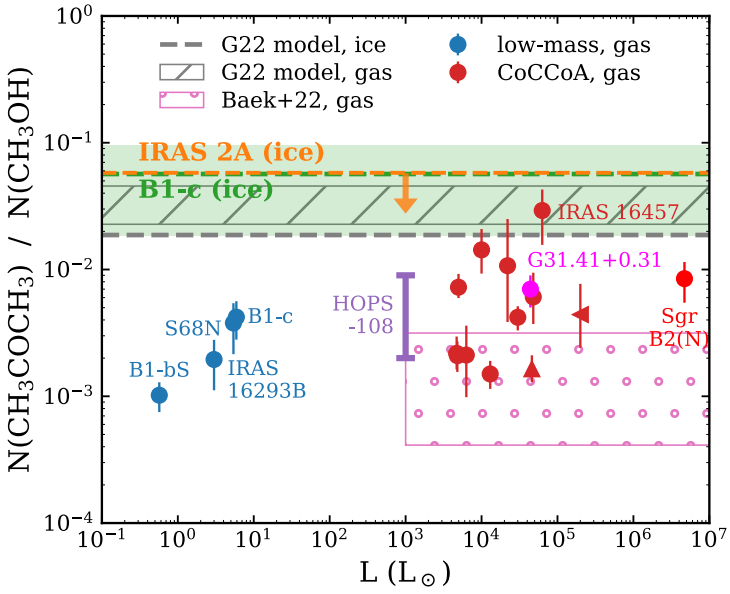


Figure 3.6: Column density ratios between acetone and methanol. The gas-phase ratios of low-, intermediate-, and high-mass sources are shown in blue, purple, and red points, respectively. The upward- and leftward-pointing triangles correspond to G34.30+0.20 (lower limit of the acetone-to-methanol ratio, see the last paragraph of Sect. 3.5.2) and NGC 6334-38 (upper limit of luminosity, see Table 3.1), respectively. The ice ratios two low-mass sources, B1-c and IRAS 2A, are indicated by dashed lines in green and orange, respectively. Since acetone ice is only considered as tentatively detected in IRAS 2A, the uncertainty is indicated by a green-shaded region only for B1-c. For the high-mass sample of Baek et al. (2022), only a range is displayed since no luminosity information is provided. The dashed line and hatched region in gray indicate the modeling results of the `MAGICKAL` astrochemical simulations in Belloche et al. (2022); the model is originally introduced in Garrod et al. (2022), hence abbreviated as G22 in the legend.

are scattered between 10^{-3} and 10^{-2}. In the high-mass sample, the CoCCoA sources are consistent with the three additional sources HOPS-108, G31.41+0.31, and Sgr B2(N). The one lower limit in the CoCCoA sample is G34.30+0.20, where the spectrum was not extracted from an acetone-rich position, so that the acetone-to-methanol ratio could be underestimated. The B22 sources show generally lower ratios, but is likely because their methanol column densities were overestimated. The low-mass samples show slightly lower acetone abundance than the high-mass one, which is different from other O-COMs that show no trend with luminosity (see Fig. 2.4). The ice abundance is about one order of magnitude higher than the mean gas abundance, following the same trend as CH_3CHO and C_2H_5OH (see Fig. 4.9). However, it should be taken into account that the sample size of ice ratios is very small (only two), and the quantification of ice abundances is much more uncertain than that of gaseous ones, as the current fitting results of COM ices may be changed in the future by possibly new identifications in the fingerprint range of 6.8–8.8 µm. Assuming that the ice abundances of acetone were reliably constrained, the higher ice ratios hint at gas-phase reprocessing

after sublimation, which reduces the amount of acetone relative to methanol, but the detailed mechanisms still need to be investigated.

3.5.3 Chemical evolution of acetone

To gain more insights into the formation and/or destruction mechanisms of acetone, we compared our observational results with the simulation results of the state-of-the-art astrochemical model `MAGICKAL` (Garrod 2013; Garrod et al. 2022). We also searched for other theoretical and experimental studies that are relevant to acetone, and a summary chemical network is shown in Fig. 3.7.

3.5.3.1 Insights from astrochemical simulations

The `MAGICKAL` model simulates the chemistry of interstellar molecules in three phases (gas phase, ice surface, and bulk ice mantle) during two stages (cold collapse stage and warm-up stage) of the formation of protostellar systems. Three timescales are considered for the warm-up stage: fast (5×10^4 yr), medium (2×10^5 yr), and slow (1×10^6 yr).

Here we used an updated version of `MAGICKAL` that was used in Belloche et al. (2022) to compare with the observational studies of propanol in Sgr B2(N). The chemical network of this model involves new gas-phase and grain-surface chemistry for propanol and related species. Although Belloche et al. (2022) only detail the changes related specifically to propanol, additional reactions are also included for acetone and propanal. As illustrated by the blue routes in Fig. 3.7, acetone is formed via two solid-phase and one gas-phase pathways. One solid-phase pathway is bottom-up from simpler species:

$$\mathrm{C} \xrightarrow{\mathrm{H,\,H_2}} \mathrm{CH_2} \xrightarrow{\mathrm{CO}} \mathrm{CH_2CO} \xrightarrow{\mathrm{2H}} \mathrm{CH_3CHO} \xrightarrow{\mathrm{CH_2}} \mathrm{CH_3COCH_3}, \tag{3.5}$$

and the other is via H abstraction of two iso-C_3H_7O radicals, $CH_3CH(\dot{O})CH_3$ and $CH_3\dot{C}(OH)CH_3$:

$$\mathrm{H + CH_3CH(\dot{O})CH_3 \rightarrow CH_3COCH_3 + H_2} \quad \text{and} \tag{3.6}$$

$$\mathrm{H + CH_3\dot{C}(OH)CH_3 \rightarrow CH_3COCH_3 + H_2}, \tag{3.7}$$

where the iso-C_3H_7O radicals are produced from iso-propanol (i-C_3H_7OH) and 2-C_3H_7 radicals:

$$\mathrm{2\text{-}C_3H_7 + O \rightarrow CH_3CH(\dot{O})CH_3} \quad \text{or} \tag{3.8}$$

$$\mathrm{i\text{-}C_3H_7OH} + h\nu \rightarrow \mathrm{CH_3CH(\dot{O})CH_3 + H} \quad \text{and} \tag{3.9}$$

$$\mathrm{i\text{-}C_3H_7OH + H \rightarrow CH_3\dot{C}(OH)CH_3 + H_2}. \tag{3.10}$$

The gas-phase pathway is:

$$\mathrm{2\text{-}C_3H_7 + O \rightarrow CH_3COCH_3 + H}, \tag{3.11}$$

with rates based on Tsang (1988). The same mechanism is available to form propanal via the 1-C_3H_7 radical. Similar grain-surface reaction to reaction (3.8) is also involved

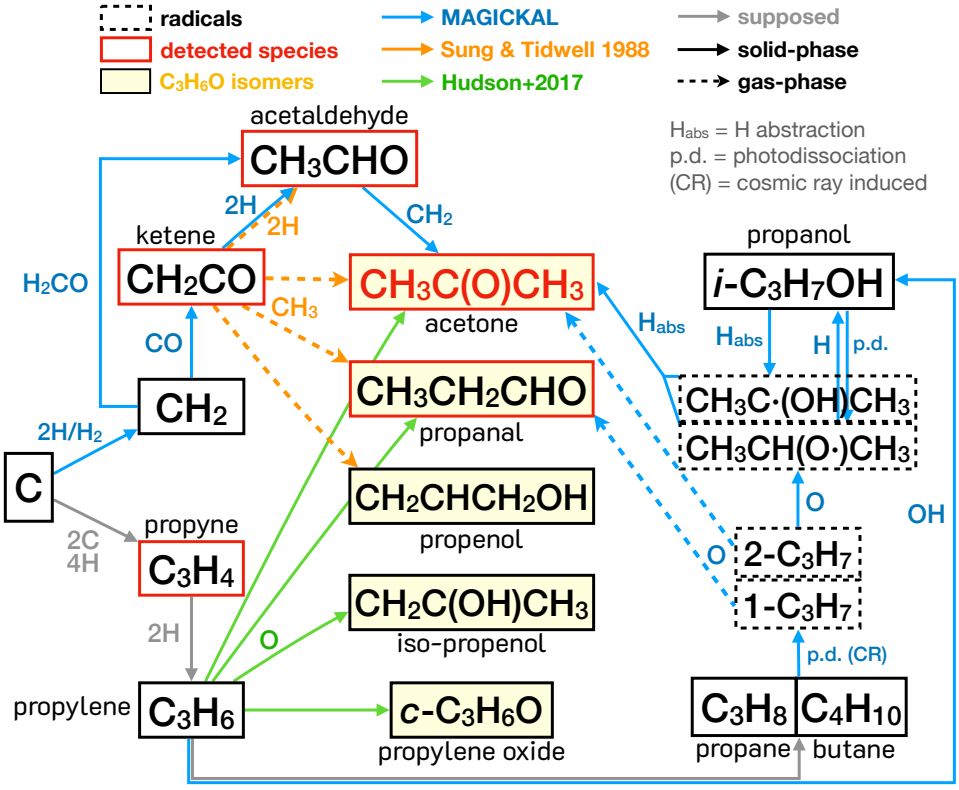


Figure 3.7: Schematics of the formation pathways of acetone, proposed by previous theoretical calculations (Sung & Tidwell 1998), astrochemical simulations (`MAGICKAL`; Garrod et al. 2022), and experiments (Hudson et al. 2017). Except the first study that assumes normal temperature and pressure, the remaining two are intended for astrochemical conditions (i.e., low temperature and pressure). Legends are provided in the figure.

in the production of propanal in the solid phase, which has origins in break-down of normal-propanol (*n*-propanol).

Figure 3.A.1 shows the rate of change of acetone abundances (summed over all phases) in the `MAGICKAL` simulations. Both the cold collapse and the warm-up stages have considerable amount of acetone formed, and the contribution from the gas-phase production via reaction (3.11) increases with the warm-up timescale (i.e., the longer the warm-up stage, the more acetone is formed in the gas phase). In the version of `MAGICKAL` models introduced in Garrod et al. (2022), only one solid-phase pathway (reaction 3.5) is included, and more than 90% of acetone is formed in the solid phase during the cold collapse stage. Without the other pathways (reactions 3.6–3.11), the peak acetone abundance would decrease by more than one order of magnitude. This difference results from the mediated energy barriers of H abstraction reactions (e.g., 3.6, 3.7, and 3.10), of which the energy barriers are poorly constrained by theoretical calculations and experiments, and many of them are set as zero.

The acetone-to-methanol ratios predicted by `MAGICKAL` simulations are plotted

along with the observational results in Fig. 3.6. The gas ratios are higher than the ice ones due to the gas-phase production in the warm-up stage (reaction 3.11). The model-predicted ratios are comparable to the observed ice ratios, but are higher than the observed gas ratios by about half order of magnitude. The overestimation by simulations hint at gas-phase destruction of acetone that is not fully considered in the current chemical models, which is also suggested by the gas-to-ice comparisons discussed in Sect. 3.5.2.

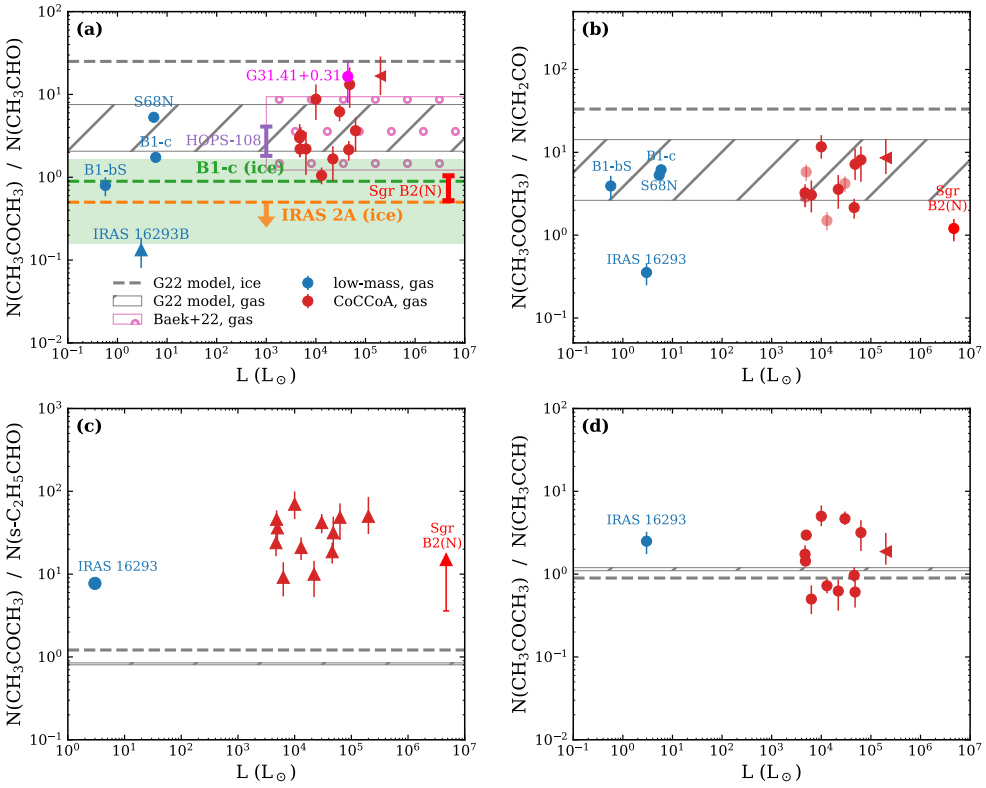


Figure 3.8: Column density ratios of acetone with respect to acetaldehyde (a), ketene (b), propanal (c), and propyne (d). In panel (b), the data points in lighter red refer to sources with less constrained column densities of ketene, of which the lines are blended with optically thick lines of CH_2DOH and NH_2CHO (see Sect. 3.3.2). Other denotations are the same as in Fig. 3.6.

In the bottom-up formation pathway (3.5) of acetone, acetaldehyde is the direct precursor of acetone. We therefore checked the column density ratios between acetone and acetaldehyde, as shown in Fig. 3.8a. The observed gas-phase ratios, especially in high-mass sources, are in general higher than the ice ratios, which means that acetaldehyde is less abundant than acetone in the gas phase. This echos the possibility of aldehydes being destroyed or converted into other species in the gas phase (Sect. 3.5.1). Interestingly, the B22 sample shows high consistency with our CoCCoA sample in the acetone-to-acetaldehyde ratios, implying that the lower acetone-to-methanol ratios are

likely due to the overestimated methanol column densities (as proposed in Sect. 3.5.2).

The MAGICKAL simulations predict the gas-phase acetone-to-acetaldehyde ratios very well, while the ice ratios are overestimated significantly. This may due to the conflicting evolution histories of acetaldehyde suggested by simulations and observations. In the MAGICKAL simulations, acetaldehyde is abundantly formed in the gas phase, which can be seen in Fig. 3.A.2 that its abundance keeps increasing after sublimation. The main formation mechanism for gas-phase acetaldehyde is given by

$$\mathrm{O} + \mathrm{C_2H_5} \rightarrow \mathrm{CH_3CHO} + \mathrm{H} \tag{3.12}$$

(Tsang & Hampson 1986). The production of C_2H_5, and of the C_3H_7 radicals that form acetone in the gas phase, is dependent on the destruction of stable hydrocarbons such as propane (C_3H_8), via cosmic-ray induced photodissociation. The rates of this process are poorly defined so far. The predicted abundances of gas-phase acetaldehyde in the MAGICKAL simulations are generally higher than in observations, in which the gas-phase ratios of acetaldehyde (with respect to methanol) are lower than those of other O-COMs (Fig. 3.5) and also the ice ratios of itself (Fig. 4.9); that is, our observations suggest the opposite case to the simulations that acetaldehyde is mainly formed in ices and then depleted in the gas phase. Taking the two discrepant evolution histories into account, it is possible that the overestimated ice ratios between acetone and acetaldehyde are due to the underproduction of acetaldehyde ice in the simulations, and the good match in the gas ratios could be just a coincidence. Although there is direct chemical link between acetone and acetaldehyde (i.e., the last step of pathway 3.5), this link is only considered in the solid phase, and both species are likely to undergo gas-phase reprocessing that seems not well simulated so far. However, we note that in high-mass protostellar systems, there might not be enough time for the gas-phase chemistry to significantly alter the abundances of these COMs if the reactions are not very efficient.

Other formation pathways (i.e., reactions 3.6–3.11) are difficult to examine from observations, since the relevant species are either radicals or large molecules that can hardly be quantified. We mentioned in Sect. 3.3.4 that the column densities of iso-propanol, even in the form of upper limits, cannot be well constrained using the CoCCoA dataset. The efficiency of reactions (3.6)–(3.11), especially the energy barriers, are also lack of constraints. However, the comparison between observations and simulations conveys an important message: the acetone abundances, along with its ratios with respect to other species, will be substantially underestimated if these reactions (3.6)–(3.11) are not included. This suggests that there must be more routes to form acetone than pathway (3.5), and we have not fully understood them yet.

3.5.3.2 Insights from computational chemistry

Besides the aforementioned pathways considered in the MAGICKAL simulations, an early theoretical study on ketene (Sung & Tidwell 1998) suggested that acetone may form directly via ketene adding CH_3 radicals (the orange routes in Fig. 3.7). Since their calculations were performed assuming normal temperature and pressure (i.e., in the gas phase), their conclusions may not fully applicable to astrochemical situations. However, some parameters in the results are irrelevant to physical conditions, and the suggested pathways can still be inspiring to future theoretical and experimental studies.

According to their calculations, reactions between ketene and CH_3 radicals can produce acetone and propanal with similar branching ratios; propenol is also one of the possible products but is less favored:

$$CH_2CO + CH_3 \xrightarrow{+H} CH_3COCH_3/C_2H_5CHO/CH_2CHCH_2OH. \quad (3.13)$$

We therefore checked the column density ratios between acetone and ketene in the gas phase, which are shown in Fig. 3.8b. The observed ratios are moderately scattered within 1–10, and are comparable between the low-mass and the high-mass sources, except the outlier IRAS 16293 where acetone is less abundant than ketene. One possibility would be that the emission of acetone and ketene is inhomogenous in IRAS 16293 as what we see in G34.30+0.20, and the ALMA spectrum of IRAS 16293 was extracted from a biased location where ketene emission is very bright. Otherwise, there should be something interesting happening to the chemistry in this source.

The observed gas-phase ratios between acetone and ketene are well reproduced by `MAGICKAL`, even though the only link between acetone and ketene in the model (i.e., pathway 3.5) is in the solid phase. According to the calculation results of Sung & Tidwell (1998), the energy barrier of reaction (3.13) is not low (∼4000 K). If this is the case, the contribution of this formation pathway would be very limited, even if it is included in the chemical network of `MAGICKAL`. Nevertheless, more reliable constraints on the efficiency as well as the branching ratios of reaction (3.13) should be set by experiments.

The simulated ice ratios between acetone and ketene are higher than the observed gas ratios by a few factors. Unfortunately, we are unable to measure the ice abundances of ketene by observations since no IR spectrum of ketene ice is available. Ketene is a very active species under normal conditions and therefore not commercially available for experiments. This also explains why ketene has not been used as an initial ingredient to form O-COMs in the laboratories.

Since it has been suggested that propanal has a similar branching ratio to acetone in reaction (3.13), we also checked the column density ratios between acetone and propanal, which are shown in Fig. 3.8c. These ratios are considered as lower limits, as only upper limits of $N(s\text{-}C_2H_5CHO)$ were constrained. Propanal was only quantified in one low-mass source IRAS 16293 B. The observations show that acetone is more abundant than propanal in the gas phase by at least one order of magnitude. If acetone is mainly formed via reaction (3.13), the ratio should be around one. The observed high ratios between acetone and propanal are also significantly underestimated by `MAGICKAL`. The `MAGICKAL` simulations tend to predict comparable or slightly higher amount of propanal than acetone, since many aforementioned formation pathways of acetone also work for propanal (e.g., if replacing 2-C_3H_7 with 1-C_3H_7 in reactions 3.8 and 3.11). The 1- and 2-C_3H_7 radicals are mainly produced from the break-down of normal- and iso-propanol, and the efficiency of relevant reactions is likely to affect the relative abundances between propanal and acetone.

So far, all versions of `MAGICKAL` simulations (Garrod 2013; Garrod et al. 2022; Belloche et al. 2022) have not been able to reproduce the high acetone-to-propanal ratios in observations. This discrepancy implies that there may be some gas-phase destruction mechanisms for propanal (or more broadly speaking, aldehydes) that we are either unaware of or not paying adequate attention to. The destruction of aldehydes should be much more efficient than that of other O-COMs. However, it is hard to

pinpoint the exact mechanisms only from observations; more investigation are needed by theorists and experimentalists.

3.5.3.3 Insights from experiments

Similar to astrochemical simulations, there is a lack of experimental studies that directly focus on acetone formation, though there are ones on large COMs such as propanol and glycine (Qasim et al. 2019; Ioppolo et al. 2021). In this case, we resort to two experimental studies on propylene oxide, an isomer of acetone, in which the discussion involves possible formation pathways of acetone (Hudson et al. 2017; Bergantini et al. 2018). They propose that propylene oxide ice can be formed through epoxidation of propylene (C_3H_6) by oxygen atoms yielded from photodissociation of CO_2:

$$C_3H_6 + O \rightarrow C_3H_6O. \tag{3.14}$$

When the oxygen atom breaks the double bond in propylene molecules, it is possible to form four C_3H_6O isomers: propylene oxide, acetone, propanal, and iso-propenol (green routes in Fig. 3.7). The branching ratios of this reaction were not measured in Hudson et al. (2017), but were provided by Bergantini et al. (2018), which conducted similar experiments (i.e., irradiating ice mixtures of C_3H_6 and CO_2 under low temperature, but with different irradiating conditions), and also performed quantum chemical calculations. In Bergantini et al. (2018), iso-propenol was not observed in the products, and the branching ratios between propylene oxide, acetone, and propanal derived from experiments are (92±4):(23±7):1.

The relation between acetone and propanal is shown in Fig. 3.8c and has been discussed in Sect. 3.5.3.1. It is worth noting that the branching ratio between acetone and propanal derived in Bergantini et al. (2018) is in good agreement with the observed ratios in the high-mass CoCCoA sample. Unfortunately, propylene oxide is not covered in our data, preventing further validation of reaction (3.14) by observations. Also, our data do not cover propylene lines, but do include propyne lines. We therefore checked the relation between acetone and propyne, a hydrogenation precursor of propylene. Figure 3.8d shows the column density ratios between acetone and propyne. There is only one low-mass source (IRAS 16293 B) that has column density measurements of both acetone and propyne (Calcutt et al. 2019), and no ice detection has been reported. The observed gas-phase ratios are moderately scattered by about one order of magnitude, and seem to have no trend with luminosity. The acetone-to-propyne ratios predicted by `MAGICKAL` simulations in both gas and ice match the observations very well.

Although we cannot rule out the possible chemical links between acetone and propyne, the evidence is very limited. In the chemical network of `MAGICKAL`, acetone and propyne are distantly related, and the possible closer relation, reaction (3.14), is not included. We still need more observations on the key species, propylene and propylene oxide, as well as modeling results from an updated chemical network.

3.5.4 Correlations among O-COMs and acetone relevant species

To summarize the discussion on the abundance statistics of O-COMs and the possible formation pathways of acetone in Sects. 3.5.1–3.5.3, we calculated the Pearson correlation coefficients (r) of the column densities of two species groups: (1) all the nine

O-COMs that have been quantified in the 12 CoCCoA sources in Chapter 2 and this work, including CH_3OH, CH_3CHO, C_2H_5OH, CH_3OCH_3, CH_3OCHO, CH_2OHCHO, $(CH_2OH)_2$, CH_3COCH_3, and C_2H_5CHO; (2) the five acetone-related species discussed in Sect. 3.5.3, which are CH_3COCH_3, C_2H_5CHO, CH_3CHO, CH_2CO, and CH_3CCH. The correlation matrices for the two groups are displayed in Fig. 3.9.

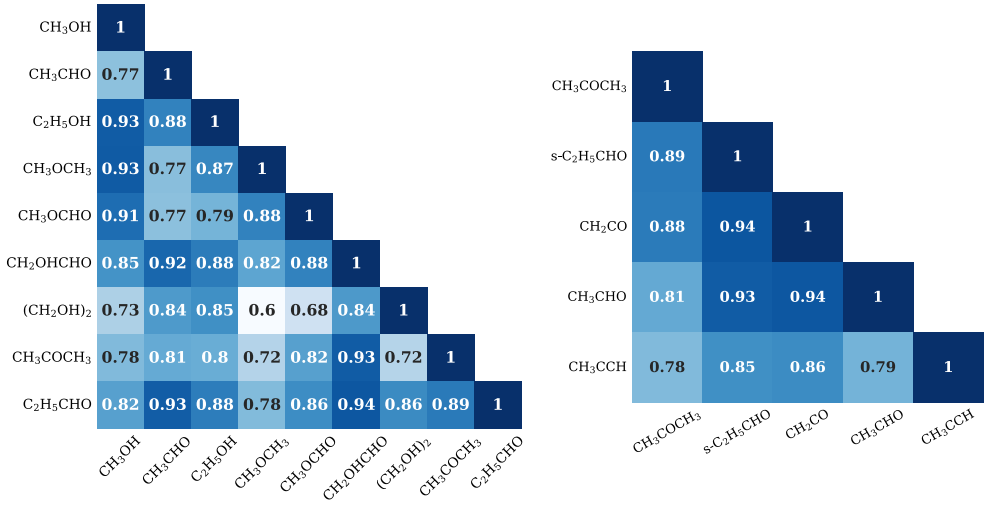


Figure 3.9: Correlation matrices of nine O-COMs (left) and five acetone-related species (right) in the 12 CoCCoA sources.

Among the nine O-COMs, the highest correlations belong to aldehyde pairs (i.e., between CH_3CHO, CH_2OHCHO, C_2H_5CHO). Methanol (CH_3OH) also have several highly correlated species, which are C_2H_5OH, CH_3OCH_3, and CH_3CHO. Ethanol (C_2H_5OH) belongs to the same group as methanol (alcohol), and the other two species are observed to have very constant ratios over methanol in both gas and ice (Chapters 2 and 4). The mean correlation coefficient is 0.85.

The five acetone-related species are in general highly correlated ($r > 0.8$) with each other. The highest correlated species with acetone are ketene and propanal. Ketene is also strongly correlated with acetaldehyde and propanal, which supports the chemical network suggested by Sung & Tidwell (1998). Propyne is less correlated with other species, but still more correlated than about half of the O-COM pairs shown in the right panel of Fig. 3.9. From an observational point of view, the discussed formation pathways of acetone in this section are possible to exist. However, more evidence is needed, in particular from the theoretical or experimental studies with the following focuses:

- destruction mechanisms and rates of O-COMs in the gas-phase, especially for aldehydes;
- theories or chemical networks that can properly reproduce the observed high gas-phase abundances of CH_3OCH_3, CH_3OCHO, and CH_3COCH_3;
- the verification or exploration of the formation pathways of acetone and other three-carbon (or even larger) O-COMs that have been detected in space.

3.6 Conclusions

In this work, we continue the observational studies by Chapter 2 on six two-carbon O-COMs in the CoCCoA sample. We derived the physical parameters (N, $T_{\rm ex}$, FWHM, and $v_{\rm lsr}$) of four additional species in the gas phase, which are acetone, propanal, ketene, and propyne. Methanol was revisited by directly fitting its optically thin lines instead of inferring from its minor isotopolog. We compared our gas-phase results with those of ice observations by JWST and astrochemical simulations, which triggered a series of discussion about the formation mechanisms of O-COMs (with a special focus on acetone) and some lingering questions in this field. The conclusions are summarized below:

- Acetone, ketene, and propyne are firmly detected in a dozen high-mass protostars as part of the CoCCoA program. Upper limits of propanal abundances were constrained.

- There are similarities in $T_{\rm ex}$, $v_{\rm lsr}$, and the emission morphology of acetone, ketene, and propanal to other O-COMs (represented by methanol), suggesting the same origin from hot cores. Unlike O-COMs, propyne has a slightly different $v_{\rm lsr}$ and tends to trace the more extended outflow cavity walls as inferred from the emission maps.

- The gas-phase column density ratios of O-COMs with respect to methanol in a combined sample of 12 high-mass CoCCoA sources and five low-mass sources from literature show systematically lower abundances of aldehydes (CH_3CHO, CH_2OHCHO, and C_2H_5CHO) than other O-COMs with the same amount of carbon atoms. In contrast, CH_3O-bearing molecules (CH_3OCH_3 and CH_3OCHO) show higher abundances than alcohols (C_2H_5OH and $(CH_2OH)_2$). These observational trends hint at more efficient destruction of aldehydes (likely in the gas phase) and more efficient formation of CH_3O-bearing O-COMs (preferably in the solid phase).

- The column density ratios between acetone and methanol are higher in the solid phase than in the gas phase by about one order of magnitude, which is in the same case as CH_3CHO and C_2H_5OH (Chapter 4). This suggests that acetone may go through some destruction that is more severe than methanol in the gas phase. However, more detection and quantification of acetone ice is need in a larger sample.

- The state-of-the-art astrochemical model `MAGICKAL` includes two solid-phase and one gas-phase formation pathways of acetone. The contribution of gas-phase production increases with the warm-up timescale. The `MAGICKAL` simulations slightly overproduce the ratios between acetone and methanol, which may due to the mediated energy barriers of some relevant reactions.

- Previous theoretical and experimental studies suggest that acetone along with other C_3H_6O isomers can be formed via ketene ($CH_2CO + CH_3$) and propylene ($C_3H_6 + O$). These pathways are plausible given that the observed gas-phase ratios between acetone and ketene or propyne (a precursor of propylene) are

not very scattered and can be well reproduced by simulations. However, these studies did not specifically focus on acetone or assume astrochemical conditions, and more direct evidence is still lacking to draw solid conclusions.

- It is urged for future COM studies, whether they are observations, simulations, or experiments, that acetone should receive adequate attention as an abundant three-carbon COM. Studying large COMs, especially those that have been detected in space, are very helpful for us to understand the broader picture of COM formation.

Observational results obtained for a variety of molecules and from a large sample of sources can bring us not only verification of existing theories, but also inspirations and guidance for future research. This observational study on acetone reveals the lack of consideration on this large but abundant molecule, and emphasizes the importance of resolving why simulations and experiments have so for not succeeded in reproducing the high abundance of CH_3O-bearing COMs. We look forward to more relevant investigation in the community, and hopefully those results will deepen our understanding of the chemical evolution of COMs in general.

Acknowledgements

This paper makes use of the following ALMA data: ADS/JAO.ALMA#2019.1.00246.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. Astrochemistry in Leiden is supported by the Nether- lands Research School for Astronomy (NOVA), by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 101019751 MOLDISK), by the Dutch Research Council (NWO) grants TOP-1 614.001.751 and 618.000.001, and by the Danish National Research Foundation through the Center of Excellence "InterCat" (Grant agreement no.: DNRF150). The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. Y.C. acknowledges Dijia Zou for the assistance in understanding chemical theories and linking them with astrochemical studies, and Julia Santos for the discussion on astrochemical experiments. R.T.G. thanks the National Science Foundation for funding through the Astronomy & Astrophysics program (grant number 2206516).

Appendix

3.A Supplementary figures for Section 3.5

Figures 3.A.1–3.A.2 provide more details about the MAGICKAL simulation results of acetone and other relevant species discussed in this paper. Figure 3.A.1 shows the net rate of change of acetone abundance as a function of time, from which we can learn at which stage acetone is efficiently produced or destroyed. Figure 3.A.2 shows the abundances of five molecules involved in the discussion (methanol, ketene, acetaldehyde, acetone, and propanal) in both ice and gas as a function of time. This can be seen as an integration results of Fig. 3.A.1.

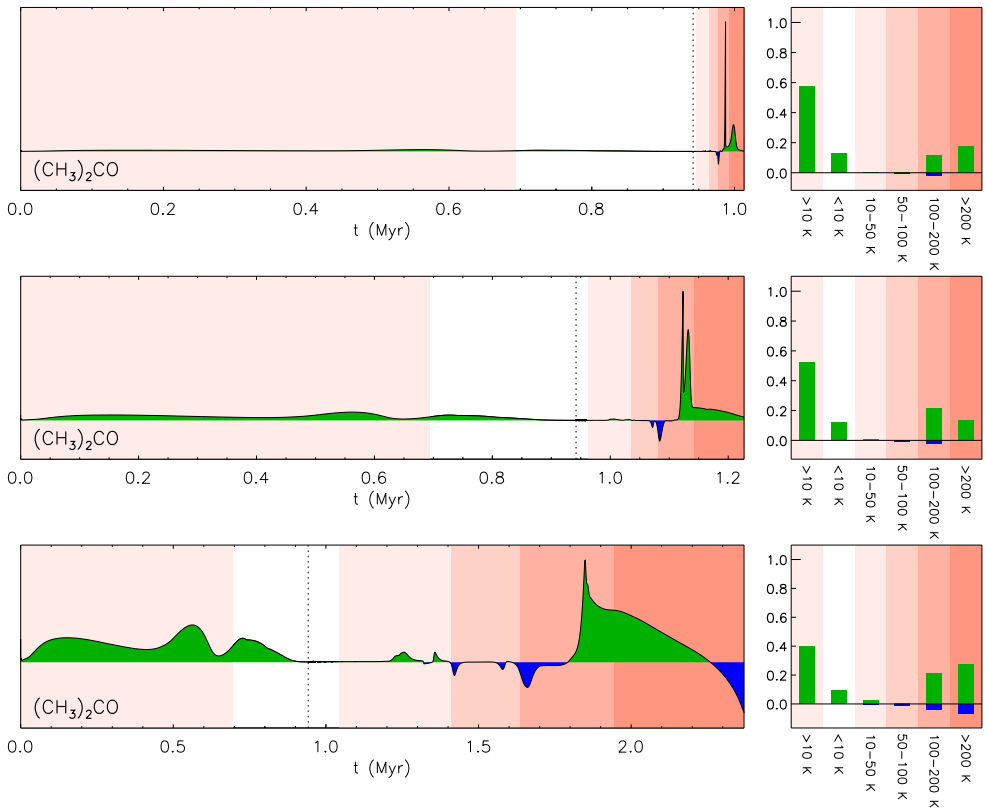


Figure 3.A.1: Net rate of change (arbitrary units) of acetone abundance as a function of time simulated by MAGICKAL. Here an updated version of MAGICKAL used in (Belloche et al. 2022) is applied . The three rows from top to bottom correspond to the simulations with a fast (5×10^4 yr), medium (2×10^5 yr), and slow (1×10^6 yr) warm-up stage, respectively. In each of the left panels, the vertical dotted line indicate the start of the warm-up stage. The panels on the right side show the net rates of change that are normalized and integrated over each temperature range.

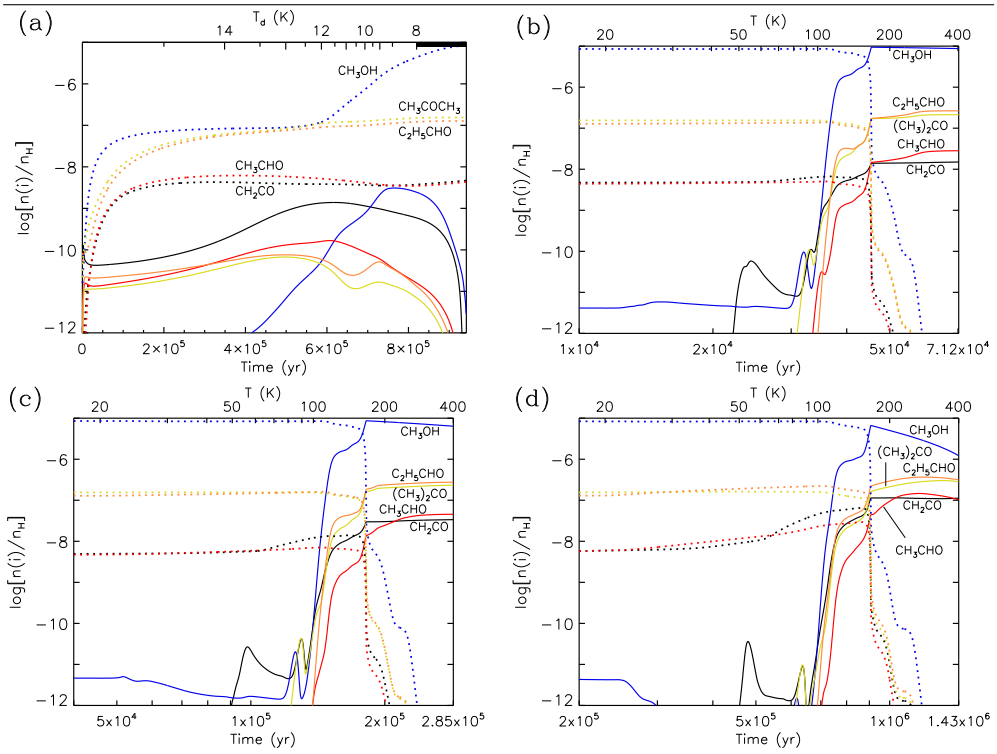


Figure 3.A.2: Abundances of CH_3OH (blue), CH_2CO (black), CH_3CHO (red), CH_3COCH_3 (yellow), and C_2H_5CHO (orange) with respect to total H abundance during the collapse stage (a) and subsequent warm-up stage with fast (b), medium (c), and slow (d) timescales. The corresponding temperatures are indicated by the upper x-axis in each panel. The abundances in the gas phase and the solid phase are plotted in solid and dashed lines, respectively. The intersection of the solid and dashed lines in the same color (i.e., ice and gas abundances of the same species) corresponds to the sublimation of this species from ice to gas.

3.B Best-fit physical parameters of molecules involved in this work in addition to acetone

Table 3.B.1 is an extension of Table 3.1, listing the best-fit physical parameters of three acetone-related species (*s*-propanal, ketene, and propyne) along with methanol, which was revisited in this work.

Table 3.B.1: Best-fit parameters of methanol, propanal, and ketene in 12 CoCCoA sources.

Source	**CH_3OH** (CDMS)				***s*-C_2H_5CHO** (CDMS)				**CH_2CO (CDMS)**				**CH_3CCH (JPL)**			
	N (cm^{-2})	T_{ex} (K)	FWHM (km s^{-1})	v_{lsr}	N (cm^{-2})	T_{ex} (K)	FWHM (km s^{-1})	v_{lsr}	N (cm^{-2})	T_{ex} (K)	FWHM (km s^{-1})	v_{lsr}	N (cm^{-2})	T_{ex} (K)	FWHM (km s^{-1})	v_{lsr}
G19.01-0.03	$1.7^{+0.6}_{-0.4}$ (18)	162±10	3.5	63.0	<1.0 (15)	[150]	3.0	61.0	1.5±*0.3* (15)	[100]	3.5	62.5	6.0±*1.5* (15)	[200]	3.0	62.3
	$3.2^{+1.4}_{-1.2}$ (18)	165±15	5.5	58.0					4.5±*1.0* (15)	[100]	4.0	58.5	$7.6^{+0.6}_{-0.5}$ (15)	80±10	3.6	58.5
G19.88-0.53	$5.5^{+1.3}_{-0.9}$ (18)	158±7	4.0	46.0	<5.0 (14)	[150]	3.0	44.0	3.7±*1.0* (15)	[100]	3.2	46.2	$6.9^{+0.7}_{-0.6}$ (15)	145±20	3.2	45.5
G22.04+0.22	$2.0^{+0.5}_{-0.3}$ (19)	135±5	6.0	52.2	<4.0 (15)	[150]	4.5	51.5	2.5±0.5 (16)	[100]	7.0	52.2	4.9±0.4 (16)	135±15	7.0	51.7
G23.21-0.37	2.8±0.5 (19)	158±8	5.5	77.0	<2.0 (15)	[150]	3.3	76.8	2.0±*0.4* (16)	[100]	4.0	77.2	$5.8^{+0.9}_{-0.6}$ (16)	180±20	5.2	77.8
									8.0±*1.6* (15)	[100]	4.0	81.3				
G34.30+0.20 [b]	1.2±0.3 (19)	130±5	3.0	56.0	<1.8 (15)	[150]	3.5	56.5	1.3±*0.3* (16)	[100]	3.8	56.2	$2.9^{+0.6}_{-0.3}$ (16)	155^{+25}_{-20}	3.2	56.2
	$5.5^{+1.4}_{-1.2}$ (18)	127±5	2.8	59.0												
G34.41+0.24	$1.7^{+0.6}_{-0.4}$ (19)	130±6	3.5	60.0												
	6.0±*1.5* (18)	[130]	[3.5]	56.8	<1.0 (15)	[150]	[3.0]	58.5	1.6±*0.3* (16)	[100]	5.5	58.3	3.2±0.2 (16)	106±10	5.7	57.5
G345.5+1.5	1.3±0.3 (18)	120±6	2.5	-15.8	<2.5 (14)	[150]	[2.0]	-16.0	$8.3^{+1.7}_{-0.5}$ (14)	75^{+45}_{-30}	2.3	-15.9	$9.5^{+1.4}_{-1.2}$ (15)	120^{+30}_{-20}	2.1	-15.3
									3.0±*0.6* (14)	[100]	2.0	-13.3	$2.8^{+0.7}_{-0.5}$ (15)	55±10	1.5	-12.7
													$8.0^{+12}_{-4.0}$ (14)	40^{+40}_{-20}	1.4	-10.0
G35.03+0.35	$2.6^{+1.4}_{-1.1}$ (18)	150±15	3.6	44.2	<6.0 (14)	[150]	3.5	47.5	1.8±0.4 (15)	[100]	3.5	46.0	6.9±0.5 (15)	90±15	3.5	47.6
													4.0±*1.0* (15)	[200]	2.5	53.0
G35.20-0.74N	1.0±0.2 (19)	146±5	5.0	31.5	<1.0 (15)	[150]	3.5	32.0	1.0±*0.2* (16)	[100]	4.5	31.8	4.8±*1.0* (15)	[200]	2.2	31.0
													$4.2^{+1.5}_{-0.7}$ (15)	135^{+50}_{-30}	2.1	33.5
IRAS 16547-4247	2.5±0.6 (18)	180±10	3.3	-35.2	<1.5 (15)	[150]	3.3	-35	9.0±*2.0* (15)	[100]	4.0	-35.3	$1.7^{+0.2}_{-0.1}$ (16)	130^{+20}_{-10}	3.3	-34.5
													6.9±1.0 (15)	115^{+35}_{-25}	2.6	-31.4
IRAS 18151-1208	$1.4^{+1.8}_{-0.7}$ (17)	150±30	2.2	35.0	<1.5 (14)	[150]	2.0	35.2	$4.2^{+1.4}_{-0.5}$ (14)	50^{+50}_{-20}	2.0	35.2	$2.4^{+0.4}_{-0.3}$ (15)	60^{+20}_{-10}	2.5	34.7
NGC 6334-38	6.8±2.3 (18)	133±10	3.8	-5.0	<6.0 (14)	[150]	3.0	-6.2	3.5±*0.7* (15)	120±20	4.0	-4.8	$1.6^{+1.5}_{-1.0}$ (16)	80^{+15}_{-10}	3.0	-3.5

Values in italics were estimated by visual inspection.

3.C Transitions covered in the CoCCoA dataset

Table 3.C.1: Transitions of acetone, propanal, ketene, propyne, and methanol covered in the CoCCoA dataset.

Frequency[(a)] (MHz)	Transition J K_a K_c s – J K_a K_c s	E_{up} (K)	$A_{ij}^{(b)}$ (s^{-1})	Frequency (MHz)	Transition J K_a K_c s – J K_a K_c s	E_{up} (K)	A_{ij} (s^{-1})
CH_3COCH_3 (JPL)		**<600**	**>5.0(-5)**	(Groner et al. 2002)			
238042.834**	32 5 27 0 – 32 5 28 1	317.01	1.05(-4)	241324.624	17 8 9 1 – 16 9 8 1	118.62	2.73(-4)
	32 6 27 0 – 32 4 28 1	317.01	1.05(-4)	241331.223	17 8 9 1 – 16 9 8 2	118.62	2.73(-4)
238044.288**	31 4 27 1 – 31 4 28 2	288.66	9.81(-5)	241420.927*	17 8 9 0 – 16 9 8 1	118.58	2.73(-4)
	31 5 27 1 – 31 3 28 2	288.66	9.81(-5)	241513.653	17 8 9 0 – 16 9 8 0	118.55	2.72(-4)
238044.373**	31 5 27 1 – 31 4 28 1	288.66	9.81(-5)	241559.887	13 10 4 0 – 12 9 3 0	77.71	3.30(-4)
	31 4 27 1 – 31 3 28 1	288.66	9.81(-5)	258070.737	17 10 8 1 – 16 9 7 2	121.93	2.63(-4)
238271.013	30 3 27 1 – 30 2 28 2	261.18	7.71(-5)	258113.895	17 10 8 1 – 16 9 7 1	121.93	2.63(-4)
	30 4 27 1 – 30 3 28 2	261.18	7.71(-5)	258338.038	17 10 8 0 – 16 9 7 1	121.89	2.63(-4)
238271.066	30 3 27 1 – 30 2 28 1	261.18	7.71(-5)	258461.738	14 10 5 1 – 13 9 5 2	87.78	5.70(-5)
	30 4 27 1 – 30 3 28 1	261.18	7.71(-5)	258472.032*†	25 1 24 1 – 24 2 23 1	168.79	7.85(-4)
238293.360	32 5 27 0 – 32 4 28 0	317.00	1.18(-4)		25 2 24 1 – 24 1 23 1	168.79	7.85(-4)
	32 6 27 0 – 32 5 28 0	317.00	1.18(-4)	258472.058*†	25 1 24 1 – 24 2 23 2	168.79	7.85(-4)
238338.954*	31 5 27 0 – 31 4 28 1	288.63	9.80(-5)		25 2 24 1 – 24 1 23 2	168.79	7.85(-4)
	31 4 27 0 – 31 3 28 1	288.63	9.80(-5)	258493.675*†	25 1 24 0 – 24 1 23 1	168.71	1.61(-4)
238354.973	17 10 7 0 – 16 11 6 0	124.02	9.37(-5)		25 1 24 0 – 24 2 23 1	168.71	6.24(-4)
238473.550	29 2 27 1 – 29 2 28 2	234.56	5.40(-5)		25 2 24 0 – 24 2 23 1	168.71	1.61(-4)
	29 3 27 1 – 29 1 28 2	234.56	5.40(-5)		25 2 24 0 – 24 1 23 1	168.71	6.24(-4)
238473.562	29 2 27 1 – 29 1 28 1	234.56	5.40(-5)	258515.262*†	25 2 24 0 – 24 2 23 0	168.63	7.85(-4)
	29 3 27 1 – 29 2 28 1	234.56	5.39(-5)		25 1 24 0 – 24 1 23 0	168.63	7.85(-4)
238611.445	17 10 7 0 – 16 11 6 1	124.07	9.40(-5)	258582.793	17 10 8 0 – 16 9 7 0	121.86	2.63(-4)
238633.098	31 5 27 0 – 31 4 28 0	288.61	9.90(-5)	259392.224	18 9 9 1 – 17 10 8 1	134.38	3.02(-4)
	31 4 27 0 – 31 3 28 0	288.61	9.90(-5)	259406.977	18 9 9 1 – 17 10 8 2	134.38	3.02(-4)
238695.516	13 10 4 1 – 12 9 3 2	77.73	6.34(-5)	259480.221	18 9 9 0 – 17 10 8 1	134.34	3.01(-4)
238708.138	17 10 7 1 – 16 11 6 1	124.11	9.57(-5)	259560.545	18 9 9 0 – 17 10 8 0	134.31	3.01(-4)
238847.035	23 1 22 1 – 22 1 21 1	144.45	6.14(-4)	259612.192*†	26 0 26 1 – 25 1 25 1	171.05	8.38(-4)
	23 2 22 1 – 22 2 21 1	144.45	6.14(-4)		26 1 26 1 – 25 0 25 1	171.05	8.38(-4)
238847.067*†	23 1 22 1 – 22 2 21 2	144.45	6.14(-4)	259612.241*†	26 0 26 1 – 25 0 25 2	171.05	8.38(-4)
	23 2 22 1 – 22 1 21 2	144.45	6.14(-4)		26 1 26 1 – 25 1 25 2	171.05	8.38(-4)
238863.500	29 2 27 0 – 29 2 28 1	234.51	5.19(-5)	259618.411*†	26 1 26 0 – 25 1 25 1	170.95	8.34(-4)
	29 3 27 0 – 29 1 28 1	234.51	5.19(-5)		26 0 26 0 – 25 0 25 1	170.95	8.34(-4)
238868.953‡	23 2 22 0 – 22 2 21 1	144.37	5.90(-4)	259624.520*†	26 0 26 0 – 25 1 25 0	170.85	8.38(-4)
	23 1 22 0 – 22 1 21 1	144.37	5.81(-4)		26 1 26 0 – 25 0 25 0	170.85	8.38(-4)
238890.805*†	23 2 22 0 – 22 2 21 0	144.29	6.14(-4)	260640.610	18 10 8 0 – 17 11 7 0	137.29	1.97(-4)
	23 1 22 0 – 22 1 21 0	144.29	6.14(-4)	260702.036	18 10 8 0 – 17 11 7 1	137.33	1.99(-4)
238952.610	30 3 27 0 – 30 3 28 0	261.09	7.79(-5)	260715.290	18 10 8 1 – 17 11 7 1	137.36	2.00(-4)
	30 4 27 0 – 30 2 28 0	261.09	7.79(-5)	260811.801	18 10 8 1 – 17 11 7 2	137.36	2.00(-4)
239022.408	17 10 7 1 – 16 11 6 2	124.11	9.33(-5)	261054.804	13 13 1 1 – 12 12 1 2	86.98	7.95(-4)
239253.257	29 2 27 0 – 29 2 28 0	234.45	5.46(-5)	261078.200	19 8 11 1 – 18 9 10 1	143.24	4.10(-4)
	29 3 27 0 – 29 1 28 0	234.45	5.46(-5)	261078.367	19 8 11 1 – 18 9 10 2	143.24	4.10(-4)
239984.730	24 0 24 1 – 23 0 23 1	146.60	6.59(-4)	261082.377	13 13 1 1 – 12 12 0 1	87.10	7.96(-4)
	24 1 24 1 – 23 1 23 1	146.60	6.59(-4)	261082.552	13 13 0 1 – 12 12 1 1	87.10	7.96(-4)
239984.779**	24 1 24 1 – 23 1 23 2	146.60	6.59(-4)	261110.448*†	13 13 0 1 – 12 12 0 2	87.22	7.96(-4)
	24 0 24 1 – 23 0 23 2	146.60	6.59(-4)	261169.099	19 9 11 1 – 18 8 10 2	143.24	4.10(-4)
239991.110‡	24 1 24 0 – 23 1 23 1	146.50	6.23(-4)	261169.839	19 9 11 1 – 18 8 10 1	143.24	4.10(-4)
	24 0 24 0 – 23 0 23 1	146.50	6.23(-4)	261175.065**	19 8 11 0 – 18 9 10 1	143.21	4.10(-4)
239997.383**	24 0 24 0 – 23 1 23 0	146.40	6.59(-4)	261268.570	19 9 11 0 – 18 8 10 1	143.21	4.11(-4)
	24 1 24 0 – 23 0 23 0	146.40	6.59(-4)	261271.656	19 8 11 0 – 18 9 10 0	143.18	4.10(-4)
240475.005	13 10 4 0 – 12 9 3 1	77.73	1.52(-4)	261328.382*†	13 13 1 0 – 12 12 1 1	87.03	7.98(-4)
240711.378*†	12 12 1 1 – 11 11 1 2	74.45	6.17(-4)	261356.943**	13 13 0 0 – 12 12 0 1	87.16	7.98(-4)
240774.176*	12 12 1 1 – 11 11 0 1	74.57	6.18(-4)	261367.487	19 9 11 0 – 18 8 10 0	143.18	4.11(-4)
240774.719*	12 12 0 1 – 11 11 1 1	74.57	6.18(-4)	261397.989	14 11 4 0 – 13 10 3 1	90.39	6.99(-5)
240822.548	11 6 5 1 – 10 5 6 1	52.99	5.05(-5)	261602.431	13 13 1 0 – 12 12 0 0	87.09	8.01(-4)
240838.029*†	12 12 0 1 – 11 11 0 2	74.69	6.19(-4)	261602.595*	13 13 0 0 – 12 12 1 0	87.09	8.01(-4)
240861.231	11 6 5 1 – 10 5 6 2	52.99	5.05(-5)	261755.985	41 11 30 1 – 41 11 31 2	576.34	2.58(-4)
240941.741	11 6 5 0 – 10 5 6 1	52.91	5.07(-5)		41 12 30 1 – 41 10 31 2	576.34	2.58(-4)
240998.755*†	12 12 1 0 – 11 11 1 1	74.49	6.20(-4)	261756.121	41 12 30 1 – 41 11 31 1	576.34	2.58(-4)
241042.018	11 6 5 0 – 10 5 6 0	52.83	5.08(-5)		41 11 30 1 – 41 10 31 1	576.34	2.58(-4)

Table 3.C.1: continued.

Frequency (MHz)	Transition J K_a K_c s – J K_a K_c s	E_{up} (K)	A_{ij} (s^{-1})	Frequency (MHz)	Transition J K_a K_c s – J K_a K_c s	E_{up} (K)	A_{ij} (s^{-1})
241042.122	13 10 4 1 – 12 9 3 1	77.80	3.26(-4)	261795.495	41 11 30 0 – 41 11 31 1	576.38	1.96(-4)
241062.732**	12 12 0 0 – 11 11 0 1	74.61	6.20(-4)		41 12 30 0 – 41 11 31 1	576.38	1.47(-4)
241286.250*	12 12 1 0 – 11 11 0 0	74.53	6.22(-4)		41 11 30 0 – 41 10 31 1	576.38	1.47(-4)
241286.761*	12 12 0 0 – 11 11 1 0	74.53	6.22(-4)		41 12 30 0 – 41 10 31 1	576.38	1.96(-4)
***s*-C_2H_5CHO, v=0 (CDMS)**		**<500**	**>1.0(-4)**	(Zingsheim et al. 2017)			
238104.321	20 11 10 1 – 20 10 11 1	172.48	1.09(-4)	259183.141	40 5 36 0 – 40 4 37 0	423.50	1.04(-4)
238104.799	20 11 9 1 – 20 10 10 1	172.48	1.09(-4)	259367.912	26 12 15 1 – 26 11 16 1	256.58	1.58(-4)
238105.199	20 11 9 0 – 20 10 10 0	172.48	1.09(-4)		26 12 14 1 – 26 11 15 1	256.58	1.58(-4)
238105.214	20 11 10 0 – 20 10 11 0	172.48	1.09(-4)		26 12 14 0 – 26 11 15 0	256.58	1.58(-4)
238350.142	19 11 9 1 – 19 10 10 1	162.33	1.05(-4)		26 12 15 0 – 26 11 16 0	256.58	1.58(-4)
238350.636	19 11 8 1 – 19 10 9 1	162.34	1.05(-4)	259787.502*	25 12 14 1 – 25 11 15 1	243.37	1.55(-4)
238351.081	19 11 8 0 – 19 10 9 0	162.33	1.05(-4)		25 12 13 1 – 25 11 14 1	243.37	1.55(-4)
238351.086	19 11 9 0 – 19 10 10 0	162.33	1.05(-4)		25 12 13 0 – 25 11 14 0	243.37	1.55(-4)
238738.609	22 6 16 1 – 21 6 15 1	149.17	2.10(-4)		25 12 14 0 – 25 11 15 0	243.37	1.55(-4)
	22 6 16 0 – 21 6 15 0	149.17	2.10(-4)	260020.036	27 1 26 1 – 26 2 25 1	181.60	2.84(-4)
238908.897	23 4 20 1 – 22 4 19 1	148.51	2.18(-4)		27 1 26 0 – 26 2 25 0	181.60	2.84(-4)
	23 4 20 0 – 22 4 19 0	148.51	2.18(-4)	260021.633	27 2 26 1 – 26 2 25 1	181.60	2.89(-4)
239035.632	22 4 19 0 – 21 3 18 0	137.05	1.20(-4)		27 2 26 0 – 26 2 25 0	181.60	2.89(-4)
	22 4 19 1 – 21 3 18 1	137.05	1.20(-4)	260022.915	27 1 26 1 – 26 1 25 1	181.60	2.89(-4)
240270.504	24 2 22 1 – 23 3 21 1	153.34	1.79(-4)		27 1 26 0 – 26 1 25 0	181.60	2.89(-4)
	24 2 22 0 – 23 3 21 0	153.34	1.79(-4)	260024.456	27 2 26 1 – 26 1 25 1	181.60	2.84(-4)
240489.453	24 3 22 1 – 23 3 21 1	153.35	2.25(-4)		27 2 26 0 – 26 1 25 0	181.60	2.84(-4)
	24 3 22 0 – 23 3 21 0	153.35	2.25(-4)	260156.935	24 12 13 1 – 24 11 14 1	230.67	1.52(-4)
240630.703*	24 2 22 1 – 23 2 21 1	153.34	2.25(-4)		24 12 12 1 – 24 11 13 1	230.67	1.52(-4)
	24 2 22 0 – 23 2 21 0	153.34	2.25(-4)		24 12 12 0 – 24 11 13 0	230.67	1.52(-4)
240676.214	40 5 35 1 – 40 4 36 1	435.05	1.04(-4)		24 12 13 0 – 24 11 14 0	230.67	1.52(-4)
240677.547	40 5 35 0 – 40 4 36 0	435.05	1.04(-4)	260480.718*	23 12 12 1 – 23 11 13 1	218.49	1.48(-4)
240849.674	24 3 22 1 – 23 2 21 1	153.35	1.81(-4)		23 12 11 1 – 23 11 12 1	218.49	1.48(-4)
	24 3 22 0 – 23 2 21 0	153.35	1.81(-4)		23 12 11 0 – 23 11 12 0	218.49	1.48(-4)
240953.122	40 6 35 1 – 40 5 36 1	435.06	1.05(-4)		23 12 12 0 – 23 11 13 0	218.49	1.48(-4)
240954.428	40 6 35 0 – 40 5 36 0	435.06	1.05(-4)	260563.102	25 4 22 1 – 24 3 21 1	172.79	1.83(-4)
241051.548	11 6 6 0 – 10 5 5 0	53.11	1.39(-4)		25 4 22 0 – 24 3 21 0	172.79	1.83(-4)
241062.630	23 3 20 1 – 22 3 19 1	148.30	2.24(-4)	260763.017	22 12 11 1 – 22 11 12 1	206.82	1.43(-4)
	23 3 20 0 – 22 3 19 0	148.30	2.24(-4)		22 12 10 1 – 22 11 11 1	206.82	1.43(-4)
241083.814	11 6 5 0 – 10 5 6 0	53.11	1.39(-4)		22 12 10 0 – 22 11 11 0	206.82	1.43(-4)
241667.998*	25 1 24 1 – 24 2 23 1	157.08	2.23(-4)		22 12 11 0 – 22 11 12 0	206.82	1.43(-4)
	25 1 24 0 – 24 2 23 0	157.08	2.23(-4)	261007.063	21 12 10 1 – 21 11 11 1	195.66	1.38(-4)
241672.910*	25 2 24 1 – 24 2 23 1	157.08	2.31(-4)	261007.453	21 12 9 1 – 21 11 10 1	195.66	1.38(-4)
	25 2 24 0 – 24 2 23 0	157.08	2.31(-4)	261007.926	21 12 9 0 – 21 11 10 0	195.66	1.38(-4)
241676.670	25 1 24 1 – 24 1 23 1	157.08	2.31(-4)	261007.927	21 12 10 0 – 21 11 11 0	195.66	1.38(-4)
	25 1 24 0 – 24 1 23 0	157.08	2.31(-4)	261069.305	24 4 20 1 – 23 4 19 1	165.80	2.85(-4)
241681.621	25 2 24 1 – 24 1 23 1	157.08	2.23(-4)		24 4 20 0 – 23 4 19 0	165.80	2.85(-4)
	25 2 24 0 – 24 1 23 0	157.08	2.23(-4)	261218.275	20 12 9 1 – 20 11 10 1	185.01	1.33(-4)
258249.922	24 7 17 1 – 23 7 16 1	179.87	2.63(-4)		20 12 8 1 – 20 11 9 1	185.01	1.33(-4)
	24 7 17 0 – 23 7 16 0	179.87	2.63(-4)		20 12 8 0 – 20 11 9 0	185.01	1.33(-4)
258358.429	28 12 17 1 – 28 11 18 1	284.55	1.61(-4)		20 12 9 0 – 20 11 10 0	185.01	1.33(-4)
258358.434	28 12 16 0 – 28 11 17 0	284.55	1.63(-4)	261290.259**	28 0 28 1 – 27 1 27 1	184.39	3.29(-4)
258358.633	28 12 16 1 – 28 11 17 1	284.55	1.61(-4)	261290.267**	28 0 28 0 – 27 1 27 0	184.39	3.29(-4)
258359.161	28 12 17 0 – 28 11 18 0	284.55	1.63(-4)	261290.271**	28 1 28 1 – 27 1 27 1	184.39	2.96(-4)
258629.931*	25 3 22 1 – 24 3 21 1	172.70	2.78(-4)	261290.280**	28 1 28 0 – 27 1 27 0	184.39	2.96(-4)
	25 3 22 0 – 24 3 21 0	172.70	2.78(-4)	261290.282**	28 0 28 1 – 27 0 27 1	184.39	2.96(-4)
258761.193	26 2 24 1 – 25 3 23 1	177.76	2.34(-4)	261290.291**	28 0 28 0 – 27 0 27 0	184.39	2.96(-4)
	26 2 24 0 – 25 3 23 0	177.76	2.34(-4)	261290.294**	28 1 28 1 – 27 0 27 1	184.39	3.29(-4)
258840.070	26 3 24 1 – 25 3 23 1	177.76	2.82(-4)	261290.303**	28 1 28 0 – 27 0 27 0	184.39	3.29(-4)
	26 3 24 0 – 25 3 23 0	177.76	2.82(-4)	261397.487	19 12 8 1 – 19 11 9 1	174.88	1.26(-4)
258865.207	16 5 12 1 – 15 4 11 1	82.91	1.02(-4)	261397.901	19 12 7 1 – 19 11 8 1	174.88	1.26(-4)
	16 5 12 0 – 15 4 11 0	82.91	1.02(-4)	261398.461	19 12 7 0 – 19 11 8 0	174.88	1.26(-4)
258893.298*	26 2 24 1 – 25 2 23 1	177.76	2.82(-4)		19 12 8 0 – 19 11 9 0	174.88	1.26(-4)
	27 12 16 1 – 27 11 17 1	270.31	1.60(-4)	261549.823	18 12 7 1 – 18 11 8 1	165.25	1.18(-4)
	26 2 24 0 – 25 2 23 0	177.76	2.82(-4)	261550.248	18 12 6 1 – 18 11 7 1	165.25	1.18(-4)
	27 12 15 1 – 27 11 16 1	270.31	1.60(-4)	261550.849	18 12 6 0 – 18 11 7 0	165.25	1.18(-4)
	27 12 15 0 – 27 11 16 0	270.31	1.61(-4)		18 12 7 0 – 18 11 8 0	165.25	1.18(-4)
	27 12 16 0 – 27 11 17 0	270.31	1.61(-4)	261677.566*	17 12 6 1 – 17 11 7 1	156.14	1.09(-4)
258971.997	26 3 24 1 – 25 2 23 1	177.76	2.34(-4)	261678.002*	17 12 5 1 – 17 11 6 1	156.14	1.09(-4)
	26 3 24 0 – 25 2 23 0	177.76	2.34(-4)	261678.447*	13 6 8 1 – 12 5 7 1	65.82	1.50(-4)

Table 3.C.1: continued.

Frequency (MHz)	Transition	E_{up} (K)	A_{ij} (s^{-1})	Frequency (MHz)	Transition	E_{up} (K)	A_{ij} (s^{-1})
259161.713	40 4 36 1 – 40 3 37 1	423.50	1.04(-4)	261678.642*	17 12 5 0 – 17 11 6 0	156.14	1.09(-4)
259163.316	40 4 36 0 – 40 3 37 0	423.50	1.04(-4)		17 12 6 0 – 17 11 7 0	156.14	1.09(-4)
259181.673	40 5 36 1 – 40 4 37 1	423.50	1.04(-4)	261679.645*	13 6 8 0 – 12 5 7 0	65.82	1.54(-4)
CH_2CO (CDMS)		**<800**		(Johnson & Strandberg 1952; Fabricant et al. 1977; Brown et al. 1990)			
240185.794*†	12 1 12 – 11 1 11	87.99	1.55(-4)	260191.982*†	13 1 13 – 12 1 12	100.47	1.98(-4)
CH_3CCH (CDMS)		**<800**		(Cazzoli & Puzzarini 2008)			
238883.441	14 9 – 13 9	670.50	2.76(-5)	239179.281**	14 4 – 13 4	201.70	4.34(-5)
238960.695	14 8 – 13 8	548.00	3.17(-5)	239211.215**	14 3 – 13 3	151.14	4.51(-5)
239028.931	14 7 – 13 7	439.85	3.54(-5)	239234.034*†	14 2 – 13 2	115.02	4.63(-5)
239088.122	14 6 – 13 6	346.07	3.85(-5)	239247.728**	14 1 – 13 1	93.35	4.71(-5)
239138.245	14 5 – 13 5	266.68	4.12(-5)	239252.294**	14 0 – 13 0	86.12	4.73(-5)
CH_3OH, vt=0–2 (CDMS)		**<2000**	**>1.0(-7)**	(Xu et al. 2008)			
238440.805**	25 5 21 3 – 24 6 19 3	1171.71	4.20(-5)	240960.557‡	5 1 5 3 – 4 1 4 3	359.95	5.76(-5)
238440.974**	25 5 20 3 – 24 6 18 3	1171.71	4.20(-5)	241042.589*†	22 6 16 2 – 23 5 19 2	775.57	2.31(-5)
238665.976*	30 3 27 0 – 30 2 28 0	1128.27	8.47(-5)	241159.199‡	5 4 2 4 – 4 4 1 4	398.11	2.15(-5)
238723.283*	7 7 1 5 – 8 6 3 5	588.21	1.33(-6)	241166.580‡	5 3 2 4 – 4 3 1 4	452.14	3.86(-5)
238729.425*	31 3 28 0 – 31 2 29 0	1200.13	8.61(-5)	241178.445‡	5 4 1 3 – 4 4 0 3	515.80	2.18(-5)
238890.424**	29 3 26 0 – 29 2 27 0	1058.72	8.37(-5)		5 4 2 3 – 4 4 1 3	515.80	2.18(-5)
239142.749*†	32 3 29 0 – 32 2 30 0	1274.32	8.80(-5)	241179.886‡	5 3 3 5 – 4 3 2 5	357.36	3.83(-5)
239219.963	24 15 10 0 – 25 12 13 3	1806.78	7.08(-6)	241184.189‡	5 4 1 5 – 4 4 0 5	440.13	2.16(-5)
239219.963	24 15 9 0 – 25 12 14 3	1806.78	7.08(-6)	241187.428‡	5 2 4 5 – 4 2 3 5	399.34	5.07(-5)
239345.054*	28 3 25 0 – 28 2 26 0	991.49	8.29(-5)	241192.856‡	5 2 4 3 – 4 2 3 3	333.40	5.03(-5)
239397.997**	16 3 13 2 – 17 0 17 1	378.28	6.77(-7)	241196.430‡	5 2 3 3 – 4 2 2 3	333.40	5.03(-5)
239659.467	33 2 31 2 – 33 2 32 1	1332.63	7.22(-6)	241198.285‡	5 3 3 3 – 4 3 2 3	430.84	3.83(-5)
239746.219‡	5 1 5 0 – 4 1 4 0	49.06	5.66(-5)	241198.291‡	5 3 2 3 – 4 3 1 3	430.84	3.83(-5)
239971.367*	33 3 30 0 – 33 2 31 0	1350.82	9.04(-5)	241203.706‡	5 1 5 4 – 4 1 4 4	326.20	5.75(-5)
239977.050⋆	27 3 24 0 – 27 2 25 0	926.58	8.24(-5)	241206.035‡	5 0 5 4 – 4 0 4 4	335.31	6.00(-5)
240241.490‡	5 3 3 1 – 6 2 5 1	82.53	1.44(-5)	241210.764‡	5 2 3 4 – 4 2 2 4	434.63	5.04(-5)
240321.199*†	27 8 19 0 – 28 7 22 0	1196.59	2.17(-5)	241238.144‡	5 1 4 5 – 4 1 3 5	448.11	5.75(-5)
240321.205*†	27 8 20 0 – 28 7 21 0	1196.59	2.17(-5)	241267.862‡	5 0 5 3 – 4 0 4 3	458.39	6.00(-5)
240454.848	5 1 5 6 – 4 1 4 6	717.41	5.69(-5)	241283.133	34 3 31 0 – 34 2 32 0	1429.63	9.33(-5)
240738.926‡	26 3 23 0 – 26 2 24 0	863.98	8.21(-5)	241364.143*	5 1 4 6 – 4 1 3 6	717.54	5.75(-5)
240752.863*	5 3 2 8 – 4 3 1 8	949.67	3.79(-5)	241441.270‡	5 1 4 3 – 4 1 3 3	360.02	5.79(-5)
240757.889	5 2 4 6 – 4 2 3 6	911.34	5.05(-5)	241588.758‡	25 3 22 0 – 25 2 23 0	803.70	8.19(-5)
240757.920*	5 2 3 6 – 4 2 2 6	911.34	5.05(-5)	241615.470	30 0 30 1 – 29 3 26 2	1082.44	2.04(-6)
240784.498	5 4 1 7 – 4 4 0 7	920.95	2.13(-5)	241700.159‡	5 0 5 1 – 4 0 4 1	47.94	6.03(-5)
240817.972*	5 1 4 7 – 4 1 3 7	833.64	5.75(-5)	258142.379	16 5 12 4 – 16 3 13 4	744.14	1.16(-7)
240861.406**	5 4 2 8 – 4 4 1 8	779.22	2.18(-5)	258567.037	17 9 8 3 – 18 7 11 3	1037.76	7.43(-7)
240869.551*	5 0 5 7 – 4 0 4 7	768.92	5.97(-5)		17 9 9 3 – 18 7 12 3	1037.76	7.43(-7)
240916.172‡	5 3 3 6 – 4 3 2 6	692.78	3.87(-5)	258780.248‡	19 3 17 0 – 19 2 18 0	490.59	9.01(-5)
240916.173‡	5 3 2 6 – 4 3 1 6	692.78	3.87(-5)	259273.686‡	17 2 15 3 – 16 1 15 3	652.67	5.58(-5)
240932.051*	5 4 1 6 – 4 4 0 6	649.19	2.14(-5)	259581.398‡	24 1 23 1 – 24 0 24 1	717.01	4.91(-5)
	5 4 2 6 – 4 4 1 6	649.19	2.14(-5)	260064.318**	20 8 13 2 – 21 7 14 2	808.20	2.13(-5)
240936.742‡	5 2 3 8 – 4 2 2 8	679.63	5.00(-5)	260381.463‡	20 3 18 0 – 20 2 19 0	536.95	9.14(-5)
240938.974‡	5 0 5 6 – 4 0 4 6	542.89	5.98(-5)	260947.124	18 14 5 5 – 19 13 7 5	1645.96	1.95(-5)
240948.343‡	5 3 3 7 – 4 3 2 7	656.13	3.81(-5)	261061.320‡	21 4 18 2 – 20 5 16 2	623.60	3.02(-5)
240952.056‡	5 2 4 7 – 4 2 3 7	620.93	5.05(-5)	261704.409‡	12 6 7 1 – 13 5 8 1	359.77	1.78(-5)
240958.922‡	5 1 5 8 – 4 1 4 8	566.72	5.75(-5)				

Notes: (a) Denotations after frequency: For the robustly detected species, $\star\star$ = key transitions (i.e., unblended and optically thin) that can be used to constrain the column densities and excitation temperatures in most sources (might be blended in a few very bright sources); $\star$ = key transitions that can be used in sources without too much line blending; $\star\dagger$ = strong but blended transitions (not considered); $\ddagger$ = optically thick transitions. For the tentatively detected species, $**$ = key transitions that are used to determine the upper limits of column densities; $*$ = unblended but weaker transitions that are checked for anti-correlation in determining the upper limits.
(b) The format of A_{ij}: $m(n) = m \times 10^n$.

3.D Figures for spectral fitting results

Figures 3.D.1–3.D.12 present the best-fit LTE-modeled spectra of acetone, propanal, methanol, and propyne, overlaid on the ALMA spectra. For each species, a number of strong transitions are selected to be displayed, including both the unblended lines and the optically thick ones that are overfit by the models.

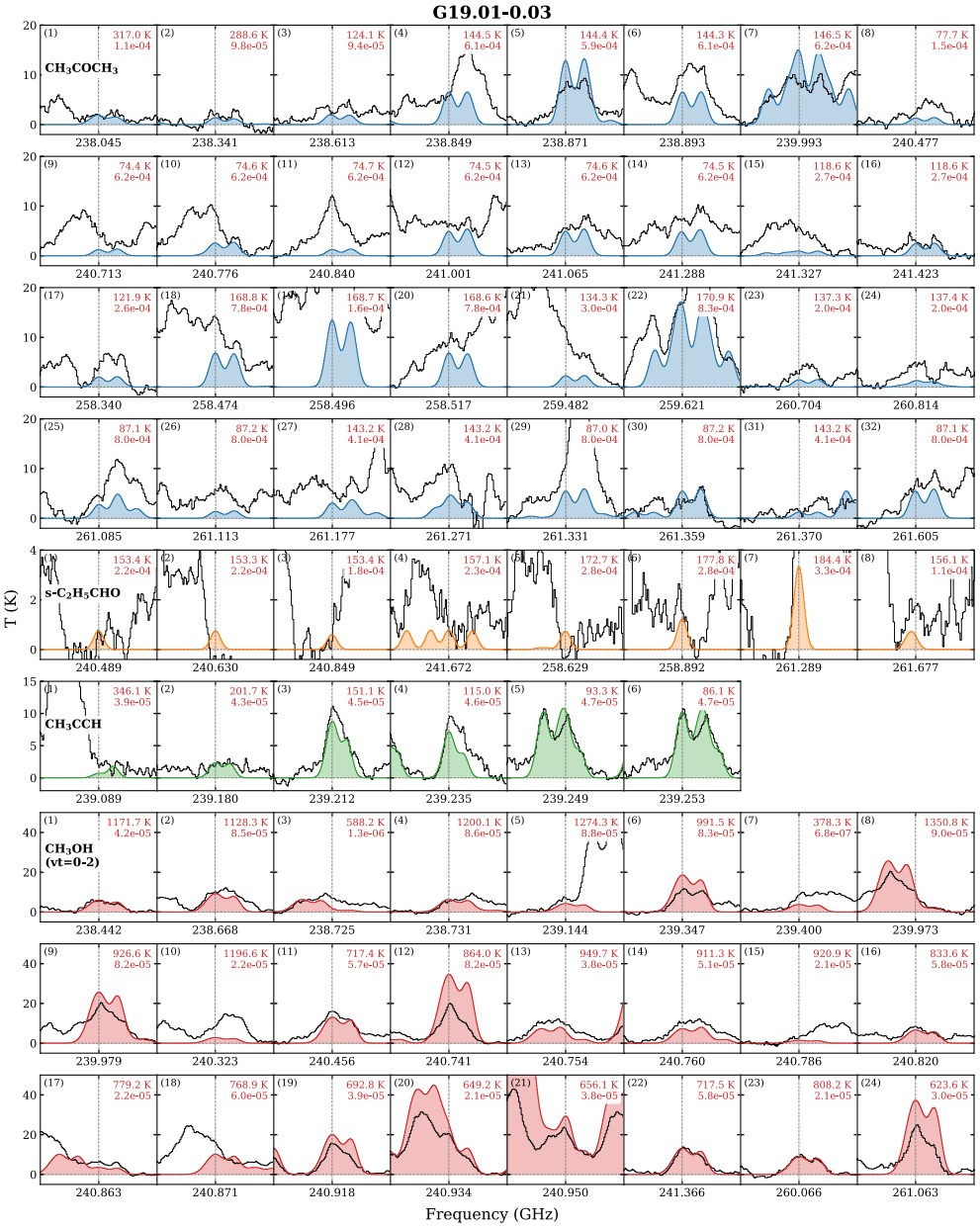


Figure 3.D.1: Best-fit LTE models of acetone (red), propanal (orange), propyne (green), and methanol (red) overlaid on the ALMA spectrum of G19.01-0.03 (black). The species name is indicated in bold black in the upper left in the first panel. The upper energy level and Einstein A coefficient are annotated in red in the upper right. We only labeled the central frequency of each line on the x axis, but the span of each panel is equal to [–15, +15] km s^{-1}. Not all transitions listed in Table 3.C.1 are displayed here; the selected transitions are either strong or not very blended. For methanol, the very optically thick transitions were now shown here, but there are still several lines (e.g., no. 9, 12, 19, 24) that are overestimated due to optical depth effects, especially in brighter sources.

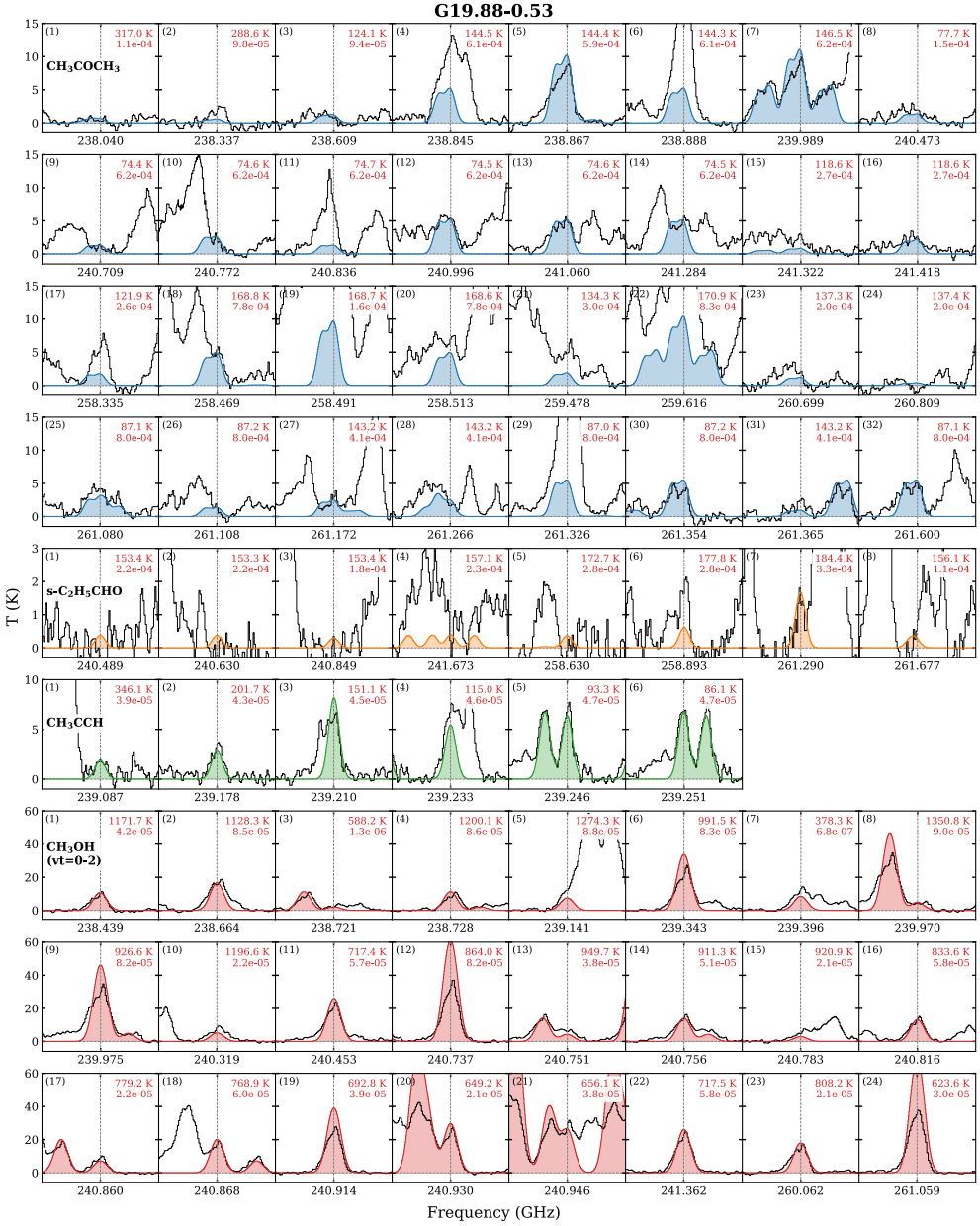


Figure 3.D.2: Same as Fig. 3.D.1 but for G19.88-0.53.

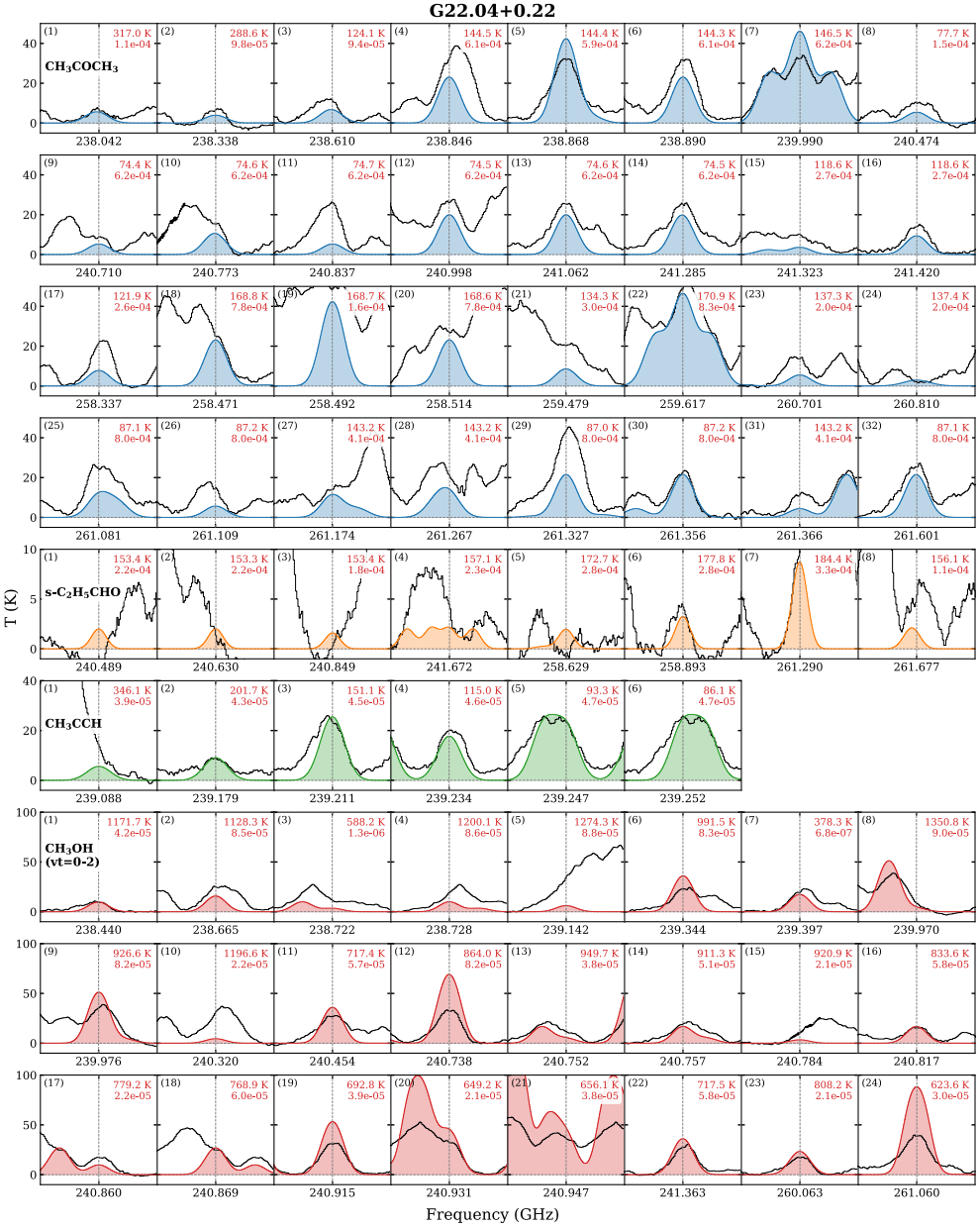


Figure 3.D.3: Same as Fig. 3.D.1 but for G22.04+0.22.

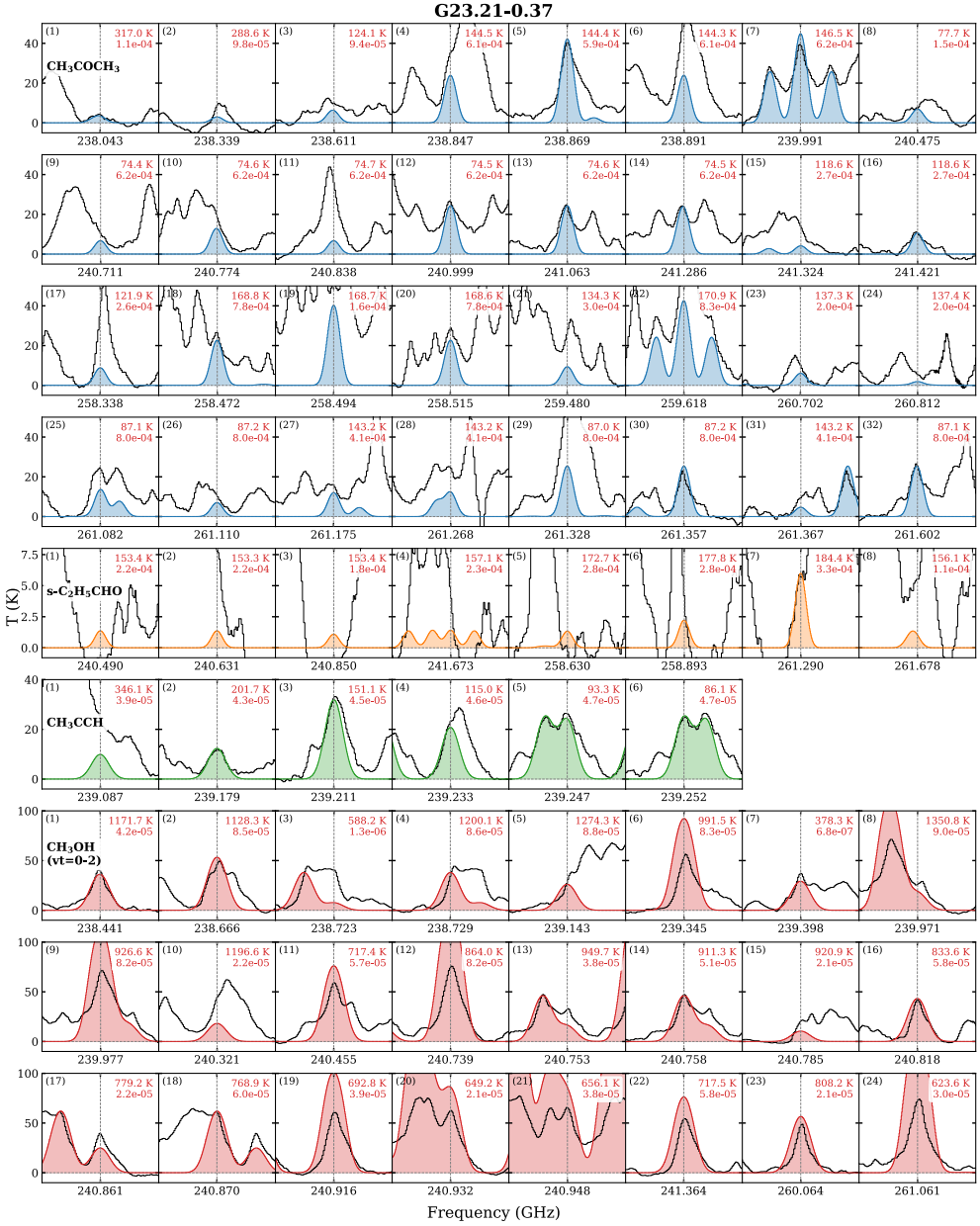


Figure 3.D.4: Same as Fig. 3.D.1 but for G23.21-0.37.

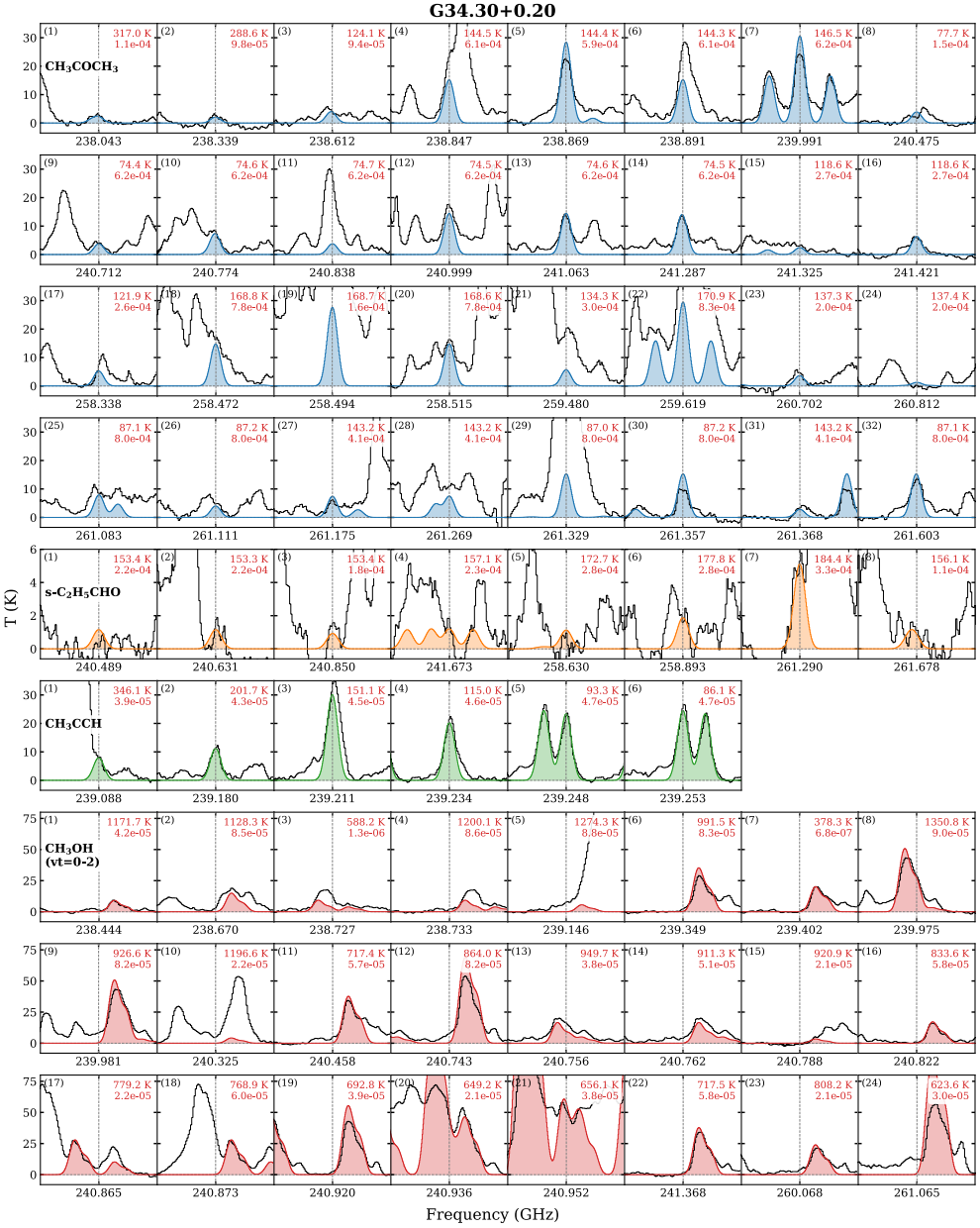


Figure 3.D.5: Same as Fig. 3.D.1 but for G34.30+0.20.

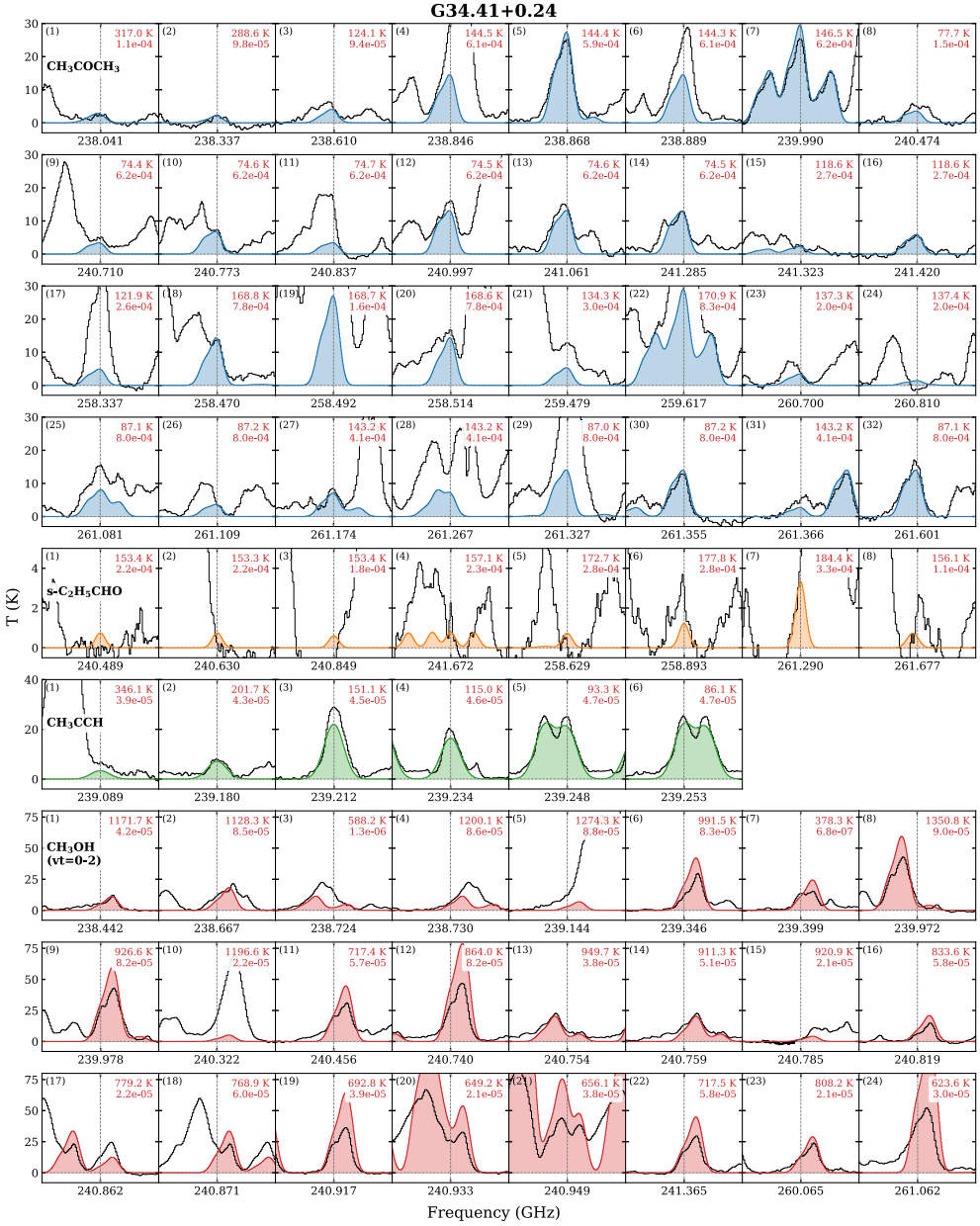


Figure 3.D.6: Same as Fig. 3.D.1 but for G34.41+0.24.

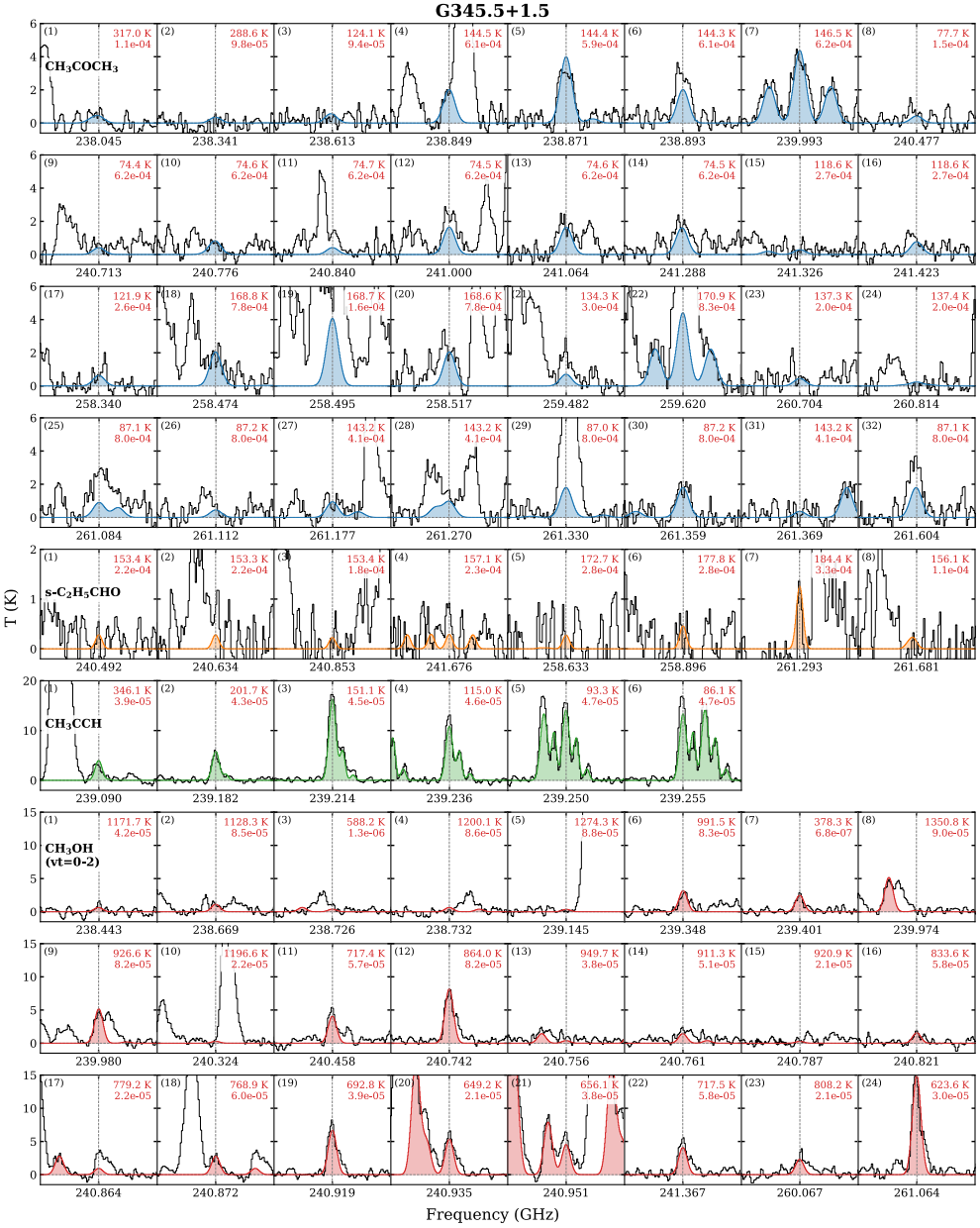


Figure 3.D.7: Same as Fig. 3.D.1 but for G345.5+1.5.

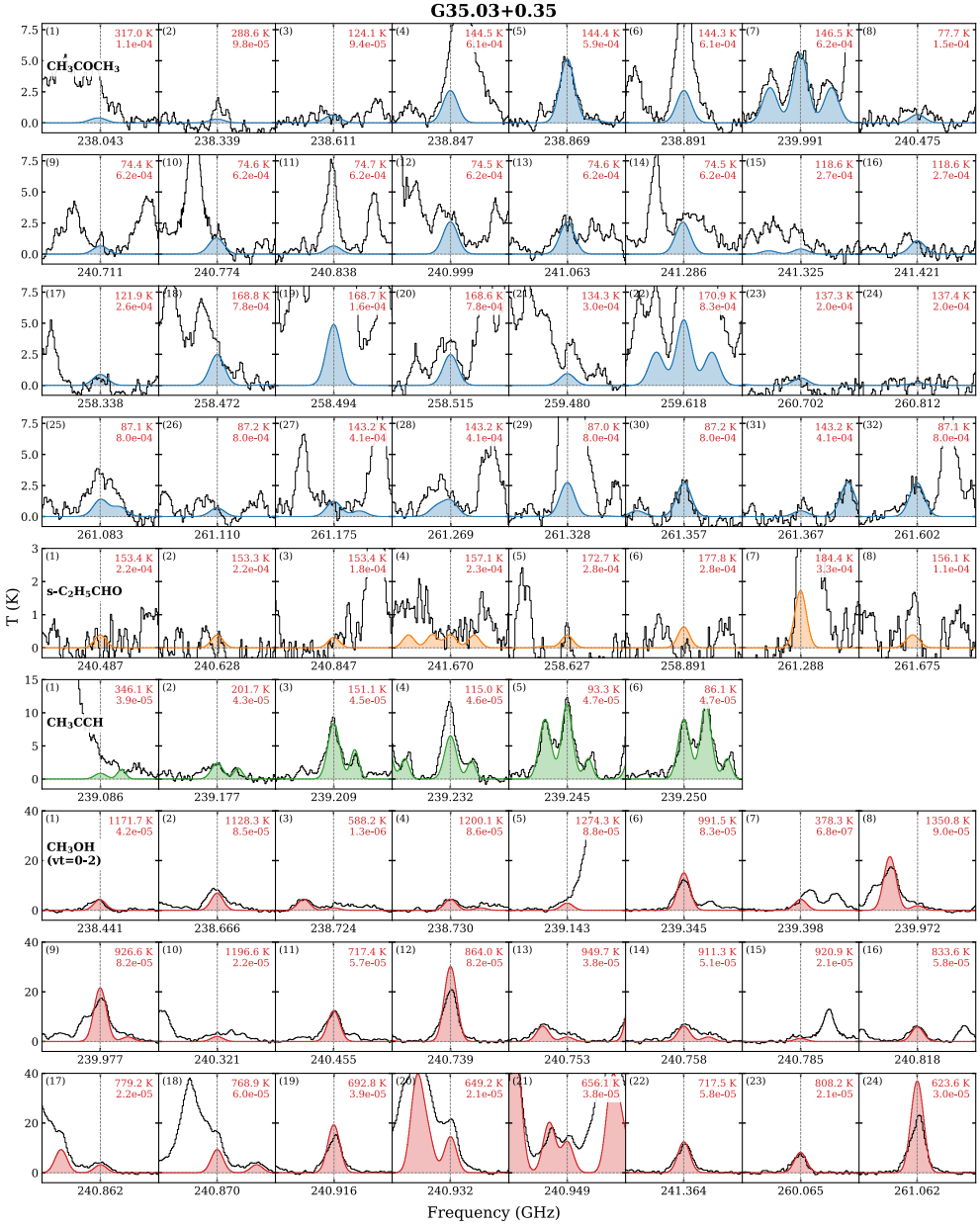


Figure 3.D.8: Same as Fig. 3.D.1 but for G35.03+0.35.

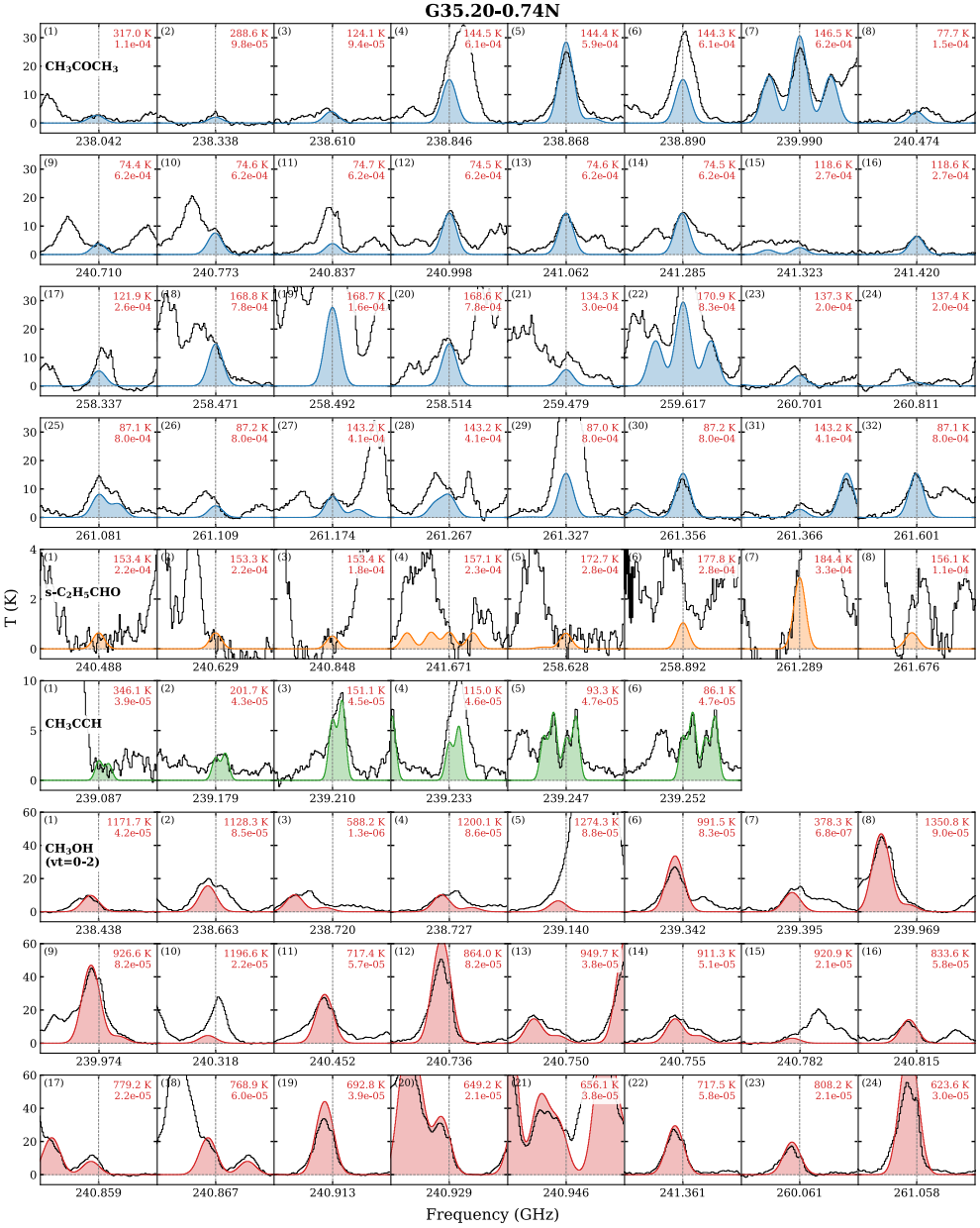


Figure 3.D.9: Same as Fig. 3.D.1 but for G35.20-0.74N.

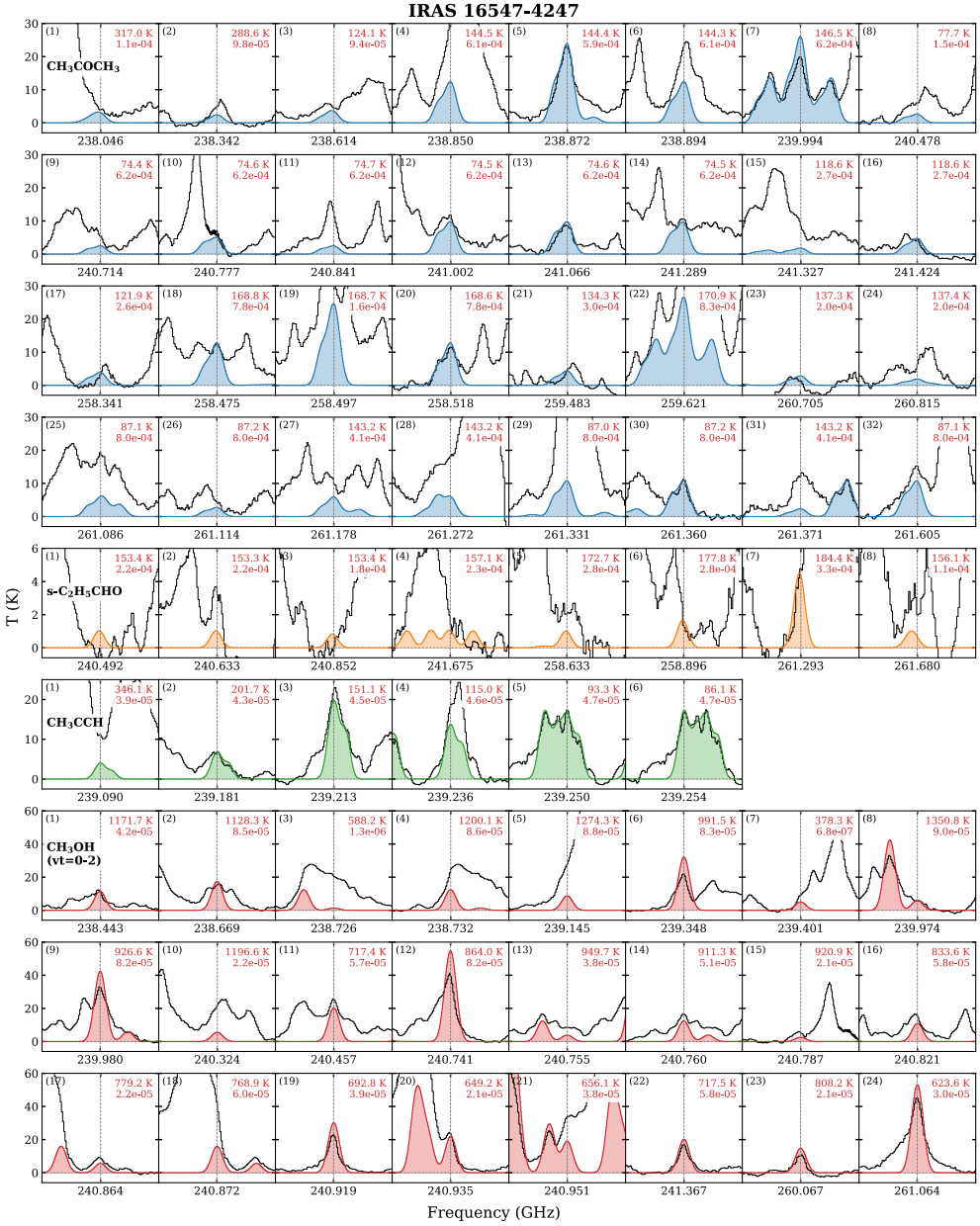


Figure 3.D.10: Same as Fig. 3.D.1 but for IRAS 16547-4247.

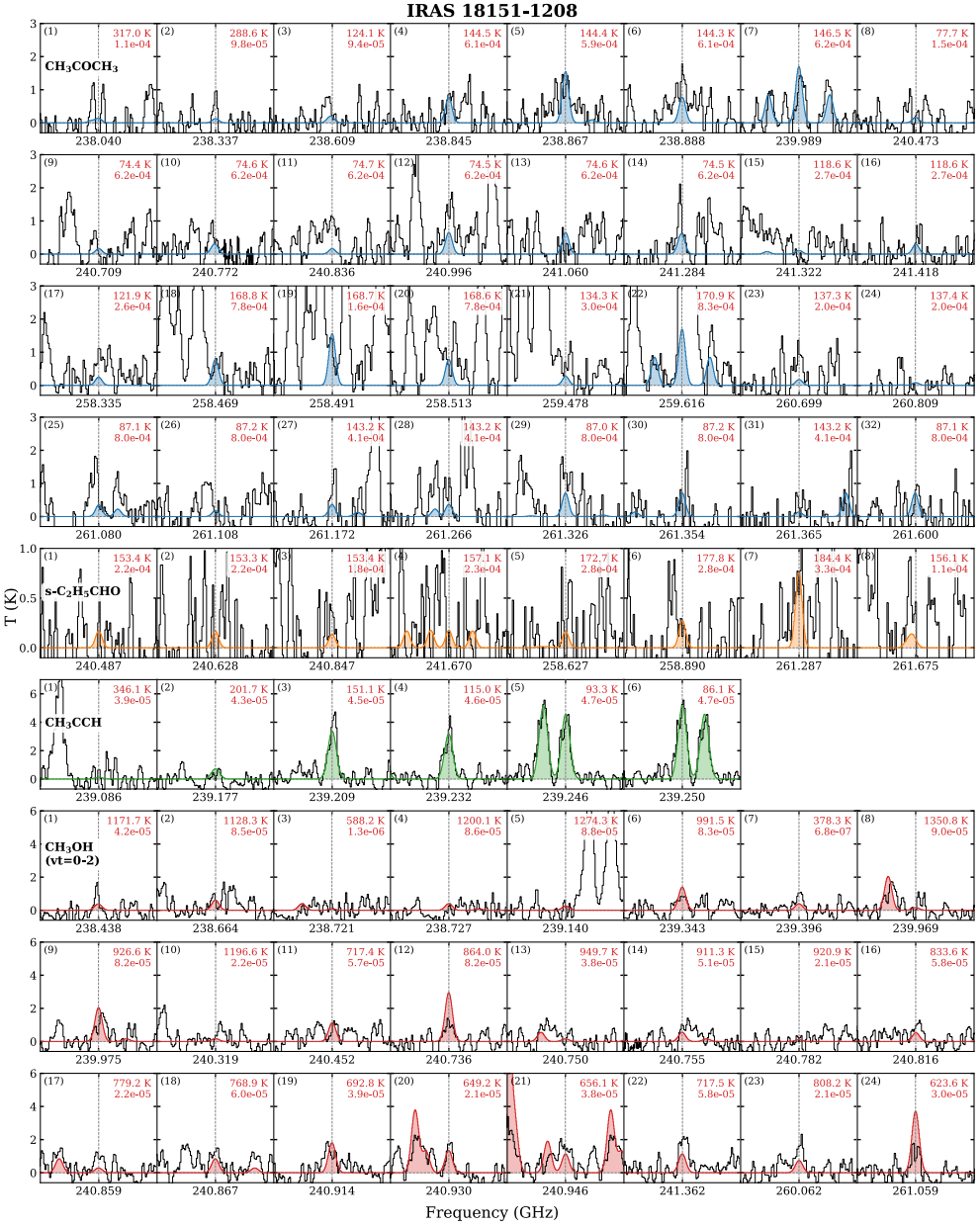


Figure 3.D.11: Same as Fig. 3.D.1 but for IRAS 18151-1208.

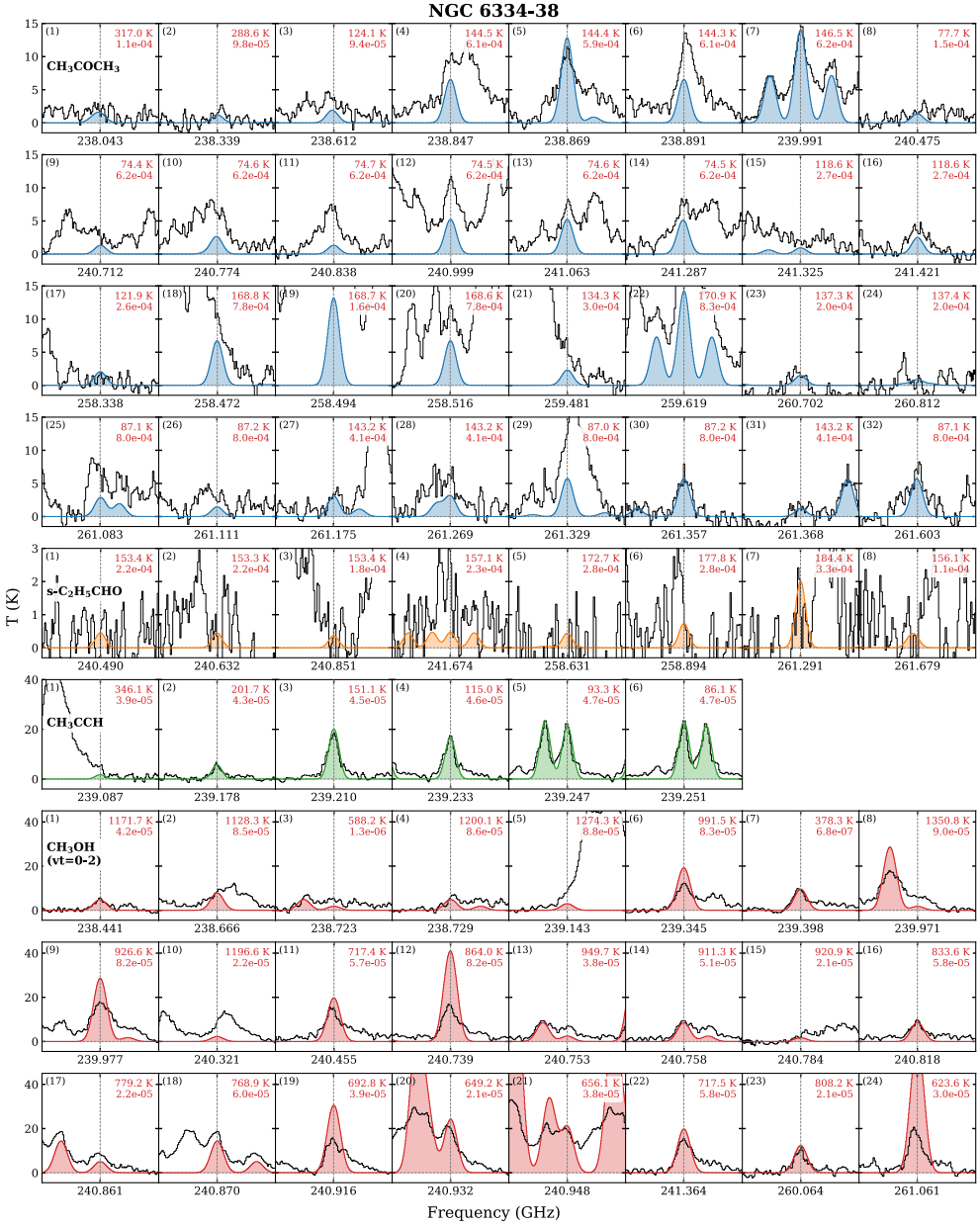


Figure 3.D.12: Same as Fig. 3.D.1 but for NGC 6334-38.

4 JOYS+: The link between the ice and gas of complex organic molecules

Comparing JWST and ALMA data of two low-mass protostars

Y. Chen, W. R. M. Rocha, E. F. van Dishoeck, M. L. van Gelder, P. Nazari, K. Slavicinska, L. Francis, B. Tabone, M. E. Ressler, P. D. Klaassen, H. Beuther, A. C. A. Boogert, C. Gieser, P. J. Kavanagh, G. Perotti, V. J. M. Le Gouellec, L. Majumdar M. Güdel, and Th. Henning

Astronomy & Astrophysics, 690, A205 (2024)

Abstract

Context. A rich inventory of complex organic molecules (COMs) has been observed in high abundances in the gas phase toward Class 0 protostars. It has been suggested that these molecules are formed in ices and sublimate in the warm inner envelope close to the protostar. However, only the most abundant COM, methanol (CH_3OH), had been firmly detected in ices before the era of the *James Webb* Space Telescope (JWST). Now, it is possible to detect the interstellar ices of other COMs and constrain their ice column densities quantitatively.

Aims. We aim to determine the column densities of several oxygen-bearing COMs (O-COMs) in both gas and ice for two low-mass protostellar sources, NGC 1333 IRAS 2A (hereafter IRAS 2A) and B1-c, as case studies in our JWST Observations of Young protoStars (JOYS+) program. By comparing the column density ratios with respect to CH_3OH between both phases measured in the same sources, we can probe the evolution of COMs from ice to gas in the early stages of star formation.

Methods. The column densities of COMs in gas and ice were derived by fitting the spectra observed by the Atacama Large Millimeter/submillimeter Array (ALMA) and the JWST/Mid-InfraRed Instrument-Medium Resolution Spectroscopy (MIRI-MRS), respectively. The gas-phase emission lines were fit using local thermal equilibrium models, and the ice absorption bands were fit by matching the infrared spectra measured in laboratories. The column density ratios of four O-COMs (CH_3CHO, C_2H_5OH, CH_3OCH_3, and CH_3OCHO) with respect to CH_3OH were compared between ice and gas in IRAS 2A and B1-c.

Results. We were able to fit the fingerprint range of COM ices between 6.8 and 8.8 μm in the JWST/MIRI-MRS spectra of B1-c using similar components to the ones recently used for NGC 1333 IRAS 2A. We claim detection of CH_4, OCN^-, $HCOO^-$, HCOOH, CH_3CHO, C_2H_5OH, CH_3OCH_3, CH_3OCHO, and CH_3COCH_3 in B1-c, and upper limits have been estimated for SO_2, CH_3COOH, and CH_3CN. The total abundance of O-COM ices is constrained to be 15% with respect to H_2O ice, 80% of which is dominated by CH_3OH. The comparison of O-COM ratios with respect to CH_3OH between ice and gas shows two different cases. On the one hand, the column density ratios of CH_3OCHO and CH_3OCH_3 match well between the two phases, which may be attributed to a direct inheritance from ice to gas or strong chemical links with CH_3OH. On the other hand, the ice ratios of CH_3CHO and C_2H_5OH with respect to CH_3OH are higher than the gas ratios by 1–2 orders of magnitude. This difference can be explained by gas-phase reprocessing following sublimation, or different spatial distributions of COMs in the envelope, which is an observational effect resulting from ALMA and JWST tracing different components in a protostellar system.

Conclusions. The firm detection of COM ices other than CH_3OH is reported in another well-studied low-mass protostar, B1-c, following the recent detection in NGC 1333 IRAS 2A. The column density ratios of four O-COMs with respect to CH_3OH show both similarities and differences between gas and ice. Although the straightforward explanations would be the direct inheritance from ice to gas and the gas-phase reprocessing, respectively, other possibilities such as different spatial distributions of molecules cannot be excluded.

4.1 Introduction

A longstanding question is how the chemistry in the Universe evolves from simple atoms in the diffuse interstellar or intergalactic medium to a prosperous biosphere on our Earth. The early phase of star formation, from molecular clouds to protostars, is probably the first key stage in this long journey. In particular, complex organic molecules (COMs, typically defined as carbon-bearing molecules with at least six atoms; Herbst & van Dishoeck 2009) have gained in popularity over the past several decades due to their importance in linking simple species with prebiotic molecules (Jørgensen et al. 2020; Ceccarelli et al. 2023).

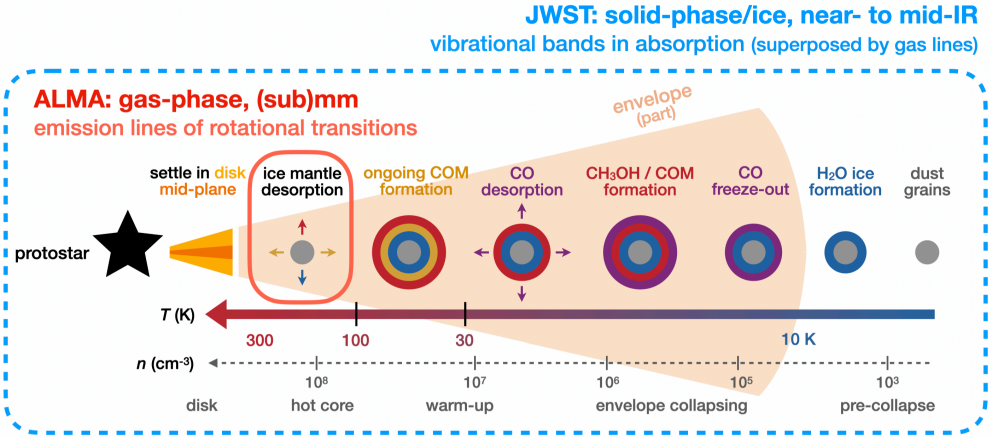


Figure 4.1: Schematic of chemical evolution on dust grains in protostellar stages, modified from Fig. 1.4 of van 't Hoff (2019) and Fig. 14 of Herbst & van Dishoeck (2009). Ice layers are dominated by different species (e.g., H_2O, CO, and COMs) that are denoted in different colors. The typical temperature and density in different evolutionary stages are labeled on arrows in the bottom. The small red box indicates the hot core region where the temperature is high enough (>100 K) to sublimate most of the volatile materials, and therefore the gas-phase molecules can be traced by radio telescopes at (sub)millimeter wavelengths. The big blue box indicates that IR telescopes such as JWST are tracing everything along the line of sight, including ice in the vast envelope and gas in the hot core.

The formation and evolution history of COMs in star-forming regions is still a topic of active study. It has been suggested that COMs are abundantly formed during cold ($T \lesssim 10$ K) early stages of star formation in the ice mantles of dust grains, and indirect evidence from gas-phase observations shows that the abundance ratios of commonly detected COMs (with respect to methanol, CH_3OH, the simplest and most abundant COM) remain consistent among sources with different luminosities (e.g., Coletta et al. 2020; van Gelder et al. 2020; Nazari et al. 2022a; Chapter 2). Figure 4.1 shows a schematic of the proposed formation history of COMs. Prior to star formation, the growth of ice mantles starts with the formation of H_2O ice and the subsequent freeze-out of CO gas as temperature and ultraviolet (UV) radiation intensity decrease toward the center of dense molecular clouds. In the icy mantles, CO is gradually hydrogenated to form formaldehyde (H_2CO), CH_3OH, and even larger COMs that contain two carbon atoms (e.g., Watanabe & Kouchi 2002; Fuchs et al.

2009; Cuppen et al. 2009; Fedoseev et al. 2015, 2022; Simons et al. 2020). As the central protostars gradually warm up their envelopes, volatile molecules such as CO in the ice mantles start to sublimate into the gas phase. In the meantime, COMs, especially those larger than CH_3OH, can keep forming via solid-phase chemistry (e.g., Garrod et al. 2022). The COMs have high binding energies and sublimate at high temperatures ($\lesssim$ 100 K; Fedoseev et al. 2015, 2022) toward the end of the warm-up stage. In hot cores in which $T > 100$ K, all the volatile ice mantles are expected to sublimate into the gas phase, and the chemistry also evolves fully in the gas phase. However, it is not certain how important the gas-phase chemistry is to the observed chemical composition of hot cores (Balucani et al. 2015).

The COMs in different phases are observed by different facilities at different wavelengths. Gas-phase COMs are usually observed at (sub)millimeter wavelengths via their rotational transitions in emission using radio telescopes such as the Atacama Large Millimeter/submillimeter Array (ALMA). With its powerful sensitivity and resolution (both spatial and spectral), ALMA has detected a rich inventory of COMs in various star-forming regions (e.g., Bacmann et al. 2012; Jiménez-Serra et al. 2016a), from starless cores to circumstellar disks (e.g., Brunken et al. 2022; Yamato et al. 2024; Booth et al. 2024b), but most commonly in protostars (e.g., Jørgensen et al. 2020; van Gelder et al. 2020; Qin et al. 2022; Nazari et al. 2022a; Baek et al. 2022; Chapter 2). These hot ($\gtrsim$ 100 K) inner regions around protostars are often referred to as hot cores, or hot corinos specifically for low-mass sources.

On the other hand, solid-phase COMs, also known as COM ices, are observed at near- and mid-infrared wavelengths via their vibrational bands in absorption. These infrared (IR) observations are better being conducted in space to avoid absorption by species such as H_2O, CO_2, and CH_4 in the Earth's atmosphere. Due to the limited sensitivity and spectral resolution of the previous Infrared Space Observatory (ISO) and *Spitzer* Space Telescope, detections of COM ices had only been confirmed for CH_3OH (the simplest and most abundant COM) before the era of the *James Webb* Space Telescope (JWST). Tentative identifications were made for two bands at 7.24 and 7.41 µm with ISO and *Spitzer*, for which the possible contributors are $HCOO^-$, C_2H_5OH, and CH_3CHO (Schutte et al. 1999; Öberg et al. 2011). With the unprecedented power of JWST, it is now promising that more COMs other than CH_3OH will be detected using the Medium Resolution Spectroscopy (MRS) mode of the Mid-InfraRed Instrument (MIRI). The wavelength range covered by MIRI-MRS (4.9–27.9 µm) contains the methanol band at 9.74 µm and a characteristic fingerprint range around 6.8–8.8 µm for COM ices, where multiple absorption bands of their vibrational modes are located (Boogert et al. 2015).

Recently, the robust detection of COMs other than CH_3OH, including acetaldehyde (CH_3CHO), ethanol (C_2H_5OH), and methyl formate (CH_3OCHO), has been reported in the MIRI-MRS spectra of two protostars, NGC 1333 IRAS 2A and IRAS 23385+6053 (hereafter IRAS 2A and IRAS 23385; Rocha et al. 2024). Now, for the first time, we are able to directly compare the gas- and solid-phase abundances of these complex organics using a combination of ALMA and JWST. In fact, some gas-to-ice comparisons have been made in previous studies (e.g., Noble et al. 2017; Perotti et al. 2020, 2021, 2023), but they were targeting simpler molecules (CO and H_2O, along with CH_3OH) in the cold envelopes in star-forming regions where the gas-phase material is nonthermally desorbed from dust grains. In this work, we are tracing the hot

core regions where most COMs have sublimated into the gas phase. As is shown in Fig. 4.1, ALMA traces the hot core region where most of the volatile species have been thermally sublimated into the gas phase. On the other hand, JWST looks through the vast envelope and traces ice mantles along the line of sight, providing so-far the best mid-IR spectra (i.e., with the highest sensitivity and spectral resolution) for COM ices. Using high-quality spectra from ALMA and JWST, we can trace the main reservoir of COMs in both phases and make a direct comparison of their abundances between gas and ice. This will provide us with valuable observational evidence of how these molecules evolve from the earlier solid-phase stage to the subsequent gas-phase stage.

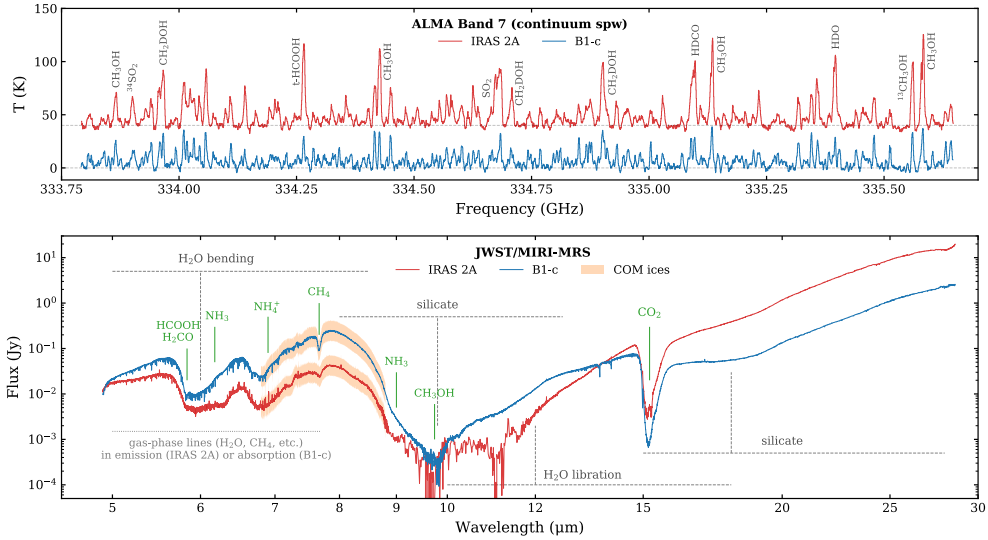


Figure 4.2: Overview of the ALMA spectra (the continuum spectral window from 333.78 to 335.65 GHz; top panel) and JWST/MIRI-MRS spectra (bottom panel) of NGC 1333 IRAS 2A and B1-c. The JWST spectrum of IRAS 2A between 8.85 and 11.7 μm was binned by a factor of ten for better visualization. In the bottom panel, the fingerprint ranges of COM ices between 6.8 and 8.8 μm are highlight in orange. Important emission lines in the ALMA spectra and absorption bands in the JWST spectra are labeled in each panel.

Ideally, the comparison should be made for the same sources that are rich in both gas and ice COMs. However, the sample of protostars for which gas-phase COMs have been detected using ALMA and other telescopes (e.g., Jørgensen et al. 2016; Mininni et al. 2020; Yang et al. 2021; Hsu et al. 2022; Nazari et al. 2022a; Chapter 2) is much larger than that for ice detections with IR telescopes. In this paper, we select two low-mass protostars, IRAS 2A and B1-c, for a case study as part of our JWST Observations of Young protoStars (JOYS+) program. These two sources are famous hot corinos that have been observed to have rich emission lines of gas-phase COMs (e.g., Taquet et al. 2015; van Gelder et al. 2020). They also have noticeable absorption features in the fingerprint range of COM ices between 6.8 and 8.8 μm, as has been revealed by previous *Spitzer* observations (Boogert et al. 2008) and our JWST/MIRI-MRS spectra.

IRAS 2A and B1-c also have a relatively strong CH_3OH ice band at 9.74 μm,

which is often severely extincted by silicate features but necessary to determine the ice column density of CH_3OH as a reference species. A higher signal-to-noise (S/N) of this band will lead to better constraints on the column density of CH_3OH ice, and hence a more accurate estimation of the COM ratios with respect to CH_3OH. Overall, IRAS 2A and B1-c are among the best candidates for the first direct comparison of COM abundances between gas and ice.

We emphasize that this observational study cannot be realized without the support of experimentalists. Over the past few years, many laboratory IR spectra (hereafter lab spectra) of COM ices have been measured under different conditions (e.g., Terwisscha van Scheltinga et al. 2018, 2021; Hudson et al. 2018; Hudson & Ferrante 2020; Hudson et al. 2021; Hudson & Yarnall 2022; Rachid et al. 2020, 2021, 2022; Slavicinska et al. 2023). Most of these lab spectra are publicly available on the Leiden Ice Database for Astrochemistry (LIDA; Rocha et al. 2022). The references for other lab spectra that are involved in this paper can be found at Table C.1 in Rocha et al. (2024).

The paper is structured as follows. Sect. 4.2 provides information on the ALMA and JWST observations and the data reduction procedures. Sect. 4.3 describes the analyzing methods of the ALMA and JWST data, and the results are displayed in Sect. 4.4. A discussion is elaborated on in Sect. 4.5, with a focus on the gas-to-ice ratios of four selected oxygen-bearing COMs. Conclusions are summarized in Sect. 4.6. A considerable amount of content is reserved in the appendices, including (i) the images of ALMA and JWST observations (Appendix 4.A); (ii) details and supplements to the fitting methodology of the JWST spectrum of B1-c (Appendices 4.B–4.D and 4.J); (iii) additional figures and table for the fitting results of ALMA spectra (Appendices 4.E and 4.L); and (iv) additional figures and table that are relevant to the fitting results of the JWST spectrum of B1-c (Appendices 4.G, 4.I, and 4.K).

4.2 Observations and data reduction

4.2.1 ALMA

The ALMA data of IRAS 2A and B1-c were taken as part of program 2021.1.01578.S (PI: B. Tabone), which targeted magnetic disk winds of five Class 0 protostars in Perseus and also observed a large number of emission lines of gas-phase COMs. The observations were taken with a combination of an extended configuration, C-6 ($\theta_{\rm beam}$ = $0.12'' \times 0.09''$), and a more compact configuration, C-3 ($\theta_{\rm beam} = 0.58'' \times 0.34''$). The maximum recoverable scales ($\theta_{\rm MRS}$) for the C-6 and C-3 datasets are $1.6''$ and $6.2''$, respectively. The data cover nine spectral windows in Band 7 between 333.8 and 347.6 GHz, including one 1.875-GHz wide continuum window with a spectral resolution of 0.87 km s^{-1}. Among the eight line windows, six windows have a spectral resolution of 0.22 km s^{-1} and a frequency coverage of 0.12 GHz or 0.24 GHz; the remaining two windows are wider (0.48 GHz) but have a lower spectral resolution of 0.44 km s^{-1}.

The data were pipeline-calibrated using CASA versions 6.2.1.7 and 6.4.1.12 (McMullin et al. 2007). The measurement sets from the C-3 and C-6 configurations were combined via concatenation and subsequently continuum subtracted and imaged using the `concat`, `uvcontsub`, and `tclean` tasks in CASA version 6.4.1.12. For B1-c, a `briggs` weighting of 0.5 was used for all the spectral windows to improve the angular resolution, yielding a synthesized beam of $\theta_{\rm beam} = 0.08'' \times 0.11''$ for the 0.88 mm

continuum and the spectral windows, which corresponds to a spatial resolution of ∼30 au at ∼320 pc (i.e., the distance of the Perseus star-forming region). For IRAS 2A, a higher `briggs` weighting of 2.0 was used to increase the S/N of the $H^{13}CO^+$ line at 348.998 GHz (customized for Nazari et al. 2024c), resulting in a larger beam size of $0.25'' \times 0.4''$ for that spectral window (336.93–337.05 GHz). The root mean square (rms) is about 1.5–2.0 mJy beam^{-1} for all the spectral windows except that the one with a larger beam has a higher rms of 5.0 mJy beam^{-1}. The rms for the continuum is about 0.2 mJy beam^{-1}.

4.2.2 JWST

The JWST/MIRI-MRS data of IRAS 2A and B1-c were taken in the Guaranteed Time Observation (GTO) programs 1236 (PI: M. E. Ressler) and 1290 (PI: E. F. van Dishoeck), respectively, as part of the JWST Observations of Young protoStars+ (JOYS+) collaboration.[2] MIRI-MRS covers 4.9–27.9 µm with a spectral resolution of $R = \lambda/\Delta\lambda$ of 1300-3700 (Rieke et al. 2015; Wright et al. 2015; Wells et al. 2015; Labiano et al. 2021; Wright et al. 2023; Argyriou et al. 2023). The dither patterns were optimized for extended sources, with the two-point pattern used for IRAS 2A and the four-point pattern for B1-c. Two pointings were observed for B1-c; one centered on the protostar itself, and one covering the blueshifted outflow. Both programs included dedicated background observations in a two-point dither pattern that allow for a subtraction of the telescope background and detector artifacts. All observations were taken using the FASTR1 readout mode and cover the full 4.9-27.9 µm wavelength coverage of MIRI-MRS. The total integration time of 333 s for IRAS 2A was equally divided over the three MIRI-MRS gratings. For B1-c, the total integration time was 8000 s, of which 4000 s were used in grating B (MEDIUM) to get a good S/N in the deep silicate absorption band, and 2000 s each for gratings A (SHORT) and C (LONG).

The data were processed through all three stages of the JWST calibration pipeline version 1.12.5 (Bushouse et al. 2023), using the procedure described by van Gelder et al. (2024b). The reference contexts `jwst_1126.pmap` and `jwst_1177.pmap` of the JWST Calibration Reference Data System (CRDS; Greenfield & Miller 2016) were used for IRAS 2A and B1-c, respectively. The raw data were processed through the `Detector1Pipeline` using the default settings. The dedicated backgrounds were subtracted at the detector level in the `Spec2Pipeline`, and the fringe flat was applied for extended sources (Crouzet et al. 2025) along with a residual fringe correction (Kavanagh et al. in prep.). An additional bad pixel map was applied to the resulting calibrated detector files using the Vortex Imaging Processing package (Christiaens et al. 2023). The final datacubes were constructed using the `Spec3Pipeline` for each band of each channel separately.

[2] `https://miri.strw.leidenuniv.nl`

4.3 Methods

4.3.1 ALMA

The ALMA spectrum of IRAS 2A was extracted from the peak pixel of CH_3OH emission at R.A. $03^h28^m55.569^s$, Dec. $+31^d14^m36.930^s$ (J2000). This position was selected based on the integrated intensity maps of a dozen of O-COM lines shown in Fig. 4.A.1, which is slightly offset from the continuum peak at R.A. $03^h28^m55.5735^s$, Dec. $+31^d14^m36.925^s$ (6 pixels offset in R.A. and 2 pixels in Dec; the pixel size is $0.01'' \times 0.01''$). The line emission of IRAS 2A near the continuum peak is attenuated, probably due to the high optical depth of dust in Band 7 (e.g., De Simone et al. 2020). B1-c does not show the same attenuation issue, and therefore its spectrum was extracted from the peak pixel of the continuum emission at R.A. $03^h33^m17.881^s$, Dec. $+31^d09^m31.740^s$, which is consistent with the emission peak of O-COMs (Fig. 4.A.2).

The ALMA spectra of the continuum window (333.78–335.65 GHz) of IRAS 2A and B1-c are shown in the top panel of Fig. 4.2. The spectra were converted to brightness temperature (K) scale by averaging over the synthesized beam. Line identification and spectral fitting were performed using the spectral analysis software CASSIS[3](Vastel et al. 2015). We mainly considered O-COMs in the fitting, including methanol and its isotopologs (CH_3OH, $^{13}CH_3OH$, and $CH_3^{18}OH$), acetaldehyde (CH_3CHO), ethanol (C_2H_5OH), dimethyl ether (CH_3OCH_3), methyl formate (CH_3OCHO), glycolaldehyde (CH_2OHCHO), and ethylene glycol ($(CH_2OH)_2$). The local thermal equilibrium (LTE) modeled spectra were fit to the observations, and a best-fit column density (N), excitation temperature (T_{ex}), and line width (full width at half maximum; FWHM) were determined for each species in each source by grid fitting or visual inspection.

We adopted the grid-fitting method when a species has more than five clean lines detected (clean means the line is unblended or partially blended and the line profile is still recognizable). For each source and each species, a grid of N, T_{ex}, and FWHM was preset, and each grid point corresponds to an LTE model. The χ^2 between the observations and the LTE model was only calculated around the fully unblended lines. The model with the smallest χ^2 gave the best-fit parameters. The 2σ uncertainties on N and T_{ex} were estimated from the N–T_{ex} contour plots. The uncertainties of N are usually 30–40% of the best-fit values, and the uncertainties of T_{ex} varies with species.

The grid-fitting method was mainly applied to B1-c. However, most of the emission lines in the IRAS 2A spectrum show double-peaked features, suggesting two velocity components. This makes the blending issue more severe, and hence makes it trickier to perform grid fittings. Instead, we fit the spectrum through visual inspection, which is more flexible in this case. For each species and each velocity component, we first determined the v_{lsr} and FWHM based on several strong and unblended lines, then manually adjusted N and T_{ex} to achieve a better fit between the LTE models and the observed spectrum. All the parameter were fine-tuned once a rough range around the best fit was found. In this way, the uncertainties on N and T_{ex} could not be calculated from contour plots, but were estimated at a level of 20–30%. More details about the fitting strategy for line-rich ALMA spectra can be found in Sect. 2.3.1.

[3] http://cassis.irap.omp.eu/

4.3.2 JWST

The JWST/MIRI-MRS spectrum of IRAS 2A has been presented and analyzed in Rocha et al. (2024). Here, we implemented a similar analysis to B1-c. The B1-c spectrum was manually extracted from the continuum peaks at R.A. $03^{h}33^{m}17.8959^{s}$ $+31^{d}09^{m}31.8578^{s}$, which is only $\sim$0.1″ offset from the ALMA continuum peak at 0.9 mm (see Fig. 4.A.3). The diameter was set to four times the size of the point spread function (PSF) of MIRI-MRS ($\mathrm{FWHM_{PSF}} = 0.033'' \times (\lambda/\mu\mathrm{m}) + 0.106''$; Law et al. 2023); that is, the extraction aperture increases toward longer-wavelength channels. In our COM fingerprint range of interest, the aperture size is $\sim$1.4″ in diameter. An additional 1D residual fringe correction was performed especially to remove the high-frequency dichroic noise in channels 3 and 4 (Kavanagh et al. in prep.). Since the photometric calibration between the bands was accurate enough, the 12 sub-bands were stitched together without any flux adjustment being applied.

Starting from the original extracted spectrum, we required five steps to reach the final goal of determining the column densities of COM ices: fit a global continuum (Sect. 4.3.2.1), subtract the silicate features (Sect. 4.3.2.2), fit a local continuum (Sect. 4.3.2.3), remove the superposed gas-phase lines (which is needed for B1-c but might be skipped for other sources if these lines are weak; Sect. 4.3.2.4), and decompose the COM fingerprint ranges between 6.8 and 8.8 µm using the lab spectra (Sect. 4.3.2.5). To calculate the ice column density ratios of COMs with respect to a reference species, we also determined the ice column densities of CH_3OH and H_2O with a slightly different routine than that for the 6.8–8.8 µm range (see Sect. 4.3.2.6).

4.3.2.1 Global continuum

The first step was to fit a global continuum level so that the original spectrum in flux scale (λF_λ) could be converted into optical depth (τ_λ) scale by taking the logarithm between the global continuum and the observed spectrum. The goal of this step is not to derive the origin of the mid-IR continuum emission, but to enable the subsequent comparison with the lab spectra that are measured in absorbance ($Abs_{\tilde{\nu}}$), which is directly linked to optical depth with a scaling factor of 2.303 (i.e., $\ln 10$):

$$\tau_{\tilde{\nu}} = 2.303 \times Abs_{\tilde{\nu}}. \tag{4.1}$$

We fit a third-order polynomial to the observed spectrum in three ranges: 4.9–5.5 µm at the short-wavelength end, 7.75–7.9 µm around the "bump" between the NH_4^+ band at 6.8 µm and the silicate band at 9.8 µm, and 27.6–27.9 µm at the long-wavelength end (see Fig. 4.3a). The flux values in the last two ranges were manually lifted by a factor of 1.3 to leave some space for the red wings of the H_2O bending mode at $\sim$6 µm and the silicate band at $\sim$18 µm. Both the fit ranges and the lifting factor of 1.3 were tuned carefully to produce a plausible global continuum. The last range of 27.6–27.9 µm was set to be narrow, since using a wider range will create an increasing feature in the polynomial longward of 20 µm, which will introduce some artifacts in the optical depth spectrum that do not match with the supposed silicate features in the next step.

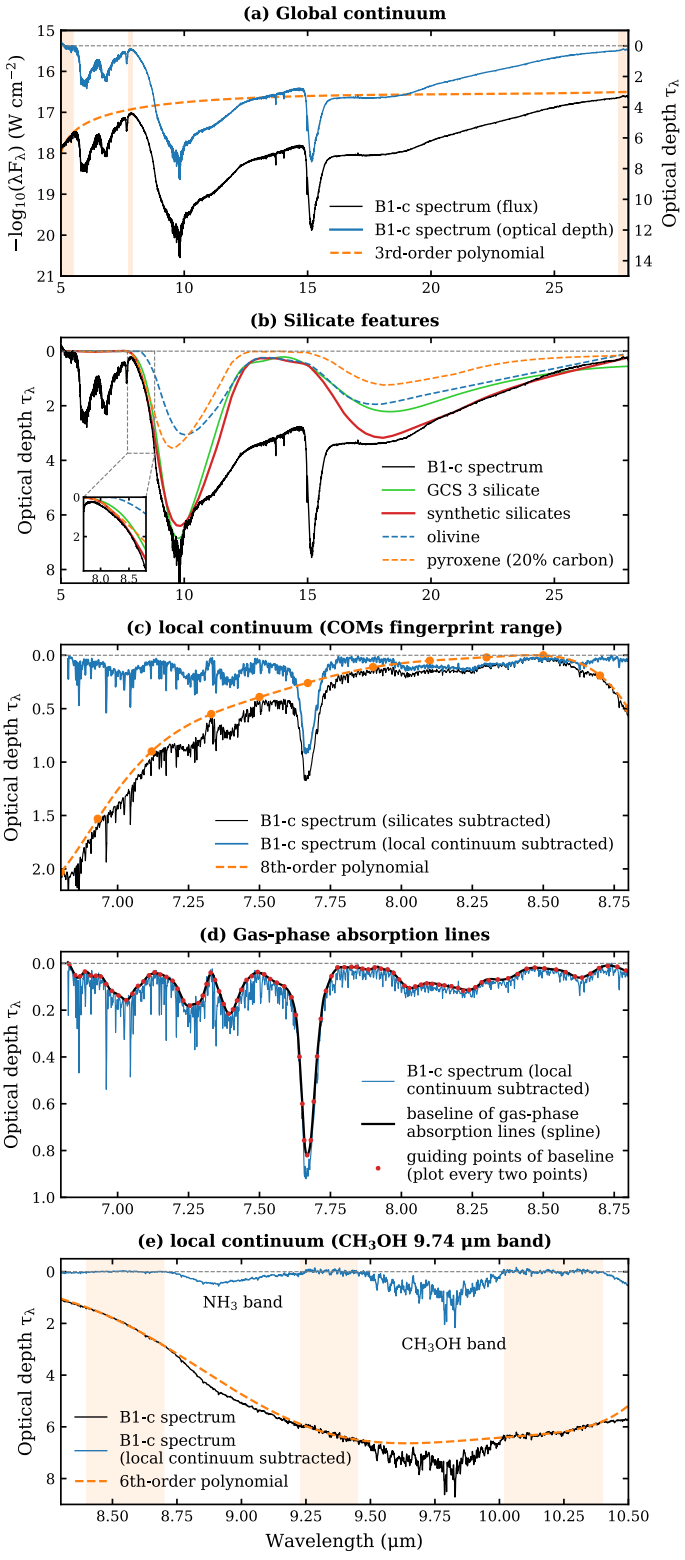


Figure 4.3: Panels (a)–(d) show the four steps to isolate the fingerprint range of COM ices between 6.8 and 8.8 µm from the original JWST/MIRI-MRS spectrum of B1-c: (a) fit a global continuum and convert the spectrum to optical depth scale; (b) subtract silicate features at ∼9.8 and 18 µm; (c) trace a local continuum in the 6.8–8.8 µm range to isolate the weak bands of COM ices from other strong features; and (d) trace a baseline of the gas-phase lines in absorption to restore the band profiles of ices. Panel (e) shows the isolation of the CH_3OH ice band at 9.74 µm by tracing a local continuum. The orange-shaded regions in Panels (a) and (e) show the selected wavelength ranges for the polynomial fitting.

4.3.2.2 Silicate bands

The second step is to fit and subtract the silicate features at 9.8 and 18 µm. This step removes the contribution from silicates to the COM bands in 8.0–8.8 µm, the CH_3OH band at 9.74 µm, and the water libration band at ∼13 µm as much as possible. The silicate features may not be fully removed, but the residuals will be further excluded in the next step of local continuum subtraction. We tried fitting with two silicate profiles: the observed spectrum toward the galactic center source GCS 3 (Kemper et al. 2004), and the synthetic silicate spectra of pyroxene ($Mg_xFe_{1-x}SiO_3$) and olivine ($MgFeSiO_4$) computed by the `optool` code (Dominik et al. 2021). The GCS 3 and the synthetic silicate spectra were scaled manually to match the optical depth spectrum of B1-c. The fitting criterion is to fit the two silicate bands at 9.8 and 18 µm as well as possible without overfitting any part of the observed spectrum, especially the blue wing of the 9.8 µm band.

Figure 4.3b shows the fitting results of the silicate features. The GCS 3 profile can be scaled deeper at the 9.8 µm band than the `optool` mixture of pyroxene and olivine, but leaves non-negligible residuals in the blue wing of the 9.8 µm band (roughly between 8 and 9 µm). This has been seen in previous ice studies (e.g., Boogert et al. 2008) and also our JOYS+ data. The reasons why are still being investigated. The scaling factors of the synthetic silicates, pyroxene and olivine, were tuned manually to achieve the best fit. Pyroxene has a more blueshifted 9.8 µm band than olivine, and its absorbance ratio between the 9.8 and 18 µm bands is higher. The contribution of pyroxene is constrained by the small bump at ∼8.5 µm in the observed spectrum (see the blow-up in Fig. 4.3b), which should not be overfit by the blue wing of the synthetic silicate spectrum. Similarly, olivine has a relatively strong 18 µm band and its scaling factor is constrained by not overfitting the 18 µm band. The synthetic silicate spectrum turned out to fit the B1-c spectrum better than the GCS 3 profile. This, however, may not be the case in general, since the observed silicate bands in independent sources can be quite different from one another. The silicate features in the mid-IR spectra of protostellar sources are worth a thorough investigation that is beyond the scope of this paper. We provide some additional details and relevant discussion in Appendix 4.B.

4.3.2.3 Local continuum

After subtracting the silicate component from the optical depth spectrum, a local continuum was fit between 6.8 and 8.8 µm to isolate the absorption bands of COM ices from other strong features (e.g., the NH_4^+ band at 6.8 µm, the red wing of the H_2O bending mode at 6 µm, and the leftover silicate features in the previous step; Schutte et al. 1999). Due to the richness of absorption features in this range, continuous absorption-free regions for fitting hardly exist. Instead, a sequence of "guiding points" are set manually at certain locations (e.g., Grant et al. 2023). We fit a seventh-order polynomial to about ten guiding points at positions that were set between two absorption bands or in the middle of a broad band to bridge the guiding points on both sides (see Fig. 4.3c). Polynomials with lower orders would slightly deviate from some of the guiding points or create artificial features.

Similar to fitting a local continuum, the choice of local continuum is somewhat subjective. Rocha et al. (2024) show in their Section 4.2.4 and appendix J that the

ice column densities of some species can be changed by using a different local continuum. In our case, we traced the local continuum as close to the observed spectrum as possible, so that the ice abundances of the targeted COMs are not likely to be overestimated.

The noise level in the optical depth around the COM fingerprint range was also estimated in this step. A second-order polynomial was fit to a small range between 8.227 and 8.240 µm that is relatively free of emission or absorption features. The noise level was calculated using

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \bar{y})^2}{N}}, \tag{4.2}$$

where y_i is the polynomial-subtracted optical depth at each channel between 8.227 and 8.240 µm, and $\bar{y}$ is the mean of y_i. The noise level of the B1-c spectrum was estimated as 2.3×10^{-3} in units of optical depth.

4.3.2.4 Gas-phase lines

There are plenty of gas-phase absorption lines superposed on the broad ice bands that need to be accounted for after subtracting the local continuum. An ideal but nontrivial solution would be simultaneously tracing a baseline and fitting LTE models to retrieve the column density and temperature of the gas-phase components. However, deriving the gas properties is not the focus of this paper, so we only considered tracing a baseline of these gas lines to isolate the ice bands that we are interested in. Similar to the local continuum, the baseline was determined by fitting a spline function to a series of guiding points that were set manually (Fig. 4.3d). The smoothing factor of the spline function was also carefully tuned.

4.3.2.5 Absorption features of complex organic molecule ices

The final step is to decompose the spectrum between 6.8 and 8.8 µm by fitting with lab spectra of different ices. However, this is not as straightforward as fitting the ALMA spectra where the LTE models can be analytically computed. When fitting the JWST spectra, the variables are not only the temperature and the ice column density of a species (or equivalently, the scaling factor of the corresponding lab spectrum), but also the mixing conditions of the ices. In the solid phase, a species can be mixed with various constituents in different ratios under different temperatures (e.g., CH_3OH mixed with H_2O in an abundance ratio of 1:10 at 15 K). The relation between the band profile and the mixing condition —that is, mixing constituents, mixing ratio, and temperature —is not linear. Therefore, we need to first select the mixing condition that matches best the observations for each species (which is introduced in this subsection), and then do a least-squares fitting with the selected lab spectra of the candidate species to get a best fit on the ice column densities, or the scaling factors (which will be introduced in the next subsection).

These two steps can be executed either simultaneously or separately. The `ENIIGMA` code implemented in Rocha et al. (2024) is the first case, which calculates the chi-square of all the possible combinations of lab spectra and searches for the global minimum using genetic algorithms. However, `ENIIGMA` is limited in the stability of fitting results due to the intrinsic randomness of genetic algorithms. In this work,

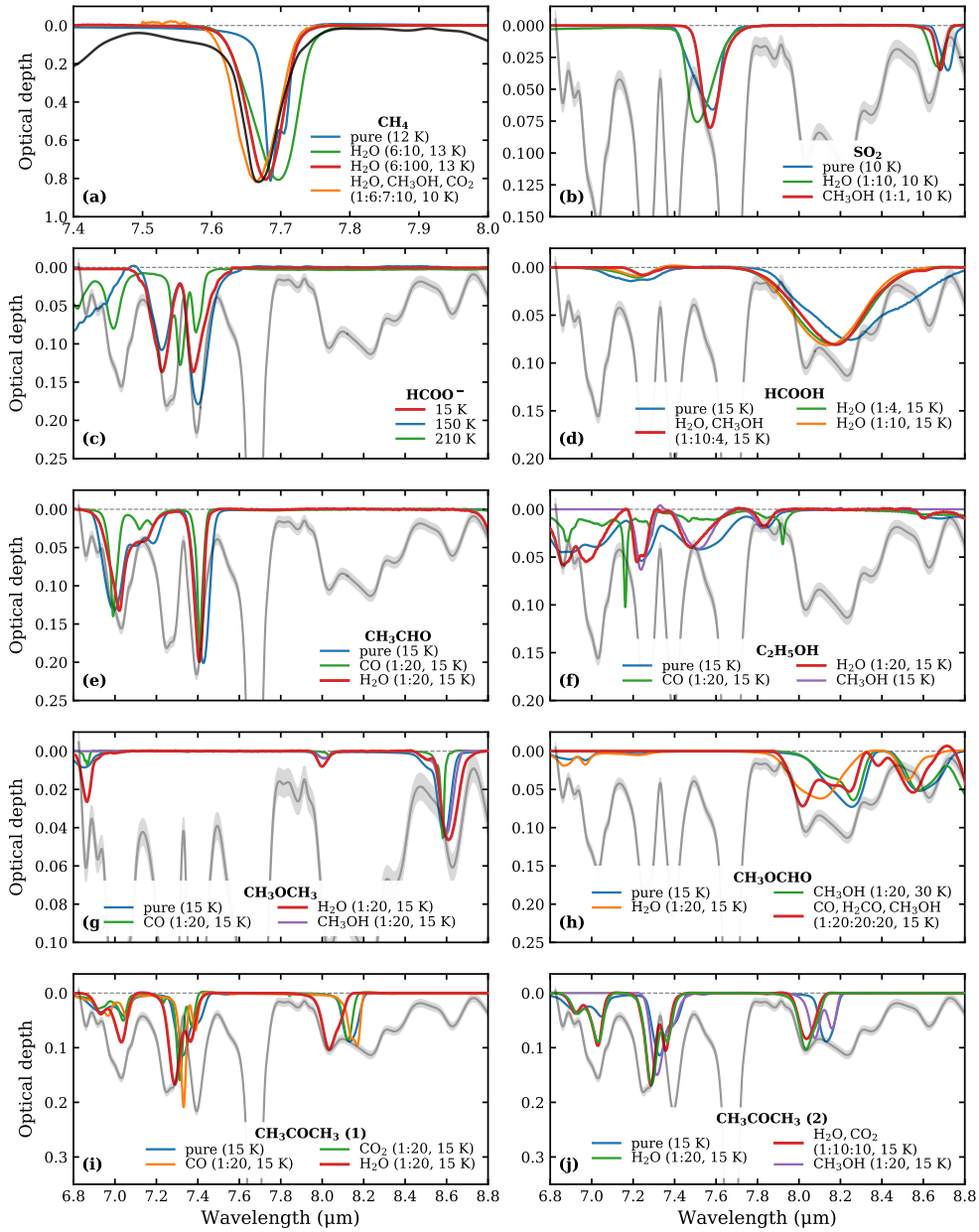


Figure 4.4: Comparison between the JWST/MIRI spectrum of B1-c (gray) and the lab spectra (colors) in the COM fingerprint range of 6.8–8.8 μm. Each panel focuses on one species and shows the comparison between the observed B1-c spectrum and the lab spectra with different mixing constituents, except for Panel (c), which shows the lab spectra of $HCOO^-$ under different temperatures. The observed spectrum along with the 3σ level is shown in light gray, except for Panel (a), which blows up the CH_4 band at ∼7.7 μm, and the observed spectrum is plotted in black for clarity. In each panel, the spectrum in blue corresponds to the pure ice, and the best-fit spectrum to the observations is highlighted with a thicker red line. A similar comparison but for lab spectra under different temperatures is shown in Figs. 4.5 and 4.C.1.

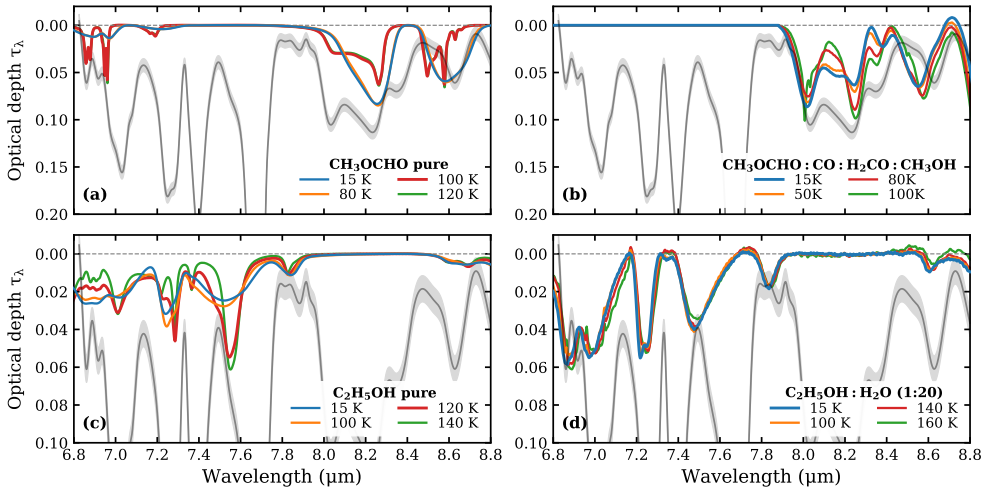


Figure 4.5: Same as Fig. 4.4 but for a comparison between different temperatures. Here, we use two species, CH_3OCHO and C_2H_5OH, as examples; the remaining species (CH_3CHO, CH_3OCH_3, and CH_3COCH_3) are shown in Fig. 4.C.1. The left and right columns show the lab spectra of pure ices and ice mixtures, respectively. In the pure-ice panels, the spectra with crystalline features are highlighted in thicker red lines. The corresponding temperatures indicate the upper limit of crystallines. In the mixed-ice panels, the spectra with the lowest temperature (15 K) are highlighted in thicker blue lines; they are also the spectra used in the overall fitting.

we performed the two steps of selecting the lab spectra and fitting the scaling factors separately, which can also be used to crosscheck the results reported in Rocha et al. (2024).

Selection of lab spectra. In the first step of this process, we selected a list of species that are likely to contribute to the absorption features between 6.8 and 8.8 μm. These species can be both COMs and simple molecules, and the promising candidates have been explored in Rocha et al. (2024). In this work, we considered 12 molecules in total, including five simple species (CH_4, SO_2, OCN^-, $HCOO^-$, and HCOOH) and seven COMs (CH_3CHO, C_2H_5OH, CH_3OCH_3, CH_3OCHO, CH_3COOH, CH_3COCH_3, and CH_3CN). Each species has a collection of lab spectra measured under different mixing conditions of ices (i.e., mixing constituent, mixing ratio, and temperature). The next step is to determine under which mixing condition the lab spectrum matches the observations best. We first did this by superposing all the lab spectra of a certain species on the JWST spectrum to see which one matches the observations best, as is indicated by Figs. 4.4–4.5 and 4.C.1, in which each panel shows the comparison between different mixing constituents and temperatures for one species, respectively. For most candidate species, this is already very informative and efficient.

However, if the differences among spectra are small and direct comparison is not straightforward, an alternative way is to compare the peak positions and FWHMs of the characteristic absorption bands of a certain species, when these data are available for the lab spectra (e.g., Boogert et al. 1997; Terwisscha van Scheltinga et al. 2018,

2021; Rachid et al. 2020, 2022). The peak positions and FWHMs of observed bands are derived by fitting Gaussian functions to the observed spectrum. Taking CH_3CHO as an example, Fig. 4.6 shows the comparison between the lab spectra and the observed B1-c spectrum for two CH_3CHO bands at ∼7.0 (CH_3 deformation mode) and ∼7.42 µm (CH_3 s-deformation and CH wagging modes). It is already straightforward to tell from the left panel that the CH_3CHO:H_2O mixture fits the observations best. The two panels on the right show that the observed spectrum has both bands obviously wider than the lab spectra, suggesting that these two bands are not only attributable to CH_3CHO, but also have contribution from other species. The peak positions of both bands match more closely the H_2O mixture than they do the pure ice and the CO mixture, which further supports that the CH_3CHO:H_2O spectrum is the most suitable one to use in the overall fitting of scaling factors. As for temperature, lower temperatures are favored in the comparison. We chose the spectrum at 15 K for the overall fitting, but 30 K can also be used alternatively, since the difference in band profiles is small under 70 K. The selection of ice mixtures of other species follows the same procedure.

For COMs, there is usually only one mixing ratio that has been measured in laboratories for a specific mixture (e.g., COM:H_2O = 1:20), so we can only select among different mixing constituents and temperatures. For some simple species such as CH_4, there are more choices for mixing ratios, but within a limited temperature range. Despite the lack of additional laboratory measurements, in Sect. 4.4.2 we shall show that there is already a lot we can do with the current measurements.

The lab spectra used in comparison had been baseline corrected to isolate the absorption bands of the targeted species from the strong features of the mixing constituents (e.g., H_2O and CH_3OH). This process is important but sometimes not trivial, especially for the ice mixtures with CH_3OH. The details are elaborated in Appendix 4.D.

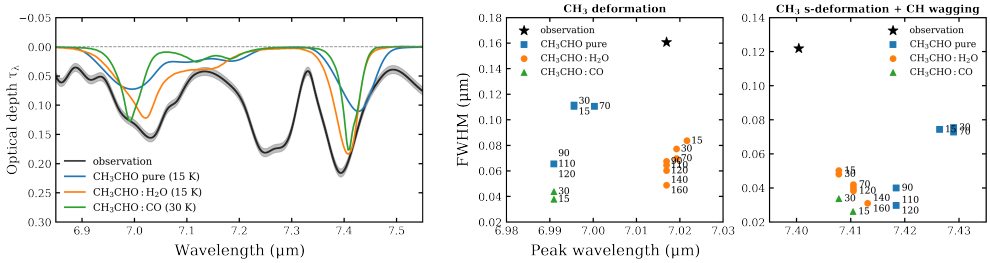


Figure 4.6: Two methods of selecting the best-fit ice mixture, taking CH_3CHO as an example. The left panel shows the direct comparison of spectral profiles between observations and experiments. The middle and the right panels show the comparison of peak position and FWHM of absorption bands between lab spectra and observations.

Overall fitting of the complex organic molecule fingerprint range 6.8–8.8 µm. In the second step, we performed an overall fitting on the scaling factors of all the selected lab spectra (each spectrum corresponding to the best-fit ice mixture of

one species), as is described in

$$\tau_{\rm obs} = \sum_{i}^{N} a_i \, \tau_{i,\rm lab}, \tag{4.3}$$

where $\tau_{\rm obs}$ is the observed spectrum between 6.8 and 8.8 µm, and a_i is the scaling factor of the i-th component; that is, the selected best-fit lab spectrum of the i-th species. N is the total number of the spectra (or species) that are considered in the fitting. We first performed a normal least-squares fitting to the observations using the *Omnifit* code (Suutarinen 2015) with all the candidate species taken into account, in order to get a general idea of the best-fit ranges. Then we selected a subgroup of significant candidates for a subsequent Markov chain Monte Carlo (MCMC; Foreman-Mackey et al. 2013) fitting to better constrain the uncertainties on their scaling factors. To reduce the computation time, we set boundaries on the scaling factors to make sure that the fitting results fall in a reasonable range. The lower limits and the initial values were set as zero, and the upper limits were determined by manually comparing the lab spectra with observations (as is done in Figs. 4.4 and 4.C.2).

4.3.2.6 Ice bands of CH_3OH and H_2O

As the simplest and the most abundant COM, CH_3OH is usually taken as the reference species to calculate the relative abundances (i.e., ratios) of other COMs. These abundance ratios are generally compared among different sources to see if there are similarities or differences. The abundance of CH_3OH ice was derived separately since its strongest mid-IR band is located at 9.74 µm, outside of the 6.8–8.8 µm range. The CH_3OH band at 9.74 µm can be isolated by fitting a local continuum on either the original optical depth spectrum (Bottinelli et al. 2010) or the silicate subtracted spectrum. However, the S/N of this band is usually much lower than other parts of the spectrum due to the strong extinction by the silicate 9.8 µm band. Here, we fit a sixth-order polynomial to the original optical depth spectrum in three wavelength ranges: 8.40–8.80, 9.23–9.45, and 10.02–10.40 µm (see Fig. 4.3d). The local continuum subtracted spectrum between 9.5 and 10.0 µm was then fit by a Gaussian function. We chose to use a Gaussian function instead of the lab spectra for simplicity, since this band is isolated and the fitting results would be similar.

H_2O is also a reference species commonly used to calculate relative abundances, especially for simple molecules. We fit the H_2O band at 13 µm in the silicate subtracted spectrum using the lab spectra recently measured by Slavicinska et al. (2024). The weaker 6 µm band was taken as a secondary reference. Since this band is very broad and strong, and its spectral profile changes prominently with temperature, it is easy to select the best fitting H_2O spectrum to the observations and adjust the scaling factor accurately through visual inspection.

4.3.2.7 Ice column density

Once the best-fit scaling factors are found for the selected lab spectra (for COMs in Sect. 4.3.2.5), we can calculate the ice column densities of each species by picking a vibrational band, usually the strongest one in the interested wavelength range, and

using:

$$N_{\rm ice} = \frac{1}{A} \int_{\tilde{\nu}_1}^{\tilde{\nu}_2} a\, \tau_{\tilde{\nu},\rm lab}\, {\rm d}\tilde{\nu}, \tag{4.4}$$

where A is the strength of this band that extends from wavenumber $\tilde{\nu}_1$ to $\tilde{\nu}_2$, $\tau_{\tilde{\nu},\rm lab}$ is the lab spectrum at the wavenumber, $\tilde{\nu}$, on the optical depth scale, and a is the scaling factor of this spectrum compared to the observations. In reality, the IR spectra are measured in absorbance (Abs), and the conversion to optical depth is given by Eq. (4.1). In the cases in which we have used a Gaussian function to represent the absorption band (e.g., for the CH_3OH band in Sect. 4.3.2.6), $\tau_{\tilde{\nu},\rm lab}$ is replaced by the best-fit Gaussian function and the scaling factor, a, is not needed.

4.4 Results

4.4.1 ALMA

4.4.1.1 Emission maps

The continuum emission of IRAS 2A and B1-c appears to be roughly symmetric and round, even at the high angular resolution of $\sim$0.1″ (i.e., 32 au at a distance of 320 pc; see contours in Figs. 4.A.1–4.A.2), which is consistent with previous observations (Taquet et al. 2015; Segura-Cox et al. 2018; van Gelder et al. 2020; Yang et al. 2021). The COM emission shows consistent morphologies among different species and transitions. In B1-c, the COM emission is symmetric and follows the morphology of the continuum, suggesting that this protostellar system is fairly quiescent. Conversely, the morphologies in IRAS 2A are asymmetric, with the emission attenuated at the continuum peak and stronger at an offset position to the southwest. This asymmetry was not revealed in previous observations with lower angular resolution, and may be related to the dynamics of the protostar and the circumstellar disk (e.g., a recent outburst; Hsieh et al. 2019).

The emission maps of multiple COM lines shown in Figs. 4.A.1–4.A.2 also reflect how large the hot cores are. The size of a hot core is usually considered as the radius at which T = 100 K, at which temperature most of the volatile species have sublimated from ice mantles into the gas phase. This radius can be estimated analytically with the following equation (Bisschop et al. 2007; van't Hoff et al. 2022):

$$R_{\rm T=100\ K} \approx 15.4 \sqrt{L/L_\odot}\ {\rm au}. \tag{4.5}$$

Taking the luminosity of 91 and 5.9 $L_\odot$ for IRAS 2A and B1-c (van Gelder et al. 2022), the $R_{\rm T=100\ K}$ are estimated as 147 au and 37 au (i.e., $\sim$4 and 1 beams), respectively. The spectra of both sources were extracted within $R_{\rm T=100\ K}$; hence, the derived column densities are also expected to be representative of the gas-phase COM abundances inside the hot cores.

4.4.1.2 Spectra

In the nine spectral windows between 333.8 and 347.6 GHz in ALMA Band 7, rich molecular lines are observed in both sources. We determined column densities (N) and

Table 4.1: Column densities and excitation temperatures of gas-phase molecules derived from the ALMA Band 7 spectra.

Sources		NGC 1333 IRAS 2A†					B1-c				
Species	Catalog	$N_{\rm gas}$ ($\rm cm^{-2}$)	$T_{\rm ex}$ (K)	FWHM (km $\rm s^{-1}$)	$v_{\rm lsr}$	X/CH_3OH (%)	$N_{\rm gas}$ ($\rm cm^{-2}$)	$T_{\rm ex}$ (K)	FWHM (km $\rm s^{-1}$)	$v_{\rm lsr}$	X/CH_3OH (%)
CH_3OH	CDMS	$(1.1\pm0.3)\times10^{19}$	165±10	3.5	5.0	≡100	$(9.0\pm3.0)\times10^{18}$	175±20	3.5	6.0	≡100
		$(7.0\pm2.0)\times10^{18}$	210±10	3.0	8.5						
$^{13}CH_3OH$	CDMS	$>5.0\times10^{16}$	170±20	3.0	5.3	>0.44	$>2.5\times10^{16}$	[175]	3.3	5.8	>0.28
		$>3.0\times10^{16}$	210±20	3.0	8.3						
$CH_3^{18}OH$	CDMS	$\leq3.0\times10^{16}$	[170]	[3.0]	[5.0]	≤0.39	$\leq3.0\times10^{16}$	100–200	[3.0]	[5.8]	≤0.33
		$\leq4.0\times10^{16}$	[210]	[3.0]	[8.5]						
CH_2DOH	JPL	$(1.3\pm1.0)\times10^{17}$	180±30	3.0	4.8	1.7	$(5.0\pm1.0)\times10^{17}$	200±40	3.0	6.0	5.6
		$(1.7\pm1.0)\times10^{17}$	220±20	3.0	7.9						
CH_3CHO	JPL	$(1.0\pm0.3)\times10^{16}$	170±50	2.8	5.3	0.14	$(2.0\pm0.4)\times10^{16}$	210±50	3.3	6.0	0.22
		$(1.5\pm0.5)\times10^{16}$	220±50	2.8	8.0						
C_2H_5OH	CDMS	$(4.5\pm1.5)\times10^{16}$	220±50	3.0	5.0	0.58	$(7.0\pm2.0)\times10^{16}$	250±50	3.0	6.0	0.78
		$(6.0\pm1.5)\times10^{16}$	220±50	3.0	8.0						
CH_3OCH_3	CDMS	$\leq2.0\times10^{17}$	[200]	[3.0]	–	≤1.1	$\leq6.0\times10^{16}$	[200]	[3.0]	–	≤0.67
CH_3OCHO	JPL	$(7.0\pm2.0)\times10^{16}$	170±30	3.0	5.0	0.83	$(1.5\pm0.4)\times10^{17}$	150±30	3.3	6.0	1.7
		$(8.0\pm2.0)\times10^{16}$	170±30	3.0	8.0						
CH_2OHCHO	CDMS	$\leq1.0\times10^{16}$	[200]	[3.0]	6.5	≤0.06	$\leq2.0\times10^{16}$	[200]	[3.0]	6.0	≤0.22
a-$(CH_2OH)_2$	CDMS	$(5.0\pm1.5)\times10^{16}$	220±50	3.5	5.8	0.47	$(2.5\pm0.5)\times10^{16}$	200±40	3.0	6.0	0.28
		$(3.5\pm1.0)\times10^{16}$	220±50	2.0	8.0						
g-$(CH_2OH)_2$	CDMS	$(2.0\pm0.5)\times10^{16}$	220±30	2.5	5.5	0.33	$(3.0\pm0.5)\times10^{16}$	200±40	2.8	6.0	0.33
		$(4.0\pm1.0)\times10^{16}$	220±30	2.5	8.0						
t-HCOOH	CDMS	$(9–12)\times10^{16}$	100–300	4.5–5.5	7.0	0.5–0.67	$(1.8–3.2)\times10^{16}$	100–300	4.0	6.5	0.2–0.36
H_2CCO	CDMS	$(3.8–6.0)\times10^{15}$	100–300	2.5	4.8	0.11–0.17	$(5.0–7.5)\times10^{15}$	100–300	3.0	6.0	0.06–0.08
		$(6.5–9.5)\times10^{15}$	100–300	2.8	7.7						
CH_2DCN	CDMS	$(2.0\pm0.5)\times10^{15}$	200±50	4.0	7.2	0.01	$(1.3\pm0.3)\times10^{15}$	200±60	3.0	6.0	0.014
NH_2CHO	JPL	$(1.3–1.9)\times10^{16}$	100–300	4.2–5.0	6.9	0.07–0.1	$(1.5–2.0)\times10^{15}$	100–300	3.0	6.0	0.01–0.02
C_2H_5CN	JPL	$(2.0–3.5)\times10^{15}$	100–300	3.0	5.0	0.04–0.08	$(3.5–6.0)\times10^{15}$	100–300	3.0	6.0	0.04–0.07
		$(2.0–3.5)\times10^{15}$	100–300	3.0	8.0						

†For IRAS 2A, some emission lines show two velocity components.

excitation temperatures ($T_{\rm ex}$) for six O-bearing COMs and one N-bearing COM that have enough clean lines detected; they are CH_3OH, CH_2DOH, CH_3CHO, C_2H_5OH, CH_3OCHO, $(CH_2OH)_2$, and CH_2DCN. In addition, we detect several strong lines of CHD_2OH and CD_3OH in both IRAS 2A and B1-c, implying a high deuteration rate of CH_3OH in low-mass protostars (e.g., Drozdovskaya et al. 2022), but this will not be studied in this paper.

For other species that are detected but that do not have enough clean lines to constrain N and $T_{\rm ex}$ independently, we provide either upper limits or a range of N assuming $T_{\rm ex}$ = 100–300 K, which is typical for hot cores. These species include the ^{13}C and ^{18}O isotopologs of CH_3OH, CH_3OCH_3, and CH_2OHCHO. $^{13}CH_3OH$ only has two strong lines covered in our spectral setup, and they are likely to be optically thick given their high Einstein A coefficients ($A_{\rm ij} > 4\times10^{-4}$ s^{-1}) and the low column density ratio with respect to the main isotopolog. The derived $CH_3OH/^{13}CH_3OH$ ratios are about 200, larger than the $^{12}C/^{13}C$ ratio in the vicinity of the Solar System (60–70), which implies that the column densities of $^{13}CH_3OH$ were underestimated due to optically thick lines. It is also difficult to constrain the $T_{\rm ex}$ with both of their upper energy levels $E_{\rm up} > 190$ K. Therefore, a lower limit of N is provided for $^{13}CH_3OH$ at the same $T_{\rm ex}$ as the main isotopolog. $CH_3^{18}OH$ and CH_3OCH_3 only have weak transitions ($A_{\rm ij}$<10^{-6} s^{-1}) covered in our data; therefore, upper limits were estimated assuming a fixed $T_{\rm ex}$ and line width (FWHM) at 3 km s^{-1}. The $T_{\rm ex}$ of $CH_3^{18}OH$ is fixed at the same value of CH_3OH, and the $T_{\rm ex}$ of CH_3OCH_3 is fixed at 200 K given that the covered transitions have high $E_{\rm up}$ of > 500 K. CH_2OHCHO have ∼10 lines detected, but most of them are blended with other stronger lines; thus, its column densities are reported as upper limits as well.

Besides CH_2DCN, we also detected two N-COMs, NH_2CHO and C_2H_5CN, and each of them have two to three strong lines covered in our data. Although these lines are unblended, they share similar $E_{\rm up}$ and therefore N and $T_{\rm ex}$ are degenerate with each other. In this case, we report N under $T_{\rm ex}$ = 100–300 K. Two 5-atom molecules, ketene (H_2CCO) and trans-formic acid (t-HCOOH), each have one strong line detected. They are often studied along with O-COMs given that they may serve as precursors of O-COMs in their formation routes. The column densities of H_2CCO and t-HCOOH are also estimated under $T_{\rm ex}$ = 100–300 K because of the degeneracy between N and $T_{\rm ex}$. Other species, such as the isotoplogs of abundant simple molecules (HDO, HDCO, $H^{13}CN$), sulfur-bearing molecules (SO, SO_2, and their isotopologs), and carbon-chain molecules (HC_3N, c- and l-C_3H_2), also have one or two strong lines detected in the spectra.

The best-fit column densities and excitation temperatures of the aforementioned species along with several other COMs and simple molecules are listed in Table 4.1. The LTE-modeled spectra along with the observed ALMA spectra of IRAS 2A and B1-c are displayed in Fig. 4.E.1–4.E.2. The transitions that were considered in the LTE fitting of ALMA spectra are listed in Table 4.L.1.

There are two special cases worth mentioning. The first is the determination of the CH_3OH column densities. Usually, $CH_3^{18}OH$ is used to infer the column density of CH_3OH assuming a $^{16}O/^{18}O$ ratio. This is because CH_3OH and $^{13}CH_3OH$ tend to be optically thick, and directly fitting their lines may lead to underestimation of their column densities. In our data, $CH_3^{18}OH$ was not robustly detected, but several CH_3OH lines with very high $E_{\rm up}$ (> 1000 K) or very low Einstein A coefficients ($A_{\rm ij}$ ∼10^{-7}

s^{-1}) were detected. These lines are intrinsically much weaker, and therefore more likely to be optically thin. The column densities of CH_3OH were determined based on those high-E_{up} or low-A_{ij} lines. Our fitting results show column density ratios of $CH_3^{18}OH/CH_3OH < 0.4\%$ for IRAS 2A and B1-c (Table 4.1), which are consistent with the expected ratio of 0.18%, assuming $^{16}O/^{18}O \sim 560$ in the local interstellar medium (Wilson & Rood 1994).

The second case is the double-peaked features observed in the emission lines of IRAS 2A, which are not rarely seen in hot cores (e.g., Chapter 2), and probably due to the dynamics of a circumstellar disk (Nazari et al. 2024c). The spectrum can be well fit by two velocity components with different N (and sometimes different T_{ex} and FWHM). By contrast, B1-c only shows one velocity component, even at the continuum peak.

The column densities ratios with respect to CH_3OH were calculated to facilitate the comparison with other sources or results in other studies. As famous hot corino sources, IRAS 2A and B1-c have been targeted in previous observations (Taquet et al. 2015; Yang et al. 2021; van Gelder et al. 2020). In particular, the COM ratios in B1-c reported by van Gelder et al. (2020) are consistent with our fitting results within 50% for O-COMs, and even within 10% for CH_3CHO and C_2H_5OH. For IRAS 2A, the COM ratios reported by Taquet et al. (2015) and Yang et al. (2021) are factors of 2–3 higher than our results, which is likely because their angular resolution or sensitivity was not as good as ours (insufficient angular resolution may lead to beam dilution, and insufficient sensitivity would hamper the robust detection of some species).

4.4.2 JWST

The continuum emission of IRAS 2A and B1-c is spatially unresolved with JWST/MIRI-MRS (as is shown in Fig. 4.A.3 for B1-c); therefore, we only focus on spectral analysis. Studies on COM ices in IRAS 2A have been carried out by Rocha et al. (2024), and here we perform a similar analysis for B1-c.

Based on the selection described in Sect. 4.3.2.5, we first selected the lab spectra with the mixing conditions that fit best with the observed B1-c spectrum (one spectrum for each candidate species), and then performed overall fittings to constrain the scaling factors of these lab spectra. A certain mixing condition refers to a certain combination of mixing constituent, mixing ratio, and temperature, among which mixing constituent has the most important influence on band profiles. There is limited availability in lab databases of mixing ratios for some species (especially COMs); therefore, we only focus on mixing constituents (Sect. 4.4.2.1) and temperatures (Sect. 4.4.2.2).

4.4.2.1 Constituents of ice mixtures

Figure 4.4 compares the lab spectra of different mixing constituents with the observations (panels a–j correspond to CH_4, SO_2, $HCOO^-$, HCOOH, CH_3CHO, C_2H_5OH, CH_3OCH_3, CH_3OCHO, and CH_3COCH_3). OCN^- and CH_3COOH are not shown in this figure since they only have lab data for one mixing constituent. CH_3CN is also excluded because its three bands between 6.8 and 7.4 μm have very similar profiles among different mixing constituents. The selection of the CH_3CN mixture was based on the results of Nazari et al. (2024a), who reported tentative detections of CH_3CN ice in the Near Infrared Spectrograph (NIRSpec) data, and the CH_3CN:H_2O:CO_2 mixture

is the main contributor of the observed band at 4.43 µm. The comparison between the B1-c spectrum and the lab spectra of OCN^-, CH_3COOH, and CH_3CN ices are shown in Fig. 4.C.2. In the following paragraphs, we introduce the comparison results for each species following the order of display in Fig. 4.4.

CH_4. In B1-c, the CH_4 band at 7.67 µm is superposed by gas-phase CH_4 lines in absorption, which have been removed before fitting with lab spectra (Sect. 4.3.2.4). However, none of the existing lab spectra fits perfectly with the observations in terms of peak position. The best two candidates are the H_2O mixture with a mixing ratio of 6:100, and a more complex mixture with H_2O, CH_3OH, and CO_2. The 7.67 µm band of these two mixtures is slightly redshifted and blueshifted from the observations, respectively. The same band of the pure CH_4 ice and the CH_4:H_2O (6:10) mixture is more redshifted and too narrow as well. This suggests that in reality the surrounding of CH_4 ice is dominated by H_2O, and there are also other species present. We finally chose the CH_4:H_2O (6:100) mixture to fit the observations considering that the mixing ratio of the CH_4:H_2O:CH_3OH:CO_2 (1:6:7:10) mixture is not very reasonable (too little H_2O and too much CH_3OH), but the fitting results are expected to be similar using either of the spectra.

SO_2. The 7.6 µm band of the CH_3OH mixture fits the blue wing of the observed CH_4 band best. Half of this band overlaps with the OCN^- band at 7.62 µm, and may lead to degeneracy in the overall fittings (see Sect. 4.4.2.4).

HCOOH. the H_2O mixtures fit the observations better than the pure HCOOH. The difference between the H_2O and H_2O:CH_3OH mixtures at ∼8.2 µm is tiny, only that the H_2O:CH_3OH mixture is slightly redshifted and fits the observed band better. The absorption features shortward of 7.0 µm in the HCOOH:H_2O:CH_3OH spectrum belong to CH_3OH, and was excluded during baseline correction (see Appendix 4.D).

CH_3CHO. It has been discussed as an example in Sect. 4.3.2.5 that the H_2O mixture fits the observations best. The 7.0 and 7.42 µm bands of the pure ice and CO mixture deviate from the observations.

C_2H_5OH. It is a bit hard to select among the pure ice, the H_2O mixture, and the CH_3OH mixture. The pure C_2H_5OH matches the observations better at the 7.2 µm band, but its other bands are too wide. The H_2O mixture fits better in 6.8–7.2 µm, but its 7.24 µm band is slightly blueshifted from the observations. The CH_3OH mixture has similar band profiles to the H_2O mixture, but the C_2H_5OH bands between 6.8 and 7.2 µm are blended with the CH_3OH bands, and it is difficult to accurately separated in the CH_3OH mixture. We finally chose the C_2H_5OH:H_2O mixture considering that H_2O is the dominant species in ice mantles, and the cases for some other O-COMs also show that H_2O-rich mixtures suit better than pure ices. The CH_3OH is a promising candidate as well, but not selected for technical reasons.

CH_3OCH_3. Except for the CO mixture, the pure ice and other two mixtures (with H_2O and CH_3OH) have similar band profiles at 8.59 µm. The H_2O has the best match

with observations in terms of peak position and band width, while the CH_3OH mixture cannot be fully excluded.

CH_3OCHO. The mixture with CO, H_2CO, and CH_3OH fits the observations at 8.1–8.4 µm better than the H_2O mixture (panel h of Fig. 4.4). In the H_2O mixture, the C–O stretching band of CH_3OCHO at 8.25 µm is significantly blueshifted and smoothed, which does not reproduce the observed profile. This suggests that CH_3OCHO is more likely to be formed in a CO-rich environment other than a H_2O-rich one. However, a caveat exists that the 8.02 µm band in the lab spectrum of the CH_3OCHO:CO:H_2CO:CH_3OH mixture is mainly contributed by H_2CO, not by CH_3OCHO. This raises the question how realistic is this mixing ratio between CH_3OCHO and H_2CO (1:20), considering that the observed 8.03 µm band also likely has a contribution from the CH_3COCH_3:H_2O mixture (panel i of Fig. 4.4). The relative strength of the H_2CO band will affect our estimation on the ice abundance of CH_3COCH_3. A more detailed discussion on how to deal with the H_2CO band blended in the lab spectrum of the CH_3OCHO:CO:H_2CO:CH_3OH mixture is provided in Appendix 4.F.

CH_3COCH_3. There are more mixtures measured in laboratories than other O-COMs (Rachid et al. 2020), and the comparison is separated into two Panels (i and j) in Fig. 4.4. Panel i shows that the H_2O mixture has all the bands blueshifted from the pure ice and the CO or CO_2 mixtures. In particular, the CCC asymmetric stretching band is significantly blueshifted from 8.14 to 8.03 µm. The double-peaked CH_3 symmetric deformation band at ∼7.33 µm is also blueshifted. These blueshifts make the H_2O mixture match the observations much better. Panel j shows that the CH_3OH mixture has the 8.14 µm band split into two peaks, not matching the observations. The mixture with H_2O and CO_2 has almost the same spectrum as the H_2O only mixture, but the 8.03 µm band is slightly weaker compared to the 7.3 µm band. The weaker 8.03 µm band is favored in the overall fittings since the observed B1-c spectrum tends to be overfit at this position by a combination of H_2CO and CH_3COCH_3 (see Sect. 4.4.2.3). We finally chose the CH_3COCH_3:H_2O:CO_2 (1:10:10) mixture for the overall fittings, but the difference would be small if using CH_3COCH_3:H_2O (1:20).

In summary, our comparison between the lab spectra and the observations reveals that most COM ices (except for CH_3OCHO) are surrounded by a H_2O-rich environment. For some species such as C_2H_5OH and CH_3OCH_3, mixing with CH_3OH cannot be ruled out. The possibility of CO-dominated mixtures is low, probably due to its desorption above 20 K. CH_3OCHO is an outlier that its surrounding is not dominated by H_2O, but rich in CO, H_2CO, and CH_3OH, implying a formation route of CO hydrogenation. However, this set of mixing constituents (CO+H_2CO+CH_3OH) is only measured for CH_3OCHO; hence, it is too early to conclude that CH_3OCHO has a different formation route than other COMs.

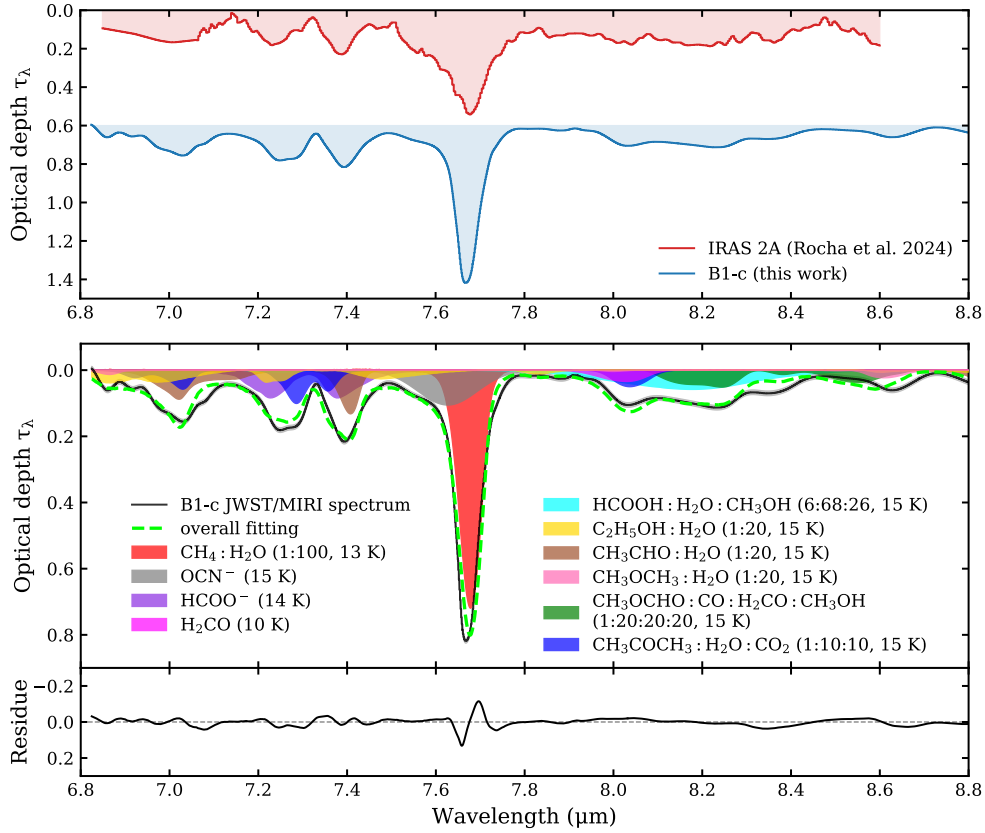


Figure 4.7: Isolated JWST/MIRI-MRS spectrum between 6.8 and 8.8 μm of IRAS 2A (Rocha et al. 2024) and B1-c (top panel), along with the best-fit decomposition of the B1-c spectrum (bottom panel). The IRAS 2A spectrum was binned by a factor of four, and therefore shows a lower spectral resolution. The B1-c spectrum is the version after removing the gas-phase lines (i.e., the black line in Fig. 4.3d). In the bottom panel, nine out of 12 candidate species that have statistically significant contribution to the 6.8–8.8 μm range are displayed, with the other three species (SO_2, CH_3COOH, and CH_3CN) excluded. The best-fit mixing conditions (constituents, ratio, and temperature) selected in Sect. 4.3.2.5 are labeled in the legend. The residue is shown below. A step-by-step version of the best-fit results of the B1-c spectrum is given by Fig. 4.K.1.

4.4.2.2 Temperature of ice mixtures

Unlike mixing constituents, varying temperature only induces very small differences in band profiles as long as the temperature is below the crystallization point. Figure 4.5 shows the comparison between the observations and the lab spectra of COM ices under different temperatures. Two species, CH_3OCHO and C_2H_5OH are shown as examples of two scenarios, and the remaining three COMs (CH_3CHO, CH_3OCH_3, and CH_3COCH_3) are shown in Fig. 4.C.1. For all the considered COMs, the pure ices show significant differences in the band profiles when transiting from amorphous to crystalline state. As temperature increases, the bands become narrower and sharper,

some even split into two bands. For some species such as CH_3OCHO, the changes in band profiles during this transition are also distinct in the ice mixtures. Because of the dilution of other constituents (e.g., H_2O), the band width may remain similar, but the relative intensity or peak position of each band will change significantly after crystallization. On the other hand, the band profiles of C_2H_5OH:H_2O mixture only show small changes after crystallization, which are very difficult to distinguish when compared with observations. The details of other three COMs are described in Appendix 4.C. In general, CH_3OCH_3 and CH_3COCH_3 are of the same type as CH_3OCHO; and CH_3CHO is more like C_2H_5OH.

By comparing the lab spectra under different temperatures, we can infer the range of crystalline temperature under laboratory conditions ($T_{\rm crystal,\,lab}$) for pure and mixed COM ices . We also compared the lab spectra with the JWST spectrum of B1-c and constrained the temperature ranges of the detected COM ices (summarized in Table 4.G.1). For pure COM ices, $T_{\rm crystal,\,lab}$ is $\lesssim$100 K. Under astrophysical conditions, $T_{\rm crystal}$ is usually 20%–40% lower. For mixed COM ices, $T_{\rm crystal,\,lab}$ is slightly higher, $\gtrsim$100 K. Despite noticeable changes in band profiles of crystalline ices, the observations can only constrain the laboratory temperature $T_{\rm lab}$ of the detected COM ices up to $\sim$100 K (equivalent to 60–80 K in space).

Similar degeneracy is also found for $T_{\rm lab} < 70$ K in IRAS 2A using the `ENIIGMA` fitting code (Rocha et al. 2024). Although IRAS 2A shows evidence of more thermal processing than B1-c by its double-peaked CO_2 band at 15.2 µm (e.g., Brunken et al. 2024a), the difference in thermal processing is hardly manifested in the band profiles of COM ices. This means that we cannot tell how many COM ices are as cold as 10 K and how many are as warm as 60 K; even if we know, we are still not able to distinguish whether these ices are formed in the cold collapse stage ($\sim$10 K in space) or the subsequent warm-up stage, since the ices could form in cold environment but then be heated. The only known information is that the observed bands of COM ices are relatively broad and smooth, which is not in favor of the sharp spectral features of crystalline ices. This does not rule out of the presence of crystalline ices, but suggests the observed ices to be mainly amorphous ($T < T_{\rm crystal}$), which is reasonable considering that only a small part of the envelope is heated by the protostar to a temperature as high as $T_{\rm crystal}$.

4.4.2.3 Decomposition of the B1-c spectrum in 6.8–8.8 µm

As introduced in Sect. 4.3.2.5, after selecting the most suitable ice mixture of each candidate species, we performed least-squares and MCMC fittings to search for the best-fit scaling factors of lab spectra to the observed B1-c spectrum between 6.8 and 8.8 µm. We focus on the spectral decomposition in this subsection, and leave the statistics to the next one. The spectral fitting of the 6.8–8.8 µm range for B1-c is similar to that for IRAS 2A; therefore, our analysis is generally based on the results of Rocha et al. (2024). The top panel of Fig. 4.7 displays the isolated JWST/MIRI spectrum in the COM fingerprint range (6.8–8.8 µm) of IRAS 2A and B1-c. Despite some difference introduced by the different methods of removing the superposed gas-phase lines, the absorption bands show very similar profiles in these two sources (see elaboration below). The bottom panel of Fig. 4.7 shows the best-fit results and residues of the B1-c spectrum, and we reserve the statistical details to Sect. 4.4.2.4.

7.5–7.8 µm. This range mainly contains absorption bands of simple molecules, and therefore is introduced first. The strongest band at 7.67 µm is attributed to the CH_4 deformation mode. The residues around this band is due to the deviation of peak position between the observations and the selected lab spectrum, which is the best among the candidates but still not optimal. The blue wing of the CH_4 band ∼7.6 µm is much weaker in B1-c than in IRAS 2A. It corresponds to one OCN^- band and one SO_2 band, which overlap with each other by ∼50%. The SO_2 component is not shown in Fig. 4.7 since OCN^- was favored in the overall fittings by least-squares and MCMC, instead an upper limit was estimated for SO_2. There is also a red wing of the CH_4 band in IRAS 2A, and is mainly made up of a broad CH_3COOH band at ∼7.75 µm. This red wing is not present in B1-c after we traced a local continuum that is close to the observed spectrum. Rocha et al. (2024) also discussed in their Sect. 4.2.4 and Appendix J that using a different local continuum could remove the CH_3COOH component from the overall fit, and CH_3COOH is only considered as tentatively detected in IRAS 2A.

6.8–7.5 µm. This range contains four absorption bands at 6.85, 7.02, 7.26, and 7.40 µm. Rocha et al. (2024) attribute them to two O-COMs (CH_3CHO and C_2H_5OH) and $HCOO^-$, which also applies to B1-c. The CH_3CHO:H_2O mixture is one of the contributors to the 7.02 and 7.40 µm bands. $HCOO^-$ has two bands at 7.23 and 7.38 µm, making up the red part of the observed bands at 7.26 and 7.40 µm. C_2H_5OH has three bands at 6.86, 6.97, and 7.24 µm, of which the contribution is less significant but not negligible.

The peak position of the third band is slightly different between IRAS 2A and B1-c. In IRAS 2A, it is observed to peak at 7.24 µm; more blueshifted than the same band in B1-c which peaks at 7.26 µm. The offset between the $HCOO^-$ 7.23 µm band and the observed 7.26 µm band in B1-c leaves an underfit area at 7.3 µm. There are a few candidate species that have absorption bands at around 7.3 µm: CH_3CN, CH_3COOH, and CH_3COCH_3. CH_3CN has recently been reported to be tentatively detected in NIRSpec spectra of several protostellar sources (Nazari et al. 2024a), but its abundance is mainly constrained by the other two stronger bands at ∼6.9 and 7.1 µm, where the absorption is weak in observations. Similarly, the abundance of CH_3COOH is constrained by the stronger band at 7.75 µm, which is degenerate with the local continuum subtraction. Even if CH_3COOH is present, its 7.3 µm band also tends to be too weak and broad to fill the gap at 7.3 µm in the observed B1-c spectrum. The best candidate turns out to be the H_2O-rich mixtures of CH_3COCH_3, of which the CH_3 symmetric deformation band at 7.33 µm can help solve the problem, and the other two bands at 7.03 and 8.03 µm also fit well with the observations (see panel i of Fig. 4.4). In Rocha et al. (2024), CH_3COCH_3 is not considered as firmly detected in IRAS 2A based on their recurrence analysis (see their Sect. 4.2.2). A possible explanation is that they trace a local continuum that is not close enough to the observed spectrum when isolating the COM fingerprint range, and therefore leave some space for CH_3COOH bands at 7.7 and 7.3 µm. However, if they trace a local continuum close to the observed spectrum as we did, there will also be an unfit area at ∼7.3 µm (shown in their Fig. J.2), which can be attributed to CH_3COCH_3.

7.8–8.8 µm. This range is composed of several blended bands between 7.9 and 8.45 µm and a small band at 8.63 µm. The band at 8.03 µm is contributed and slightly overproduced by a combination of H_2CO band at 8.02 µm and CH_3COCH_3:H_2O band at 8.03 µm. The contribution of H_2CO was fixed by fitting its another band at 6.67 µm (see Appendix 4.F). To alleviate the overestimation, we chose the CH_3COCH_3 mixture with H_2O and CO instead of the H_2O-only mixture, since the H_2O:CO_2 mixture has a weaker 8.03 µm band (Fig. 4.4j), although the difference is small. The broad band peaking at 8.24 µm is composed of the broad HCOOH band at 8.17 µm and the double-peaked CH_3OCHO band at ∼8.2 µm. The observed band at 8.63 µm matches best with the CH_3OCH_3:H_2O band at 8.6 µm, with potential contribution from the SO_2:H_2O band at 8.67 µm. Similar to CH_3COCH_3, CH_3OCH_3 is also considered as tentative detection for IRAS 2A based on the recurrence analysis; however, this is likely because Rocha et al. (2024) only perform the fittings over 6.8–8.6 µm, missing a large portion of the observed 8.63 µm band. After extending the considered wavelength range to 8.8 µm, CH_3OCH_3 is likely to be considered as a firm detection (as is shown in Fig. 4.4g). Besides the contribution of the aforementioned species, there is still an unfit band between 8.3 and 8.4 µm, also seen in other JOYS+ sources. It is recently found to be attributable to CH_2OH (priv. comm. with W. Rocha), and will be studied in a future paper. On the other hand, the region at ∼8.5 µm is overfit by the CH_3 rocking band of CH_3OCHO. The reason is unclear, and could be related to the local continuum subtraction.

The decomposition results of the COM fingerprint range between 6.8 and 8.8 µm are generally the same for B1-c (this work) and IRAS 2A Rocha et al. (2024), in spite of different fitting strategies. The only difference is that we tend to consider CH_3OCH_3 and CH_3COCH_3 as firmly detected and provide constraints on their ice column densities instead of upper limits (see Sect. 4.4.2.4).

4.4.2.4 Fitting statistics

We adopted least-squares and MCMC fittings to find the best-fit values and uncertainties of the scaling factors of lab spectra (see Table 4.I.1). In the least-squares fitting, we considered all the 12 candidate species: CH_4, SO_2, OCN^-, $HCOO^-$, HCOOH, CH_3CHO, C_2H_5OH, CH_3OCH_3, CH_3OCHO, CH_3COOH, CH_3COCH_3, and CH_3CN (i.e., $N = 12$ in Eq. (4.3)). The scaling factor of H_2CO was determined from the band at 6.67 µm, and was fixed when fitting the 6.8–8.8 µm range (see details in Appendix 4.F).

The least-squares fitting results show that the lab spectrum scaling factors of most of the candidate species were well constrained with relative errors smaller than 10% (see Table 4.I.1). There are four exceptions: SO_2, CH_3OCH_3, CH_3COOH, and CH_3CN, and particularly, SO_2 and CH_3COOH have very little contribution. The absence of CH_3COOH is because we traced a local continuum close to the observed spectrum, which eliminated the red wing of the CH_4 band at 7.67 µm. However, the lack of SO_2 is more likely because its band at 7.6 µm is too close to the OCN^- band at 7.62 µm, and hence highly degenerate with each other. The blue wing of the CH_4 band in B1-c is less prominent than in IRAS 2A, making it more difficult to distinguish between the contribution from SO_2 and OCN^-. The degeneracy origin prevents us from drawing the conclusion that SO_2 is not present or more depleted in B1-c than in IRAS 2A.

Table 4.2: Ice column densities of candidate species derived from the least-squares and MCMC fittings to the JWST/MIRI-MRS spectra of NGC 1333 IRAS 2A and B1-c.

	Source				B1-c					NGC 1333 IRAS 2A [a]				
Species	Band λ (μm)	Mode [b]	$A/10^{-18}$ (cm molec^{-1})	Ref. [c]	$N_{\rm ice}$ (10^{17} cm^{-2})	X/CH_3OH (%)	X/H_2O (%)	$T_{\rm ice}$ (K)	mixture	$N_{\rm ice}$ (10^{17} cm^{-2})	X/CH_3OH (%)	X/H_2O (%)	$T_{\rm ice}$ (K)	mixture
H_2O	13.2	libration	32	1,2	250±17	–	≡100	15	pure	300±12	–	≡100	15,160	pure
CH_3OH	9.74	CO str.	15.6 [d]	3	30±10	≡100	11.9±4.1	–	Gaussian	19±4	≡100	6.3±1.4	–	Gaussian
CH_4	7.67	CH_4 def.	8.4	1,2,4	9.3±5.1	30.9±19.8	3.7±2.0	13	w/H_2O	$4.9^{+7.5}_{-3.2}$	26^{+40}_{-18}	$1.6^{+2.5}_{-1.1}$	15	w/H_2O
SO_2	7.60	SO_2 str.	34	4	<0.47	<1.6	<0.19	10	w/CH_3OH	$0.6^{+1.9}_{-0.4}$	$3.2^{+10}_{-2.2}$	$0.2^{+0.6}_{-0.1}$	10	w/CH_3OH
OCN^-	7.62	$2\nu_2$	7.45	5	4.1±2.2	13.6±8.7	1.6±0.9	15	NH_3+HNCO [e]	$3.7^{+6.6}_{-3.3}$	19^{+35}_{-17}	$1.2^{+2.2}_{-1.1}$	15	NH_3+HNCO [e]
H_2CO	8.02	CH_2 rock.	1.5	1	4.5±1.3	14.9±6.7	1.8±0.6	15	pure	$12.4^{+19.7}_{-6.6}$	65^{+105}_{-37}	$4.1^{+6.6}_{-2.2}$	15	w/CO, CH_3OH, CH_3OCHO
$HCOO^-$	7.38	CO str.	17	7	0.93±0.51	3.1±2.0	0.37±0.21	14	H_2O+NH_3 +HCOOH [e]	$1.4^{+2.4}_{-0.4}$	$7.4^{+12.7}_{-2.6}$	$0.5^{+0.80}_{-0.1}$	14	H_2O+NH_3 +HCOOH [e]
HCOOH	8.23	CO str.	29	1	1.1±0.6	3.7±2.4	0.44±0.25	15	w/H_2O	$3.0^{+5.3}_{-1.7}$	16^{+28}_{-10}	$1.0^{+1.8}_{-0.6}$	15	w/H_2O, CH_3OH
CH_3CHO	7.42	CH_3 sym. def. + CH wag.	4.1	8,9	1.9±1.0	6.2±4.0	0.74±0.41	15	w/H_2O	$2.2^{+2.8}_{-1.4}$	12^{+15}_{-8}	$0.7^{+0.9}_{-0.5}$	16	w/H_2O
C_2H_5OH	7.24	CH_3 sym. def.	2.4	10	2.1±1.2	7.1±4.6	0.85±0.47	15	w/H_2O	$3.7^{+4.5}_{-0.5}$	19^{+24}_{-5}	$1.2^{+1.5}_{-0.2}$	15	w/H_2O
CH_3OCH_3	8.59	COC str. +CH_3 rock.	5.55 [d]	9	0.51±0.28	1.7±1.0	0.20±0.11	15	w/H_2O	<5.8	<13	<0.8	–	
CH_3OCHO	8.25	CO str.	24.9	11	0.48±0.26	1.6±1.0	0.19±0.11	15	w/CO, H_2CO, CH_3OH	$0.2^{+0.4}_{-0.1}$	$1.0^{+2.1}_{-0.6}$	$0.07^{+0.13}_{-0.03}$	15	w/CO, H_2CO, CH_3OH
CH_3COOH	7.82	OH bend.	45.7	5	<0.21	<0.7	<0.08	10	w/H_2O	$0.9^{+1.3}_{-0.6}$	$4.7^{+6.9}_{-3.3}$	$0.3^{+0.4}_{-0.2}$	10	w/H_2O
CH_3COCH_3	7.33	CH_3 sym. str.	10.2 [d]	12,13	1.7±1.0	5.8±3.7	0.69±0.38	15	w/H_2O, CO_2	<1.1	<6	<0.4	15	w/H_2O
CH_3CN	7.27	CH_3 sym. def.	1.2	14	<3.1	<10	<1.2	15	w/H_2O, CO_2	<5.0	<26	<1.7	15	w/H_2O

[a] Results obtained from Rocha et al. (2024).
[b] Abbreviation of vibrational modes: def. = deformation; str. = stretching; rock. = rocking; sym. = symmetric; wag. = wagging; bend. = bending.
[c] References of band strengths A: [1] Bouilloud et al. (2015); [2] Hudgins et al. (1993); [3] Luna et al. (2018); [4] Boogert et al. (1997); [5] Rocha et al. (2024); [7] Schutte et al. (1999); [8] Hudson & Ferrante (2020); [9] Terwisscha van Scheltinga et al. (2018); [10] Boudin et al. (1998); [11] Terwisscha van Scheltinga et al. (2021); [12] Hudson et al. (2018); [13] Rachid et al. (2020); [14] Rachid et al. (2022).
[d] See Appendix 4.H for additional details.
[e] These constituents are not added intentionally, but are ingredients to form the target ions (i.e., OCN^- and $HCOO^-$).

Instead, an upper limit was estimated for SO_2 by only scaling the lab spectrum of SO_2 to the observations (Fig. 4.4b). For CH_3OCH_3 and CH_3CN, the best-fit scaling factors are not negligible, but the relative errors are slightly larger than other species, mainly because their bands correspond to weak and blended absorption features in the observations.

In the least-squares fitting, eight out of 12 candidate species have fairly small relative uncertainties (<10%). In the next step, we performed an additional MCMC fitting on these eight species plus CH_3OCH_3 (i.e., nine species in total) to get a better understanding of the uncertainty level. CH_3OCH_3 was taken into consideration because it is the main contributor to the observed band at 8.63 µm, and it is one of the most abundant O-COMs observed in the gas phase. We did not include CH_3CN in the MCMC fitting since it does not have a characteristic band in the 6.8–8.8 µm range; all the three bands are weak and blended with others. Including CH_3CN would also make the MCMC fitting less convergent, suggesting that its contribution is less significant. Instead, we report upper limits for CH_3CN like for SO_2 and CH_3COOH.

The best-fit scaling factors derived from the MCMC fitting are consistent with the least-squares within 5%. Fig. 4.K.2 displays the posterior distributions of each component. Most components are rather independent from each other, except for pairs that have similar band locations (e.g., CH_4 and OCN^-, HCOOH and CH_3OCHO). The relative uncertainties generally increase, but still at a low level of ∼10%, except for CH_3OCH_3, whose relative uncertainty is ∼20%. For the left three species, SO_2, CH_3COOH, and CH_3CN, we report only upper limits. We manually scaled the lab spectrum of each species to the observed B1-c spectrum (e.g., Fig. 4.C.2), and the maximum scaling factors allowed by the observations were converted to upper limits of ice column densities.

Based on the fitting results of scaling factors, the ice column densities of each species can be calculated using Eq. (4.4). The absolute ice column densities are on the same order of magnitude in B1-c (this work) and IRAS 2A (Rocha et al. 2024). The uncertainties of the ice column densities are propagated from scaling factors (∼10%), band strengths (∼20%), and the steps taken to isolate the COM bands between 6.8 and 8.8 µm (Sects. 4.3.2.1–4.3.2.4). The uncertainties of the isolation steps should be dominant but also difficult to quantify, since that they are more or less subjective. Here we estimated a conservative uncertainty level of 50% for isolating the COM fingerprint range, which resulted in a ∼55% total uncertainties of the ice column densities. This uncertainty level is consistent with that of IRAS 2A reported in Rocha et al. (2024). The derived ice column densities and uncertainties, along with the temperatures and mixing constituents of the best-fit lab spectra are summarized in Table 4.2 for both B1-c (this work) and IRAS 2A (Rocha et al. 2024).

4.4.2.5 CH_3OH and H_2O bands

The strategies for fitting the CH_3OH band at 9.74 µm and the H_2O and at 13 µm have been introduced in Sect. 4.3.2.6. Here we report the fitting results and the derived ice column density of CH_3OH and H_2O. The CH_3OH band at 9.74 µm is isolated and was fit by a Gaussian function (see top panel of Fig. 4.8). However, it is likely that the best-fit Gaussian function underestimates the real intensity of this band because of the silicate extinction. The extinction correction factor is not trivial to estimate, and a better strategy would be combining the two CH_3OH bands in the NIRSpec range

(at 3.4 and 3.9 µm) and fit all three bands simultaneously. Unfortunately, the CH_3OH band at 3.4 µm is below the detection limit of our observations of B1-c, probably due to the strong extinction by H_2O at ∼3.05 µm.

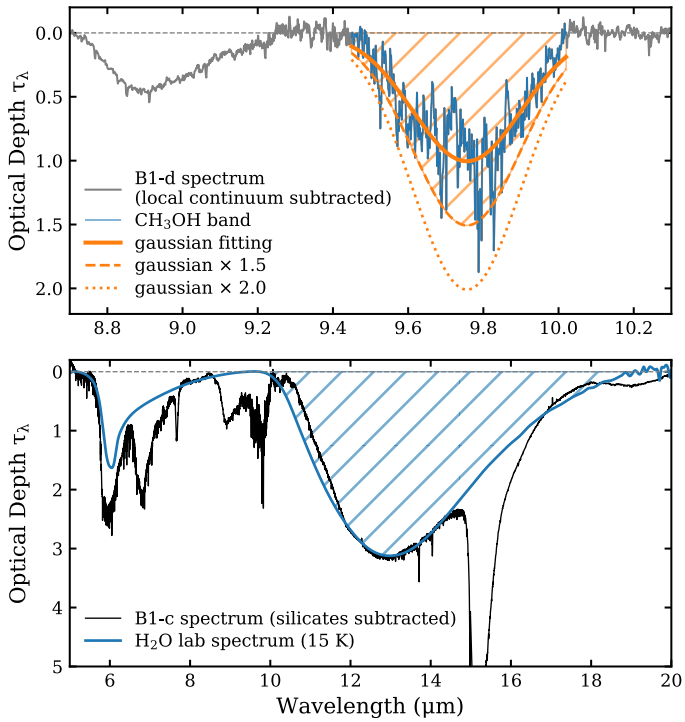


Figure 4.8: Fitting results of the CH_3OH band at 9.74 µm (top panel) and the H_2O band at 13 µm (bottom panel). The hatched regions indicate the integrated areas for calculating the ice column densities of CH_3OH and H_2O.

The strength of the 3.9 µm band is also not well constrained, since there is strong indication that a grain shape correction is needed to properly trace the continuum of this band (Dartois et al. 2022, 2024). Considering that NIRSpec data are not the focus of this paper and will be reserved for following studies, we performed a preliminary analysis of the 3.9 µm band and derived an upper limit of 4.0×10^{18} cm^{-2} for the CH_3OH ice column density (see details in Appendix 4.J). For comparison, the best-fit Gaussian to the 9.74 µm band resulted in a column density of 2.0×10^{18} cm^{-2}. We finally adopted a value of 3.0×10^{18} cm^{-2}, which corresponds to 1.5 times the best-fit Gaussian. As is shown in the top panel of Fig. 4.8, this correction factor of 1.5 is reasonable when considering the extinction of the silicate band at 9.8 µm. The uncertainty was estimated as $\pm1.0\times10^{18}$ cm^{-2}.

The fitting result of the H_2O band at 13 µm is shown in the bottom panel of Fig. 4.8. A scaling factor of 14.5 was estimated through visual inspection for the lab spectrum of H_2O ice at 15 K (Sect. 4.3.2.6), giving a column density of 2.5×10^{19} cm^{-2}. The slight excess between 10 and 12 µm is likely due to the over-subtraction of silicate features, as the synthetic silicates have a wider red wing of the 9.8 µm band than the

GCS 3 profile (Fig. 4.3b).

With the ice column densities of H_2O and CH_3OH derived, we are able to calculate the relative abundances of other species, which are presented in Table 4.2. The total ice abundance of the detected six COMs, namely CH_3OH, CH_3CHO, C_2H_5OH, CH_3OCH_3, CH_3OCHO, and CH_3COCH_3, is 14.6% with respect to H_2O ice in B1-c. As the most abundant COM, the relative abundance of CH_3OH is 11.9%, dominates more than 80% of all COMs. The upper limit of the most abundant N-COM, CH_3CN, is 1.2% with respect to H_2O. According to Rocha et al. (2024), the ice abundances of all the O-COMs and CH_3OH with respect to H_2O are 9.8% and 6.3% in IRAS 2A, respectively; the upper limit of CH_3CN is 1.7%. The ice ratios between CH_3OH and other COMs will be discussed in Sect. 4.5. We provide the statistics but leave out the discussion on the simple species as they are beyond the scope of this study.

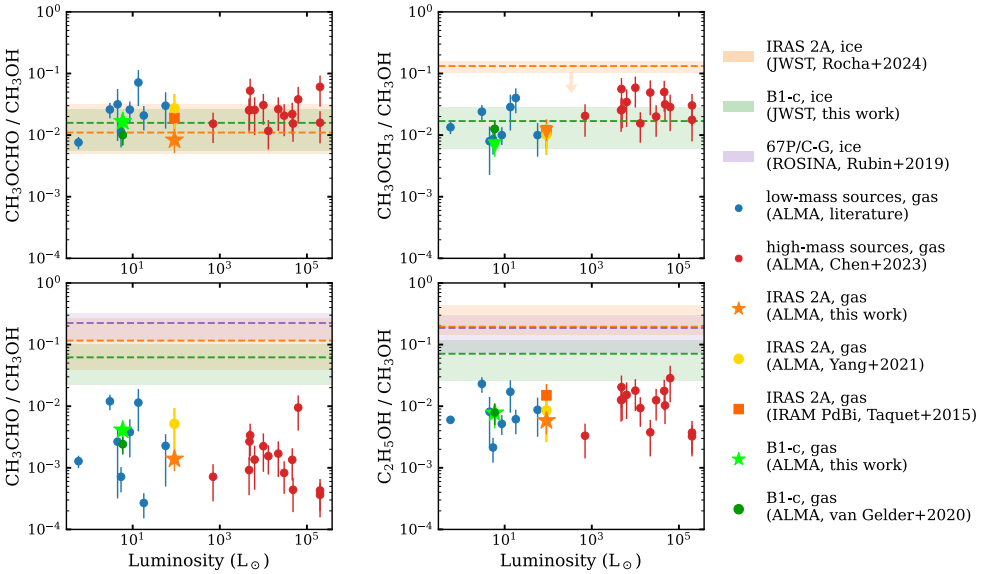


Figure 4.9: Ratios of four O-COMs with respect to CH_3OH in gas (data points) and ice (shaded regions). Gas ratios are collected from a total sample of nine low-mass sources from this work and literature (blue circles; Jørgensen et al. 2018; Manigand et al. 2020; van Gelder et al. 2020; Yang et al. 2021), and 14 high-mass sources from the CoCCoA survey (Chapter 2). Ice ratios are available for only two protostars (IRAS 2A and B1-c; Rocha et al. 2024, this work) and one comet (67P/C-G; Rubin et al. 2019b). All data of the two focused low-mass sources, NGC 1333 IRAS 2A and B1-c, are colored in orange and green, respectively. The gas ratios of IRAS 2A and B1-c derived in this work and previous observations are highlighted in stars and squares (Taquet et al. 2015; van Gelder et al. 2020), respectively. Upper limits are denoted by downward triangles or arrows.

4.5 Discussion

4.5.1 Column density ratios: Ice versus gas

The detection of COM ices by JWST (this work and Rocha et al. 2024) strongly suggests that the formation of COMs starts in the solid phase, as is illustrated in Fig. 4.1. However, it remains unclear to what extent the gas-phase chemistry will alter the composition of COMs after their sublimation from dust grains. Now, with COMs observable in both gas and ice, we are able to gain insights into this phase transition by comparing the COM ratios in both phases.

Our comparison is focused on four O-COMs: CH_3CHO, C_2H_5OH, CH_3OCH_3, and CH_3OCHO, which are detectable in both gas and ice. Their ratios with respect to CH_3OH in the two phases are summarized in Fig. 4.9. The number of sources that have COM ice ratios derived is limited. Therefore, in addition to the two case-studied sources IRAS 2A and B1-c, we also include the comet 67P/C-G for comparison, of which the ice abundances of detected molecules are reported in Rubin et al. (2019b). For the gas ratios, in addition to previous observations of IRAS 2A and B1-c (Taquet et al. 2015; Yang et al. 2021; van Gelder et al. 2020), we also collected literature data of other sources with different masses and luminosities from previous line surveys (i.e., PILS, PEACHES, and CoCCoA; Jørgensen et al. 2018; Manigand et al. 2020; van Gelder et al. 2020; Yang et al. 2021; Chapter 2) in order to better represent the gas-phase trend using a larger sample. For the PEACHES sources, we adopted the CH_3OH column density derived from optically thin minor isotopologs (i.e., $^{13}CH_3OH$ or $CH_3^{18}OH$) by van Gelder et al. (2022).

Figure 4.9 shows two different cases: CH_3OCHO and CH_3OCH_3 show constant gas-phase ratios among the ALMA sample, which also matches well with the ice ratios. In contrast, CH_3CHO and C_2H_5OH show larger scatter in the gas-phase ratios, and the ice ratios are higher than the gas ones by 1–2 orders of magnitude. The possible reasons are discussed in the following subsections.

4.5.2 Origin of the similarity in complex organic molecule ratios between ice and gas

The top panels in Fig. 4.9 show constant gas-phase ratios of CH_3OCHO and CH_3OCH_3 with respect to CH_3OH in both low- and high-mass sources, and the good consistency in the ratios between gas and ice. These can be explained by direct inheritance or strong chemical links among CH_3OCHO, CH_3OCH_3, and CH_3OH. Direct inheritance means that the bulk of CH_3OCHO, CH_3OCH_3, and CH_3OH remain intact during the transition from ice to gas, and their relative abundances also stay on the same level in the gas phase. In the hot core phase, either their absolute abundances remain stable, or their abundances are altered similarly due to a strong chemical link. For instance, they are involved in a similar chemical network and share similar routes of formation and destruction. The numerical simulations on grain-surface chemistry by Simons et al. (2020) suggest that CH_3OH and CH_3OCHO could share the same precursor (CH_3O radical) in their formation pathways. Although CH_3OCH_3 is not discussed in Simons et al. (2020), studies by Garrod et al. (2022) who present simulation results of a three-phase chemical network (bulk ice, grain surface, and gas phase) show that

the CH_3O radical plays an important role in the formation of CH_3OCH_3 ice. The possibility of having chemical links among CH_3OH, CH_3OCHO, and CH_3OCH_3 is also supported by the remarkably constant ratio between CH_3OCHO and CH_3OCH_3 observed in a large sample of protostellar sources with different masses and luminosities (Coletta et al. 2020). However, it is difficult to clearly distinguish between the cases of direct inheritance and strong chemical links only by observations; more studies by simulations and experiments are needed to break this degeneracy.

There is also observational evidence that COMs other than CH_3OH can be inherited to later stages (e.g., the recent detection in the protoplanetary disk around Oph IRS 48; Brunken et al. 2022; Yamato et al. 2024; Booth et al. 2024b). The trend of O-COM abundances in the protoplanetary disks seem to be consistent with those observed in protostars (e.g., there are more CH_3OCH_3 and CH_3OCHO than CH_3CHO and C_2H_5OH in the gas phase), though the exact ratios with respect to methanol are different. Theoretically, the gas-phase COMs observed in the Class 0 stage can be accreted to the circumstellar disk, and some of them will freeze out in the disk midplane at the Class II stage. However, it is not clear to what extent the high temperature environment generated by the accretion shock at the disk-envelope interface would destroy these molecules.

4.5.3 Origin of the difference in complex organic molecule ratios between ice and gas

4.5.3.1 Gas-phase reprocessing

In contrast to CH_3OCHO and CH_3OCH_3, CH_3CHO and C_2H_5OH show not only large scatters in their gas-phase ratios with respect to CH_3OH, but also inconsistent ratios between ice and gas (bottom panels in Fig. 4.9). The large scatter in gas-phase ratios among different sources may be caused by the physical structures (e.g., the presence of a circumstellar disk can lower the temperature in the hot core; Nazari et al. 2022b, 2023a). It is also possible that gas-phase chemistry after ice sublimation plays an important role in the chemical evolution of CH_3HO and C_2H_5OH, in which case the variance in physical properties (density, temperature, and UV intensity) are reflected in the chemical abundances. Chapter 2 suggest that C_2H_5OH may go through gas-phase destruction in the hot core phase. For example, C_2H_5OH can be H-abstracted by OH radicals or halogen atoms and finally converted into HCOOH, H_2CO, and two other O-COMs, CH_2OHCHO and CH_3COOH (Skouteris et al. 2018). The situation of CH_3CHO is more complicated. Simulation results reported by Garrod et al. (2022) show that a significant proportion of CH_3CHO is formed in the gas phase, leading to an increasing gas-phase abundance of CH_3CHO during the warm-up stage, compared to the ice abundance in the cold collapse stage. However, our observations reveal the opposite case. The CH_3CHO ratios with respect to CH_3OH are actually lower in gas than in ice by more than one order of magnitude, which implies more CH_3CHO is consumed (either destructed into smaller species or converted into larger COMs) than reproduced in the gas phase.

It is not clear to what extent the gas-phase chemistry in hot cores can change the abundances of COMs after they sublimate from dust grains (e.g., Balucani et al. 2015). The influence of gas-phase chemistry in the COM abundances can be a few factors or orders of magnitude, and it is likely to vary among different species. Astrochemical

models and experiments are required to further investigate the gas-phase chemical network in hot cores.

4.5.3.2 Spatial distribution

Besides the influence of gas-phase chemistry, it is also important to consider observational effects (e.g., beam dilution). ALMA and JWST are tracing different parts of protostellar systems: ALMA only observes the emission (or absorption) lines of gas-phase molecules, mostly from the central hot core region; in comparison, JWST traces materials along the line of sight in a pencil-like beam, and the observed materials can be in both phases, as long as their spectral features fall between 0.6 and 27.9 μm (i.e., the wavelength coverage of NIRSpec plus MIRI). The molecules detected by JWST, in particular those in the solid phase, are mostly coming from the extended envelope instead of the central hot core.

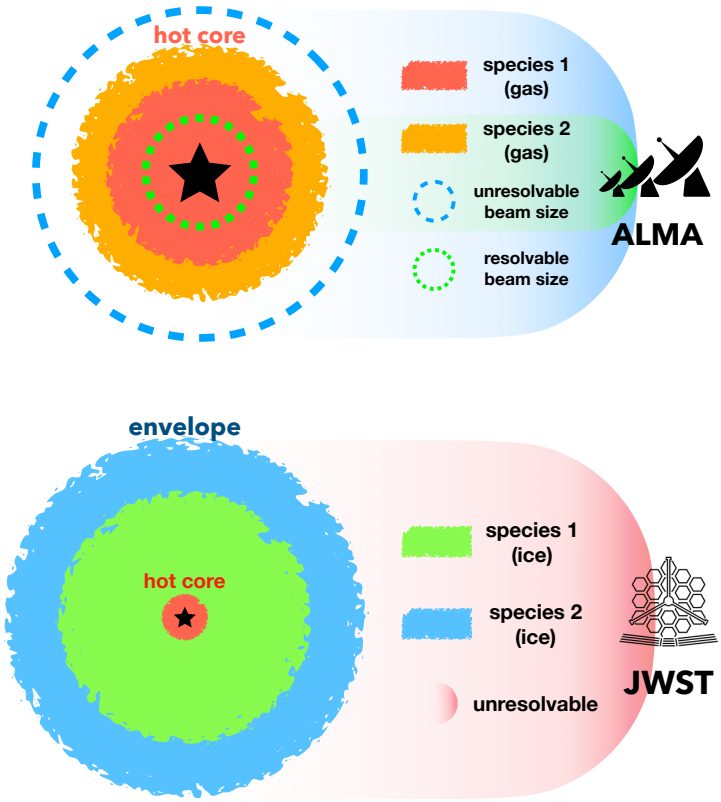


Figure 4.10: Schematics of how different spatial distributions of two species can affect the column density ratios estimated from ALMA (top) and JWST (bottom) observations. In each panel, two colors are used to indicate two species with different spatial distribution. In reality, the distance between telescopes and sources is very large, and the beam should be like a thin cylinder. For visualization, the beam sizes are exaggerated here.

Figure 4.10 shows how spatial distribution can affect the observed column density ratios between two species. In ALMA observations (top panel of Fig. 4.10), it is possible to spatially resolve a hot core; that is, the spatial scale corresponding to the beam size is smaller than the region where most of the ice mantles has sublimated into the gas phase. Assume species 1 has higher sublimation temperature than species 2, then its gas material will be present and emit in a more compact region (i.e., within a smaller radius). If the beam size is small enough to resolve the emitting regions of both species 1 and 2, the observed ratio between species 1 and 2 will be representative of the actual ratio. On the other hand, if the beam is not able to resolve the emitting region of any of the species, then its (or their) emission will be diluted within the beam, and the ratio between species 1 and 2 will be underestimated, since the beam dilution effect is more severe for the more compactly distributed species. In the case of our ALMA observations, the angular resolution is $\sim$0.1″, equivalent to $\sim$32 au at the distance of 320 pc for IRAS 2A and B1-c. As already estimated in Sect. 4.4.1.1 using Eq. (4.5), the hot core sizes of IRAS 2A and B1-c are 147 au and 37 au, respectively, both spatially resolvable. This means that the gas-phase COM ratios derived from our ALMA data are likely to be free from beam dilution effects.

In JWST observations (bottom panel of Fig. 4.10) the observed ice column densities of COMs are integrated along the line of sight in the protostellar envelope, and therefore dependent on the spatial distribution of these species. Assume that the ice of species 1 is present in a less extended region in the envelope than species 2. This situation is possible when species 2 starts to form in the solid phase at lower density than species 1 due to different formation mechanisms. Since JWST observes an integrated abundance (i.e., column density) along the line of sight, the integration is done over a more extended region for species 2, and as a result, the ice column density ratio between species 1 and 2 estimated from JWST observations is lower than the value if we only pick a small region inside the envelope. If species 1 appears to be CH_3OH and species 2 corresponds to CH_3CHO or C_2H_5OH, and the resolution of ALMA observations is high enough to resolve the emitting regions of both species, than the ice ratios between CH_3CHO (or C_2H_5OH) and CH_3OH estimated from JWST observations will be higher than the gas ratios derived from ALMA observations, as we find in this work.

More quantitatively, we test the influence of spatial distribution on column density ratios between two species with a toy model. We took the 1D spherical model created by Kristensen et al. (2012) for IRAS 2A. The temperature and H_2 number density profiles of this model are shown in the top left panel of Fig. 4.11, with corresponding cartoons shown in the bottom panels. We calculated the H_2 column density (i.e., number density integrated over radius) of an envelope where ices are present between an inner and an outer radius ($R_{\rm in}$ and $R_{\rm out}$). For each species, $R_{\rm in}$ and $R_{\rm out}$ can be different; $R_{\rm in}$ is considered as the sublimation boundary of ices, and $R_{\rm out}$ corresponds to the formation boundary of ices; that is, ices are only present within the envelope. Inside of $R_{\rm in}$, ices have sublimated into the gas phase, and outside of $R_{\rm out}$, ices have not formed yet. We assumed that the density distributions of all the species (in both gas and ice) are proportional to that of H_2, and we compared the H_2 column densities of envelopes with different sizes.

Specifically, we considered two cases. In the first case, envelopes have a fixed $R_{\rm in}$ but different $R_{\rm out}$ (as is shown in the bottom middle panel of Fig. 4.11). $R_{\rm in}$ is fixed

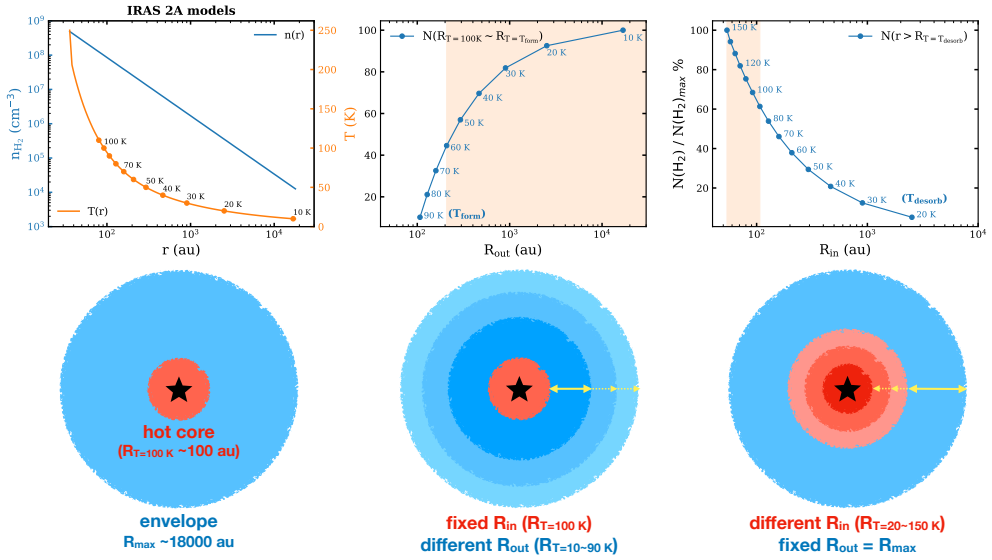


Figure 4.11: One-dimension model of IRAS 2A and the integrated H_2 column densities (top panels), along with the explanatory cartoons (bottom panels). Top left: H_2 number density (blue) and temperature (orange) profiles of the physical model of IRAS 2A (Kristensen et al. 2012). Top middle: normalized H_2 column densities in envelopes with a fixed inner radius (where $T = 100$ K) and different outer radii (where $T = 10$–90 K). Top right: normalized H_2 column density in an envelope with a fixed outer radius (the max radius in the model, ∼18000 au) and different inner radii (where $T = 20$–150 K). In the top middle and right panels, each point corresponds to an H_2 column density; that is, the integrated H_2 number density (top left panel) from an inner radius to an outer one. The orange shaded regions indicate the likely inner or outer radii in real cases. Bottom panels show corresponding cartoons of the toy model. Hot cores and envelopes are colored in red and blue, respectively. In the bottom middle and right panels, yellow arrows indicate the span of envelopes with different outer or inner radii.

as the radius where $T = 100$ K, and $R_{\rm out}$ are set as radius where $T = 10$, 20, ..., and 90 K. These temperatures are also noted as $T_{\rm form}$ since they by design correspond to the beginning of ice formation. The top middle panel of Fig. 4.11 shows how the H_2 column density $N(H_2)$ changes with $R_{\rm out}$. As introduced in Sect. 4.1 and also supported by studies on gas-phase observations (e.g., Coletta et al. 2020; Nazari et al. 2022a; Chapter 2), O-COMs are suggested to start forming in the cold dense pre-stellar phase through CO hydrogenation. According to the constraints on temperature set by the observations (last column in Table 4.G.1), the COM ices must have been formed under 100 K in laboratories, which is equivalent to ∼60 K in space. The difference in $N(H_2)$ introduced by different $T_{\rm form}$ is 50–60% (i.e., about a factor of two).

In the second case, we compared $N(H_2)$ in envelopes with a fixed $R_{\rm out}$ but different $R_{\rm in}$. This is to study the influence of desorption temperatures ($T_{\rm desorb}$) on ice abundances of different species, as is shown in the right panels in Fig. 4.11. For the O-COMs that we focused on in the gas-to-ice comparison (i.e., CH_3OH, CH_3CHO, C_2H_5OH, CH_3OCH_3, and CH_3OCHO), $T_{\rm desorb}$ are similar to each other, about 130–

160 K under lab conditions (Fedoseev et al. 2015, 2022), and will be $\gtrsim$ 90 K under astrophysical conditions. For some larger COMs such as $(CH_2OH)_2$, T_{desorb} can be higher up to 200 K under laboratory conditions (equivalent to $\sim$ 150 K in space). Although the density is higher toward the protostar, the column densities of envelopes with different R_{in} can only vary by $\sim$ 40% assuming $T_{desorb} \sim$ 90–150 K for O-COMs.

In reality, each species have different T_{form} and T_{desorb} (R_{in} and R_{out}, equivalently). In the extreme case, species 1 in Fig. 4.10 has T_{form} = 60 K and T_{desorb} = 90 K, and species 2 has T_{form} = 10 K and T_{desorb} = 150 K, then the column density of species 2 traced by JWST will be a factor of 4.3 as that of species 1. For CH_3CHO and C_2H_5OH, the difference should be smaller given their similar T_{form} and T_{desorb}, in which case different spatial distributions is not able to fully explain the observed order-of-magnitude difference in the column density ratios with respect to CH_3OH between gas and ice. However, all the estimations were based on the assumption that the density distributions of COMs are proportional to that of H_2. This may not be true if the formation or evolution mechanisms of COMs are reacting differently to the H_2 density as well as the temperature. The details behind the discrepancy cannot be pinned down only by observations, but require more studies on the spatial distribution of COM ices in protostellar envelopes from the side of simulations and experiments.

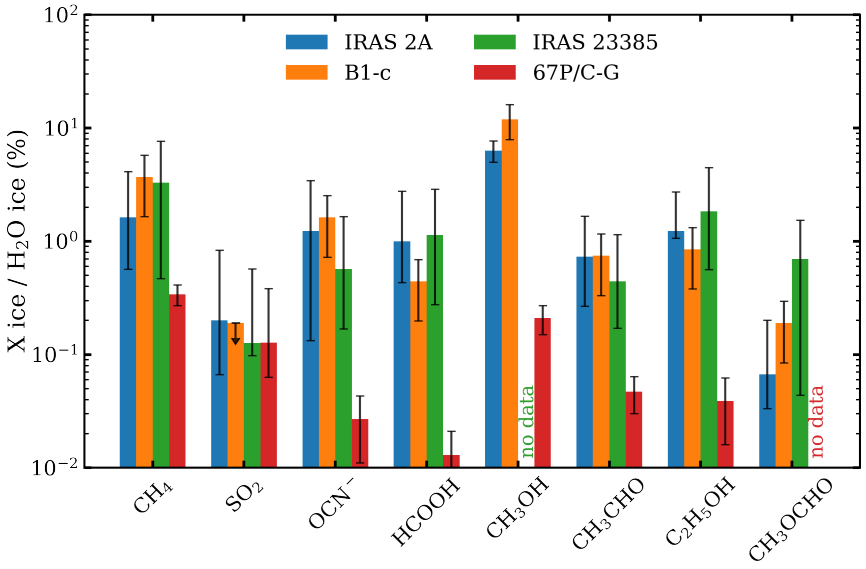


Figure 4.12: Ice column density ratios of seven species with respect to H_2O in three protostars (IRAS 23385, IRAS 2A, and B1-c; Rocha et al. 2024, this work) and one comet (67P/C-G; Rubin et al. 2019b). The ice column densities of CH_3OH for IRAS 23385 and CH_3OCHO for 67P/C-G are not reported in the references.

4.5.4 Ice ratios in protostars and comets

Additional to the COM ratios with respect to CH_3OH, we also compared the ice column density ratios of simple molecules and COMs with respect to H_2O among the

three protostellar sources (IRAS 2A, B1-c, and IRAS 23385; Rocha et al. 2024) and the comet 67P/C-G (Rubin et al. 2019b). We find results similar to those in Rocha et al. (2024) that the ice ratios of most species with respect to H_2O in protostars are similar and generally higher than those of the comet, except for SO_2 (Fig. 4.12). The COM ratios match better between protostars and the comet when they are calculated with respect to CH_3OH instead of H_2O (Fig. 4.9). This may suggest different evolutionary pathways between simple molecules and COMs, but more sources and studies are needed to draw robust conclusions.

4.6 Conclusions

We studied COMs and several simple molecules in both the gas and solid phases for two low-mass protostars, NGC 1333 IRAS 2A and B1-c. We derived the column densities in gas and ice by fitting the gas emission lines in the ALMA Band 7 spectra and the ice absorption bands in the JWST/MIRI-MRS spectrum, respectively. We made the first direct comparison in the column density ratios of four O-COMs with respect to CH_3OH between the gas and ice, which helps us gain a better understanding of the chemical evolution of these molecules in protostellar systems. Our main conclusions are:

1. We derived the gas-phase COM column densities for B1-c and IRAS 2A using a set of high-resolution ALMA Band 7 data, and the COM ratios with respect to CH_3OH are consistent with previous studies within a factor of two. A small difference may be introduced by the lower spatial resolution and sensitivity in previous observations.

2. Ices of COMs and other simple molecules were detected in the JWST/MIRI-MRS spectrum of B1-c, following the detection in NGC 1333 IRAS 2A and IRAS 23385+6053 (Rocha et al. 2024). We constrained the total abundance of the detected COM ices with respect to H_2O ice, $N_{\rm ice}({\rm COM})/N_{\rm ice}({\rm H_2O})$, as $\sim$15%. In particular, CH_3OH is the dominate species, with $N_{\rm ice}({\rm CH_3OH})/N_{\rm ice}({\rm H_2O}) \sim$ 12%. The upper limit of the most abundant N-COM, CH_3CN, is estimated as $N_{\rm ice}({\rm CH_3CN})/\ N_{\rm ice}({\rm H_2O}) < 1.2\%$.

3. By directly comparing the lab spectra of different COM ice mixtures and the observed JWST spectrum of B1-c, we find that most COM ices are likely present in a H_2O-rich environment. For some COMs such as CH_3OCH_3, CH_3OH-rich mixtures cannot be ruled out. CH_3OCHO is special case in that its H_2O-rich mixture matches the observations worse than the $CO{:}H_2CO{:}CH_3OH$ mixture, which implies a formation route of CO hydrogenation.

4. The temperature of COM ices cannot be well constrained by the observations, since the band profiles of COM ices are highly degenerate before crystallization, which usually occurs at around 100 K under laboratory conditions ($\sim$60–80 K in space). For mixed ices of some COMs such as C_2H_5OH, the degeneracy remains even after crystallization.

5. The fitting results of the COM fingerprint range (6.8–8.8 µm) using least-squares and MCMC for B1-c are similar to those reported in Rocha et al. (2024) for

IRAS 2A and IRAS 23385 using genetic algorithms. The derived column densities and uncertainties are consistent between B1-c and IRAS 2A. The only difference is that CH_3OCH_3 and CH_3COCH_3 are considered to be firmly detected in B1-c, but only tentatively detected in IRAS 2A. We expect to further confirm the detection of these COMs in a larger sample.

6. The comparison of COM ratios with respect to CH_3OH between gas and ice shows two cases: CH_3OCHO and CH_3OCH_3 have consistent ratios in both phases, while the ice ratios of CH_3CHO and C_2H_5OH are higher than the gas ones by about one order of magnitude.

7. The consistency in COM ratios between gas and ice suggests a direct inheritance from ice to gas, and probably a common formation history for CH_3OCH_3, CH_3OCHO, and CH_3OH ices. On the other hand, the difference in COM ratios between gas and ice implies the participation of gas-phase chemistry after ices sublimate in hot cores. Another possible but less dominant cause is the different spatial distributions of COM ices in protostellar envelopes, which are related to the formation time and desorption temperatures of each species.

Thanks to the successful operation of JWST and the high-quality mid-IR spectra it provides, it is now becoming feasible to have robust detections and quantitative analyses of COMs in the solid phase, allowing direct comparison to the gas-phase counterparts that have been intensively studied by ALMA and other telescopes. As JWST opens a new door to the ice world, it is both intriguing and important to connect it with the already well-explored gas world. The pilot study of two famous hot corinos has shed some light on how COMs evolve through the transition from ice to gas, from the cold envelope to the hot core region. We look forward to further verifying our hypotheses in a larger sample, and additional interpretations from simulation or laboratory studies are highly welcome.

Acknowledgements

This paper makes use of the following ALMA data: ADS/JAO.ALMA#2021.1.01578.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. This work is also based on observations made with the NASA/ESA/CSA James Webb Space Telescope. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. These observations are associated with program #1290. Astrochemistry in Leiden is supported by the Netherlands Research School for Astronomy (NOVA), by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 101019751 MOLDISK), by the Dutch Research Council (NWO) grants TOP-1 614.001.751 and 618.000.001, and by the Danish National Research Foundation through the Center of Excellence "InterCat" (grant agreement no.: DNRF150). The work of M.E.R. was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National

Aeronautics and Space Administration (80NM0018D0004). H.B. acknowledges support from the Deutsche Forschungsgemeinschaft in the Collaborative Research Center (SFB 881) "The Milky Way System" (subproject B1). P.J.K. acknowledges support from the Science Foundation Ireland/Irish Research Council Pathway programme under Grant Number 21/PATH-S/9360. L.M. acknowledges the financial support from DAE and DST-SERB research grants (SRG/2021/002116 and MTR/2021/000864) of the Government of India. T.H. acknowledges support from the ERC grant no. 832428 Origins.

Appendix

4.A Images of ALMA and JWST observations

Figures 4.A.1–4.A.2 display the integrated intensity maps (i.e., moment 0 maps) of 12 COM lines in IRAS 2A and B1-c. Five lines are from CH_3OH with different upper energy levels (E_{up}) and Einstein A coefficients (A_{ij}). For the other seven O-COMs, the strongest unblended line of each species was selected to display. Figure 4.A.3 shows the continuum images in 12 MRS channels along with the extracted spectrum.

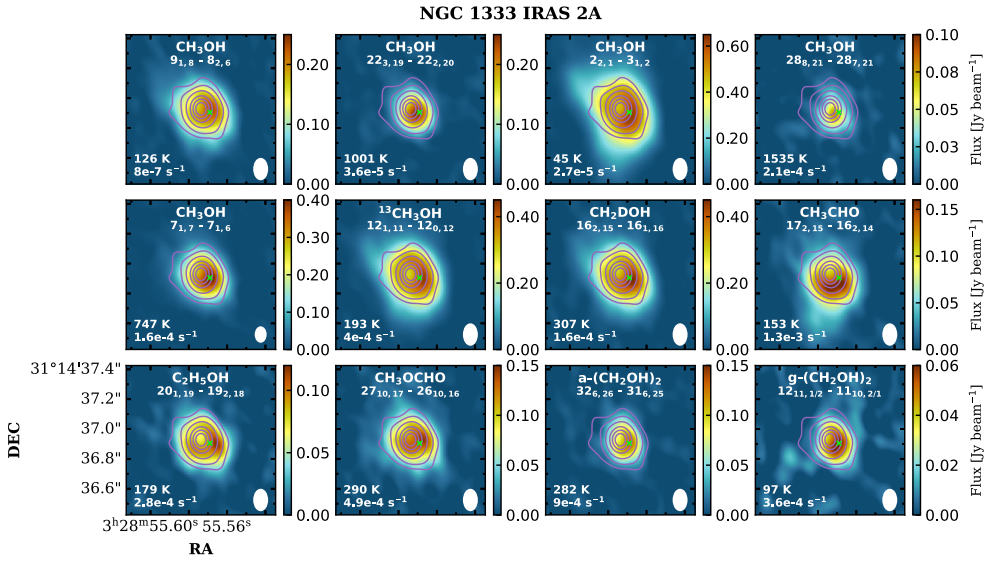


Figure 4.A.1: Integrated intensity maps (i.e., moment 0 maps) of 12 emission lines of eight O-COMs in NGC 1333 IRAS 2A. The quantum numbers, E_{up}, A_{ij} of each transition are labeled in white text. The contours in light purple indicate the continuum emission at the 30, 50, 100, 200, 300, 400, and 500σ levels (σ = 0.2 mJy beam^{-1}). The beam size ($0.1'' \times 0.15''$) is denoted by white ellipses in the lower right, and the pixel where the spectrum was extracted is marked by a green cross in each panel. The beam size in the first panel of the second row is slightly smaller ($0.08'' \times 0.11''$), since this transition lies in a different spectral window than the others.

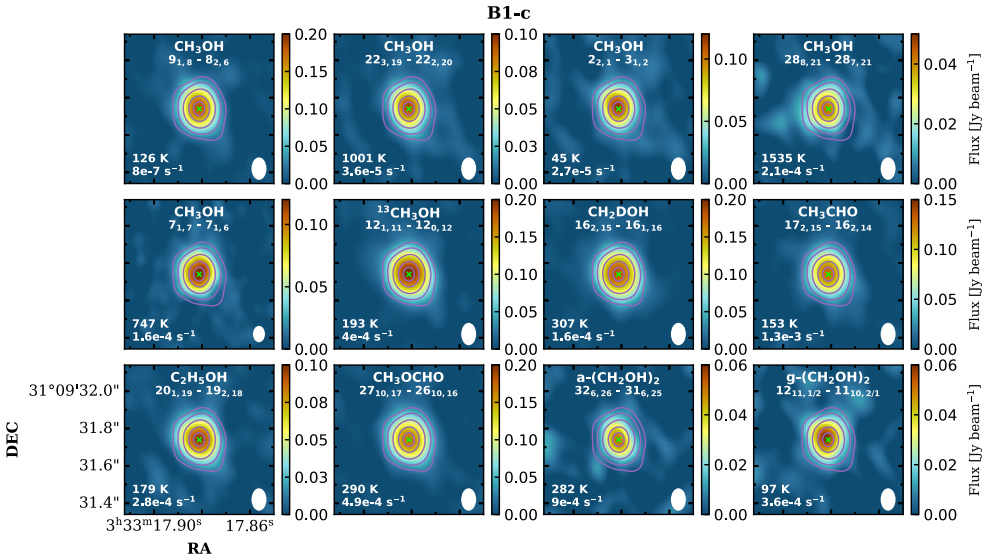


Figure 4.A.2: Same as Fig. 4.A.1 but for B1-c. The continuum contours are set at 3, 5, 10, 20, 30, and 40σ levels (σ = 2 mJy beam^{-1}).

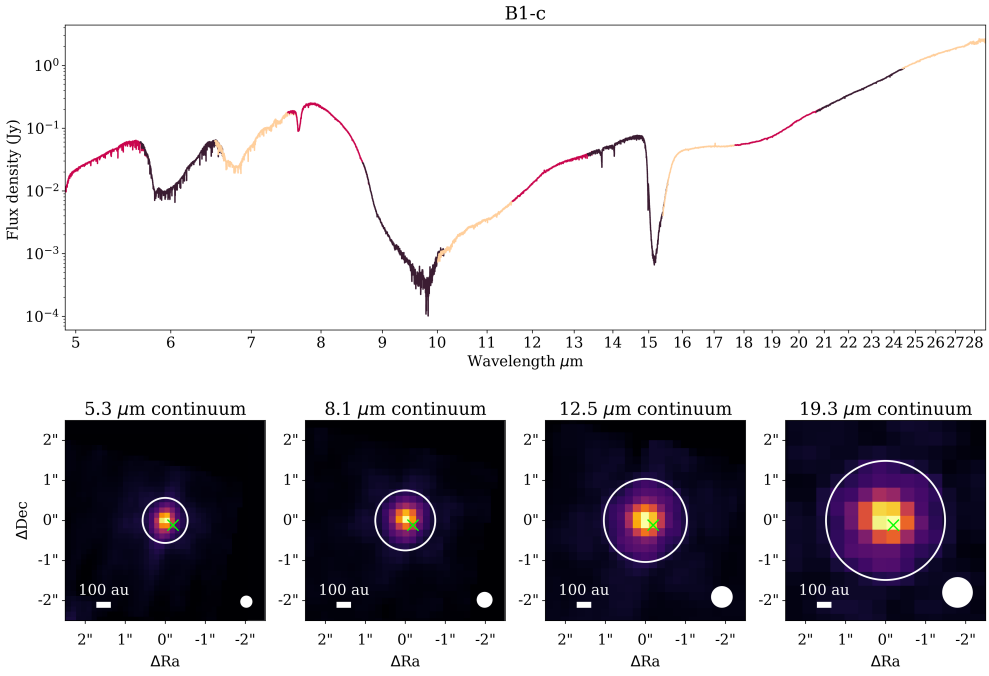


Figure 4.A.3: Extracted spectrum (top) and the continuum images (bottom) of JWST/MIRI-MRS observations of B1-c. Top panel: JWST/MIRI-MRS spectrum of B1-c extracted from the reduced datacubes. The spectrum slices of three sub-bands in each of the four MIRI-MRS Channels are distinguished in three different colors. Bottom panels: continuum images in four Channels (4.9–7.65 µm, 7.51–11.7 µm, 11.55–17.98 µm, and 17.7–27.9 µm). The sizes of the PSF and the aperture for spectrum extraction increase as a function of wavelength, and are denoted by white circles, respectively. The extraction position of the ALMA spectrum (also the continuum peak) in Fig. 4.A.2 is marked here using the same green cross. The ALMA extraction position is well within the MIRI apertures, and the small offset to the aperture center is likely due to the lower accuracy of the JWST pointing.

4.B The silicate features

All the silicate spectra used to fit the observed spectrum of B1-c in Sect. 4.3.2.2 have no features shortward of 7.6 µm. The GCS 3 silicate spectrum shown in Fig. 4.3b is a smoothed version of the original spectrum, in which the emission and absorption lines that are not attributed to silicates were removed (see the left panel of Fig. 4.B.1).

When fitting the silicate features, we primarily focused on whether the profiles of the 9.8 and 18 µm bands and their optical depth ratios match between the observations and the silicate spectra used for fitting. Figure 4.B.1 shows that the optical depth ratio between the two bands in the GCS 3 profile is ∼3.3, larger than those of the computed spectra of olivine and pyroxene ($\lesssim 2$). However, the relative intensity of the 18 µm band in the pyroxene spectrum can be reduced by increasing the mass fraction of carbon. The computed spectrum of pyroxene is by default made up of $Mg_{0.7}Fe_{0.3}SiO_3$ and mixed with carbon in a mass fraction of 87% and 13%, respectively. We finally used the spectrum of 80% pyroxene mixed with 20% carbon to get a better fit for the 18 µm band. The 18 µm band in the olivine spectrum can also be widened by increasing the grain size (right panel of Fig. 4.B.1). In Sect. 4.3.2.2 we only came up with a plausible fitting of the silicate features, but potentially there is a huge parameter space to explore.

Another caveat is that even small difference in the global continuum fitting (Sect. 4.3.2.1) can change the spectral profile at long wavelengths, which makes it degenerate to fit the global continuum and the 18 µm silicate band. Fortunately, our analysis on the COMs fingerprints between 6.8 and 8.8 µm is little affected by the fitting of the 18 µm band, and we are safe to proceed with a plausible fitting shown in Fig. 4.3b.

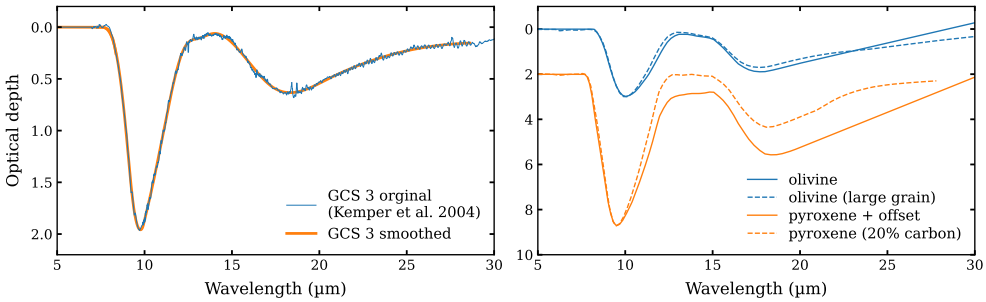


Figure 4.B.1: Silicate spectra between 5 and 30 µm. Left panel: original and smoothed silicate spectra of GCS 3. Right panel: computed spectra of olivine (blue) and pyroxene (orange) with increased grain size or mass fraction of carbon (dashed).

4.C Selection of ice mixtures

Figure 4.C.1 is the same as Fig. 4.5 but shows the temperature comparison for the lab spectra of CH_3CHO, CH_3OCH_3, and CH_3COCH_3. Same as C_2H_5OH and CH_3OCHO, the band of pure ices will become significantly sharper after crystallization. For the mixed ices, the bands of CH_3CHO:H_2O mixture become narrower as temperature increases, but the peak positions remain similar and cannot be ruled out by the observations. The major change of CH_3OCH_3:H_2O mixture is the peak position of the

8.63 µm band. The band is blueshifted as temperature increases, and anti-correlates with the observations when $T > 120$ K. For CH_3COCH_3:H_2O mixture, as temperature increases, the two band at 7.3 and 8.03 µm are redshifted, and the latter splits into two bands. The anti-correlation occurs at 7.3 µm when T >70–90 K.

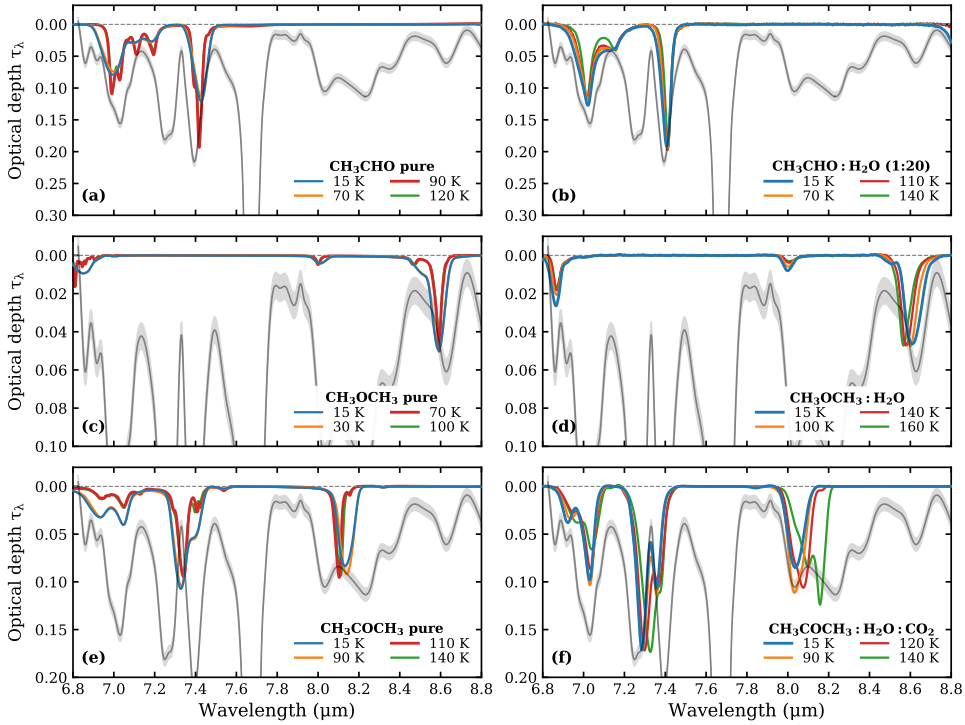


Figure 4.C.1: Same as Fig. 4.5 but for CH_3CHO, CH_3OCH_3, and CH_3COCH_3.

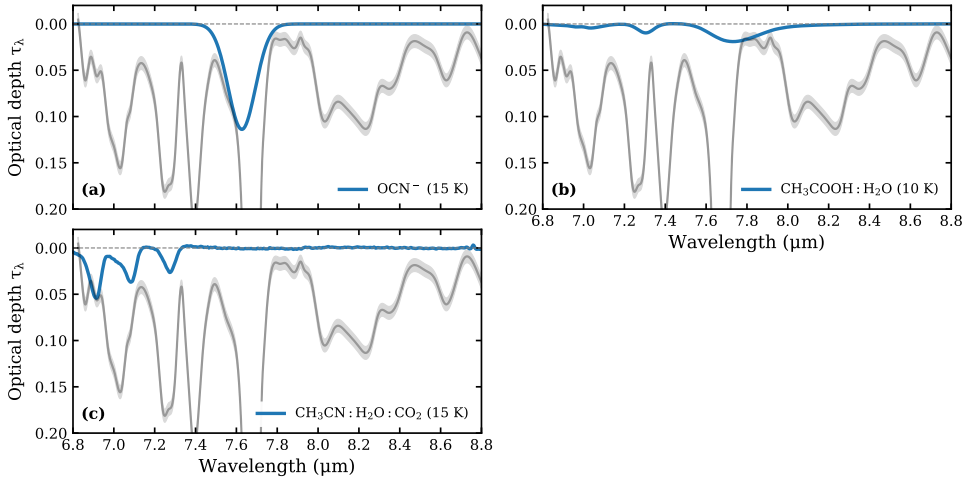


Figure 4.C.2: Same as Fig. 4.4 but for OCN^-, CH_3COOH, and CH_3CN.

4.D Baseline correction of lab spectra

The lab spectra of COMs mixed with H_2O or CH_3OH require baseline correction to isolate the weak COM bands from the strong H_2O band at 6 µm or the CH_3OH band at 6.75 µm. A polynomial was fit to wavelength regions that are free of absorption bands. The baseline correction was only applied between ∼6.0 and 10.0 µm, considering that we only fit the observation between 6.8 and 8.8 µm.

For H_2O-rich mixtures, the baseline is just the red wing of the H_2O bending mode at 6.0 µm, and it is easy to select and fit a polynomial to those ranges that do not have absorption features of the interested species, as is shown in the left panels of Fig. 4.D.1.

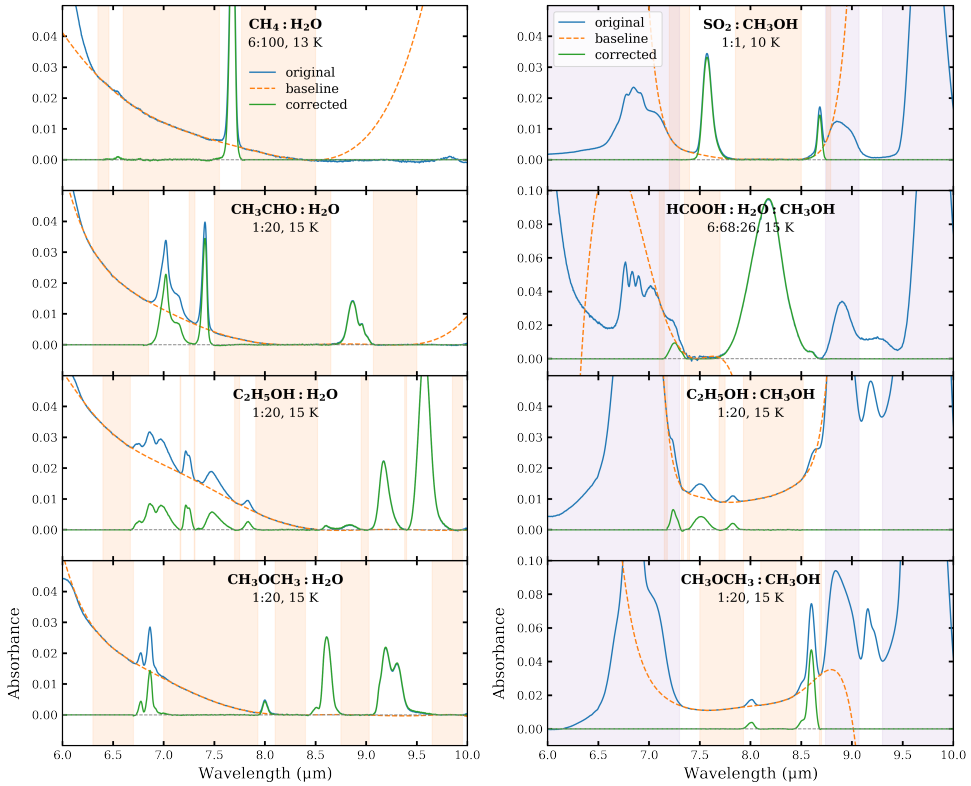


Figure 4.D.1: Baseline correction to lab spectra of ice mixtures. In the left panels, we take the H_2O-rich mixtures of CH_4 and three O-COMs (CH_3CHO, C_2H_5OH, CH_3OCH_3) as examples. The original and the corrected spectra are plotted in solid blue and green, respectively. The polynomial is plotted in dashed orange, and the wavelength ranges that were selected to fit the polynomial are highlighted in orange panels. The right panels are similar to the left ones but for CH_3OH mixtures. SO_2, HCOOH, C_2H_5OH, and CH_3OCH_3 are taken as examples. The purple-shaded regions indicate the frequency ranges of CH_3OH bands.

However, it is usually not straightforward to do baseline correction for CH_3OH-rich mixtures, especially in 6.5–7.3 µm where there is a CH_3OH band with wavy features

(see right panels of Fig. 4.D.1; the feature shown in the top panel are CH_3OH-only). If the interested species has no overlapping band with CH_3OH (e.g., SO_2), the routine is the same as for H_2O-rich mixtures, except that a higher order polynomial may be needed. Unfortunately, most O-COMs have weak bands around 7.0 µm, and it is difficult to isolate those features from the uneven CH_3OH features. We tried to reserve the weak bands between 7.1 and 7.3 µm where the CH_3OH spectrum is still smooth and it is feasible to trace a local continuum; this was done for HCOOH and C_2H_5OH which have a weak band at ∼7.2 µm. For those bands shortward of 7.1 µm, we discard them in the corrected spectra (e.g., for C_2H_5OH, CH_3OCH_3, and CH_3OCHO).

4.E Fitting results of the ALMA spectra of IRAS 2A and B1-c

Figures 4.E.1–4.E.2 show the full ALMA spectra of IRAS 2A and B1-c and the best-fit LTE models of the detected COMs, respectively. Important strong lines of simple molecules are labeled in gray text.

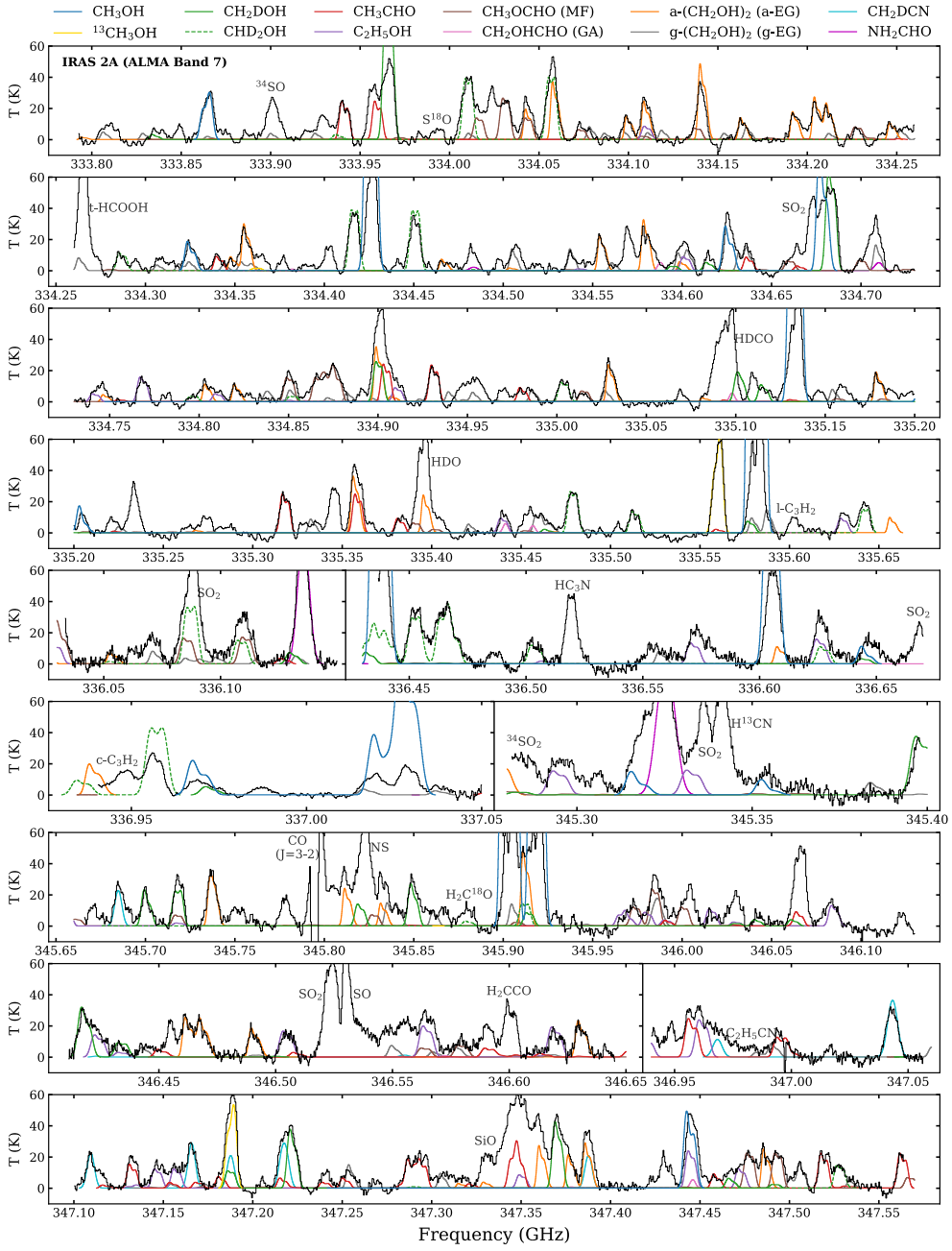


Figure 4.E.1: Best-fit LTE models overlaid on the observed ALMA spectrum of IRAS 2A.

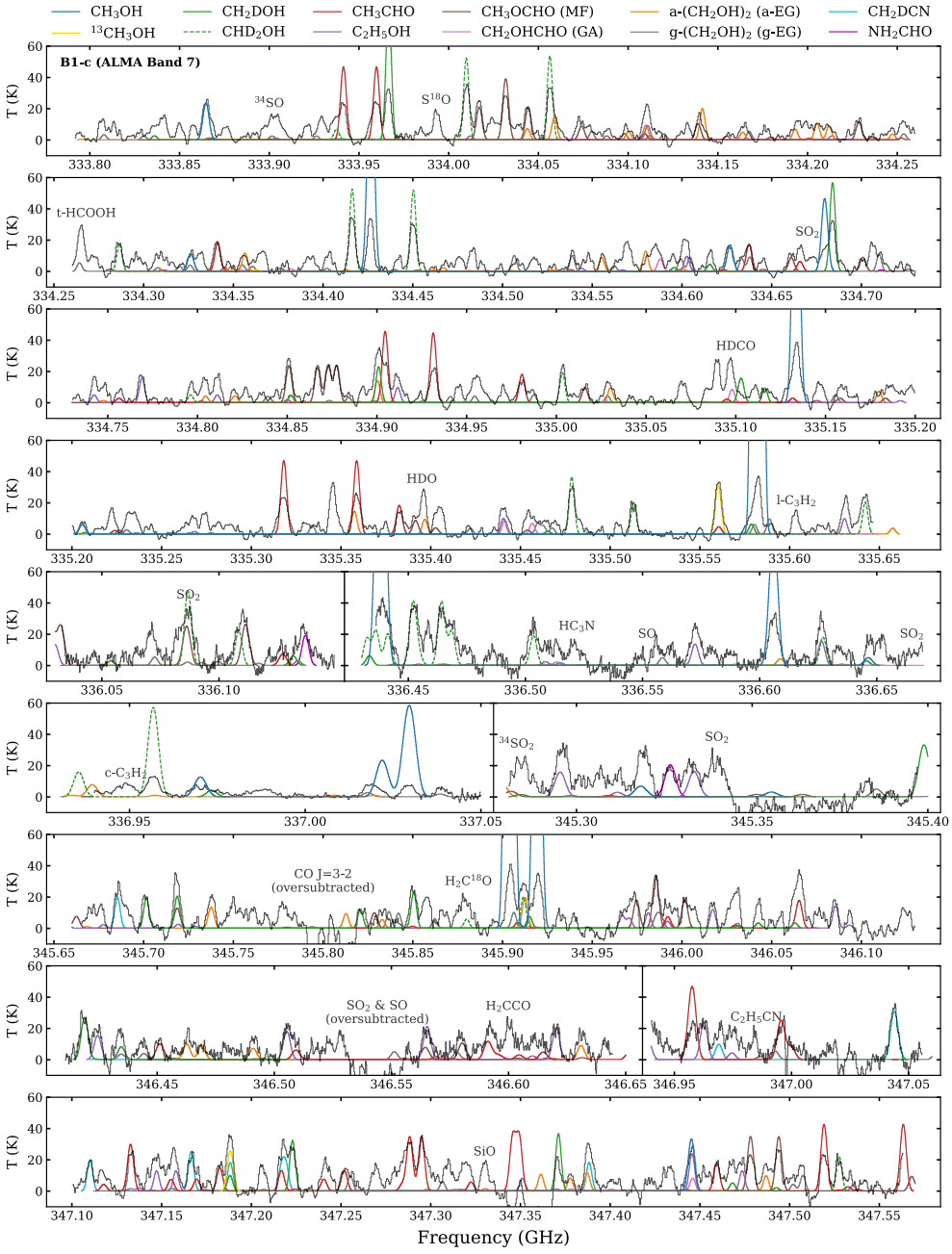


Figure 4.E.2: Same as Fig. 4.E.1 but for B1-c.

4.F H_2CO bands in the CH_3OCHO mixture

As was mentioned in Sect. 4.4.2.1 and shown in Fig. 4.4h, the lab spectrum of CH_3OCHO mixed with CO, H_2CO, and CH_3OH in a mixing ratio of 1:20:20:20 (hereafter CO-rich mixture) fits the observations best. However, the band at 8µm does not appear in the pure ice and other ice mixtures with H_2O and CH_3OH, suggesting that this absorption feature comes from neither CH_3OCHO nor CH_3OH. In fact, this 8µm band belongs to H_2CO. Figure 4.F.1 compares the lab spectra of pure H_2CO, pure CH_3OCHO, and the CO-rich mixture of CH_3OCHO under 15 K and 80 K. It is shown that the 8 µm band in the CO-rich mixture remains almost the same as in the pure H_2CO, which implies that this band is not affected by ice mixing, and therefore can be treated separately from the CH_3OCHO band at 8.25 µm.

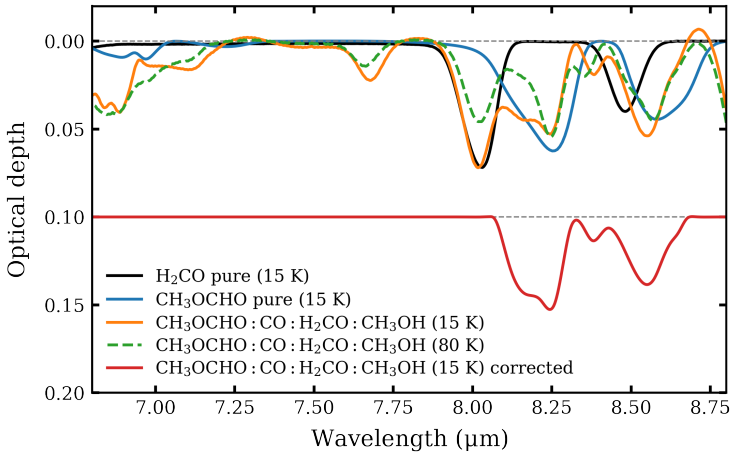


Figure 4.F.1: Comparison between the lab spectra of pure H_2CO ice (black), pure CH_3OCHO ice (blue), and CH_3OCHO:CO:H_2CO:CH_3OH mixture under 15 K (orange) and 80 K (dashed green). The corrected spectrum for the CH_3OCHO mixture under 15 K (red) is plotted at an offset place.

The problem of using the original lab spectrum of the CH_3OCHO:CO:H_2CO:CH_3OH mixture is that the mixing ratio between CH_3OCHO and H_2CO is fixed at 1:20, and therefore the intensity ratio between the CH_3OCHO band at 8.25 µm and the H_2CO band at 8.02 µm is fixed as well. The lab spectra of the CH_3OCHO:CO:H_2CO:CH_3OH mixture at 80 K also shows that the intensity ratio between the 8.25 µm and 8.02 µm bands would increase as part of H_2CO has sublimated from the ice mixture at 80 K. However, the 1:20 ratio between H_2CO and CH_3OCHO may not be realistic; it was found in test fittings that the 8.02 µm band would be overfit and the 8.25 µm band would be underfit the observations if we used the original lab spectrum. As a result, we wanted to break the correlation between H_2O and CH_3OCHO in the CO-rich mixture to achieve a better fit at both bands.

To address this problem, we fit a Gaussian to the H_2CO band at 8.02 µm and subtracted it from the lab spectrum of the CH_3OCHO:CO:H_2CO:CH_3OH mixture, in order to isolate the CH_3OCHO band from the H_2CO band. There is also a weaker

band of H_2CO at 8.48 µm, which splits the 8.59 µm band in pure CH_3OCHO into two peaks at 8.38 and 8.55 µm in the CO-rich mixture. The presence of H_2CO also increase the relative intensity between the bands at 8.55 µm and 8.25 µm from 0.73 (in pure CH_3OCHO) to 1 (in CO-rich mixture). To further separate the influence by H_2CO, we scaled down the relative intensity of the 8.55 µm band back to 0.73. Finally, we removed other weak feature shortward of 7.8 µm that are either absent in the pure CH_3OCHO spectrum or blended with CH_3OH features at ∼6.9 µm (considering that the ratio between CH_3OCHO and CH_3OH is 1:20, the contribution from CH_3OCHO is negligible). The corrected lab spectrum of the CH_3OCHO:CO:H_2CO:CH_3OH mixture is shown in Fig. 4.F.1, and was the one that used in the overall fitting.

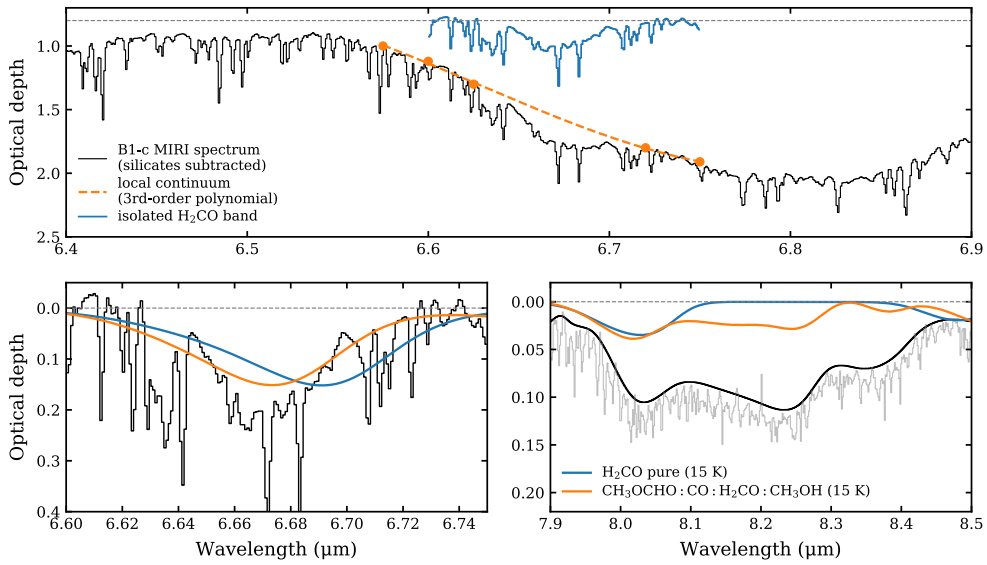


Figure 4.F.2: Isolation of the CH_3OCHO bands from the H_2CO band in the lab spectrum of the CH_3OCHO:CO:H_2CO:CH_3OH mixture. Top panel: H_2CO band at 6.67 µm (blue) isolated by tracing a local continuum (orange) to the observed spectrum (black); superposed gas-phase absorption lines of H_2O are not removed here. Bottom-left panel: lab spectra of pure H_2CO ice (blue) and the CH_3OCHO:CO:H_2CO:CH_3OH mixture (orange) manually scaled to the isolated H2CO band at 6.67 µm in observations (black). Bottom-right panel: same as the middle panel but for a different wavelength range (7.9–8.5 µm).

On the other hand, the 8.03 µm band observed in the B1-c spectrum has contribution from not only H_2CO, but also H_2O-rich mixtures of CH_3COCH_3 (Sect. 4.4.2.1 and panels i–j in Fig. 4.4). To determine how much H_2CO accounts for this band, we turned into another characteristic band of H_2CO at 6.67 µm, which is outside of the COM fingerprint range. We first fit a local continuum following a similar routine to Sect. 4.3.2.3 to isolated this band (but without removing the gas-phase absorption lines), and then manually scaled the lab spectra of pure H_2CO and the CH_3OCHO:CO:H_2CO:CH_3OH mixture to the observations (see Fig 4.F.2). The observed H_2CO band at 6.67 µm is about 0.02 µm blueshifted from pure H_2CO, but matches the CH_3OCHO:CO:H_2CO:CH_3OH mixture well (where CO:H_2CO:CH_3OH = 1:1:1), which implies that H_2CO is likely to have surroundings rich in CO and

CH_3OH in interstellar ices. The two H_2CO-relevant lab spectra were scaled to have the same intensities at 6.67 µm. The scaling factor of pure H_2CO was then determined by visual inspection, and was fixed in the overall fitting. The ice column density of H_2CO was calculated from the 8.02 µm band.

4.G Temperature analysis of COM ices

Table 4.G.1 summarizes the temperature-relevant information of the lab spectra of the five O-COMs that are considered detected in the JWST spectrum of B1-c.

4.H Band strengths

The band strength A describes how strong is the absorbance of a specific band given a unit of ice column density (cm molecule^{-1}). In Table 4.2, Col. 3 lists the characteristic absorption bands of the species (Col. 1), and Cols. 3–4 provide values and references of the corresponding band strengths. Most of these values and references are the same as those given in Rocha et al. (2024), except for the following species:

- CH_3OH: for the 9.74 µm band, we adopted $A = 1.56 \times 10^{-17}$ cm molec^{-1} from Luna et al. (2018), instead of 1.8×10^{-17} cm molec^{-1} from Bouilloud et al. (2015), considering that the former one is more updated.
- CH_3OCH_3: the band strength of the 8.59 µm band is adopted from Table 1 in Terwisscha van Scheltinga et al. (2018), and A is corrected for the H_2O mixture by a factor of ∼0.5. This factor, known as the relative band strength, is not directly provided in Terwisscha van Scheltinga et al. (2018) but can only be read from the lower right panel of Fig. C.11 therein. Rocha et al. (2024) considered it as 0.5, and the corrected A is 4.9×10^{-17} cm molec^{-1}. However, a more accurate estimation of this factor would be 0.56–0.57, yielding a slightly larger A of 5.55×10^{-17} cm molec^{-1}.
- CH_3COCH_3: the band strength of the 7.33 µm band is originally measured as 1.39×10^{-17} cm molec^{-1} by Hudson et al. (2018). The relative A of the H_2O mixture is not directly stated, but plotted in the lower right panel of Fig. A.11 in Rachid et al. (2020). The estimated relative A is 0.73, and the corrected A is 1.02×10^{-17} cm molec^{-1}, a bit smaller than 1.2×10^{-17} cm molec^{-1} used by Rocha et al. (2024)[4]

4.I Statistics of fitting the JWST/MIRI-MRS spectrum of B1-c

Table 4.I.1 lists the statistics of fitting the JWST/MIRI-MRS spectrum of B1-c, including the best-fit values and uncertainties of the scaling factors of the selected lab spectra.

[4]In Table 1 of Rocha et al. (2024), the identification of the 7.33 µm band of CH_3COCH_3 is mistakenly noted as the CCC asymmetric stretching mode, which is actually the band at 8.03 µm.

Table 4.G.1: Inferred range of crystallization temperature of COM ices under laboratory conditions ($T_{\mathrm{crystal,\,lab}}$). All the considered lab spectra of the five COM ices are obtained from the LIDA database (Rocha et al. 2022).

	pure ices		mixed ices [a]		
Species	measured T_{lab} (K) [b]	$T_{\mathrm{crystal,\,lab}}$ (K)	measured T_{lab} (K) [b]	$T_{\mathrm{crystal,\,lab}}$ (K)	fit B1-c (K) [c]
CH_3CHO	15, 30, 70, 90, 110, 120	70–90	15, 30, 70, 90, 110, 120, 140, 160	110–140	all [d]
C_2H_5OH	15, 30, 70, 100, 120, 130, 140, 150	100–120	15, 30, 70, 100, 120, 130, 140, 150, 160	100–160	all [d]
CH_3OCH_3	15, 30, 70, 90, 100	30 [e]–70	15, 30, 70, 90, 100, 120, 140, 160	100–140	$<$ 100–120
CH_3OCHO	15, 30, 50, 80, 100, 120	80–100	15, 30, 50, 80, 100, 120	50–80 [f]	$<$ 100
CH_3COCH_3	15, 30, 70, 90, 110, 120, 130, 140	90–110	15, 30, 70, 90, 100, 110, 120, 140, 160	100–140	$<$ 90

[a] The mixing constituents are provided in the legends of Figs. 4.5 and 4.C.1.
[b] In experiments, temperature is usually sampled at small intervals (e.g., 1 K), but only a small part of the spectra measured at certain temperatures are provided on databases.
[c] The range of T_{lab} in which the lab spectra can fit into the B1-c spectrum, as long as there is no anti-correlation.
[d] The changes in the band profiles around crystallization are small; therefore, it is hard to exclude the crystalline features when compared with observations.
[e] The lower limit of $T_{\mathrm{crystal,\,lab}}$ of pure CH_3OCH_3 ice is likely higher than 30 K.
[f] For CH_3OCHO, the $T_{\mathrm{crystal,\,lab}}$ of mixed ice is lower than that of pure ice. This is likely due to the crystallization or desorption of CO and H_2CO, not the crystallization of CH_3OCHO itself.

Table 4.I.1: Fitting statistics of the scaling factors[†] of lab spectra. The best-fit values, standard errors (σ_a) and relative errors (a/σ_a) are listed for least-squares and MCMC fittings. All the 12 candidate species were considered in the least-squares fitting, and nine of them were included in the MCMC fitting.

Ice mixture				**Least-squares**			**MCMC**		
Constituents	Ratio	T (K)	Fit range[a]	Best-fit a	σ_a	a/σ_a (%)	Best-fit a	σ	a/σ_a (%)
CH_4:H_2O	1:100	13	0–5	3.61	0.03	0.8	3.66	0.27	7.3
SO_2:CH_3OH	1:1	10	0–1.0	4.2×10^{-7}	0.56	1.3×10^{8}	–	–	–
OCN^-	–	15	0–8.0	6.58	0.34	5.2	6.49	0.45	6.9
$HCOO^-$	–	14	0-0.08	0.04	0.001	2.4	0.037	0.002	4.3
HCOOH:H_2O:CH_3OH	6:68:26	15	0–1.0	0.64	0.03	4.0	0.61	0.05	7.9
CH_3CHO:H_2O	1:20	15	0–2.5	1.65	0.04	2.2	1.66	0.13	7.6
C_2H_5OH:H_2O	1:20	15	0–5.0	1.35	0.10	7.7	1.98	0.14	7.5
CH_3OCH_3:H_2O	1:20	15	0–0.8	0.40	0.05	12.9	0.38	0.07	19.3
CH_3OCHO:CO:H_2CO:CH_3OH	1:20:20:20	15	0–2.0	0.94	0.06	6.0	0.98	0.09	9.2
CH_3COOH:H_2O	1:20	10	0–0.2	0.08	0.02	30.7	–	–	–
CH_3COCH_3:H_2O	1:20	15	0–2.0	0.99	0.03	3.4	1.13	0.08	6.8
CH_3CN:H_2O:CO_2	1:5:2	15	0–1.5	0.70	0.08	10.9	–	–	–
H_2CO	–	10	0.8[b]	0.8[b]	–	–	0.8[b]	–	–

[†] The scaling factors refer to a_i in Eq. (4.3), which serve as the variables in the least-squares and MCMC fittings. They are proportional to the ice column densities (Eq. 4.4).

[a] The fit range of scaling factors were determined by manually comparing the lab spectra to the observations, as is shown in Figs. 4.4, 4.5, and 4.C.1.

[b] The scaling factor of the lab spectrum of H_2CO was determined by fitting to an H_2CO band at 6.67 µm, outside of the 6.8–8.8 µm range. The H_2CO band at 8.02 µm was scaled with the same factor of 0.8, which was fixed in both the least-squares and MCMC fittings.

4.J The CH_3OH band at 3.9 µm

To better quantify the ice column density of CH_3OH, we checked the 3.9 µm CH_3OH band falling in the wavelength coverage of NIRSpec. The NIRSpec integral field unit (IFU) observations of B1-c were taken as part of GTO 1290 in a four-point dither pattern. Details on the reduction of NIRSpec data and the spectra extraction will are not the focus of this work and will be introduced in future publication.

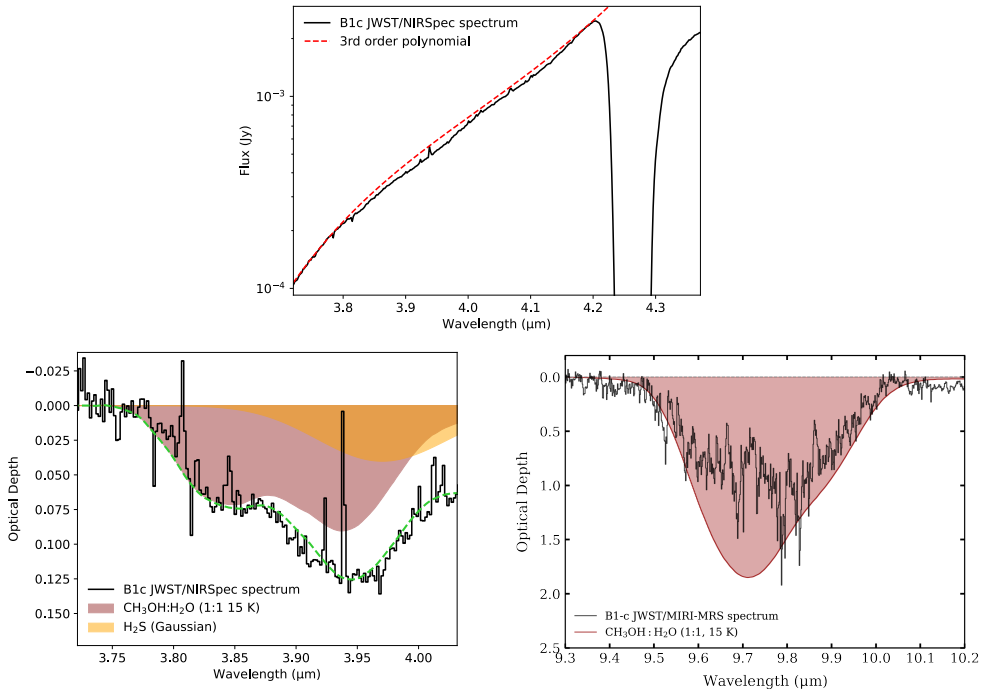


Figure 4.J.1: Fitting of the CH_3OH bands at 3.94 and 9.74 µm. Top panel: the NIRSpec IFU spectrum of B1-c (solid black) and the local continuum of the 3.94µm CH_3OH band fit by a 3rd-order polynomial (dashed red). Bottom-left panel: best-fit results of the observed absorption complex between 3.75 and 4.03 µm (black, the same as in the top panel but in optical depth scale). The shaded regions in red and yellow correspond to the CH_3OH and H_2S components, respectively. The overall best-fit spectrum is shown in dashed green. Bottom-right panel: the observed 9.74 µm band in the MIRI-MRS spectrum (black) overlaid with the CH_3OH:H_2O lab spectrum (red) scaling with the same factor as in the middle panel.

As is shown in the top panel of Fig. 4.J.1, the 3.9 µm band is next to the strong CO_2 band at 4.27 µm, of which the blue wing tends to be warped by grain scattering and may need grain size correction (Dartois et al. 2022, 2024). We tried fitting a 3rd-order polynomial to the observed NIRSpec spectrum in 3.72–3.8 µm and 4.17–4.2 µm. This local continuum is only considered as a preliminary analysis without a proper grain shape correction, and it is deviating from the observations longward of 4.2 µm.

The 7.33 µm band corresponds to the CH_3 symmetric stretching mode.

The observed spectrum between 3.72 and 4.2 µm was converted to optical depth scale by taking a logarithm with the local continuum.

The absorption features in this range were attributed to CH_3OH, H_2S, and HDO (Slavicinska et al. 2024, the HDO feature will be discussed in a future paper). The CH_3OH band was fit by a lab spectrum of CH_3OH:H_2O (1:1) ice mixture, recently measured by Slavicinska et al. (2024). The contribution of H_2S is represented by a Gaussian function. We performed a least-squares fitting on the scaling factor of the CH_3OH:H_2O spectrum and the coefficients of the Gaussian function (best-fit results displayed in the bottom-left panel of Fig. 4.J.1). The ice column density of CH_3OH was then calculated using Eq. (4.4), where a band strength of 1.56×10^{17} cm molec^{-1} (Luna et al. 2018) was taken. The best-fit scaling factor of the 3.9 µm CH_3OH band in the observed NIRSpec spectrum yields an ice column density of 4.0×10^{18} cm^{-2} for CH_3OH. We scaled the CH_3OH:H_2O lab spectrum using the same factor as in the middle panel and compared it with the observed MIRI-MRS spectrum at 9.74 µm (the bottom-right panel of Fig. 4.J.1). The lab spectrum overestimates the observed 9.74 µm band, but within a reasonable extent considering the extinction by the silicate band at 9.8 µm.

As was mentioned in Sect. 4.4.2.5, we finally adopted a CH_3OH ice column density of 3.0×10^{18} cm^{-2}, corresponding to 1.5 times the best-fit Gaussian to the observed 9.74 µm band (top panel in Fig. 4.8). This result is based on the Gaussian fitting to the 9.74 µm band, with silicate extinction taken into account. The difference between the ice column densities derived from fitting the 3.9 µm and the 9.74 µm bands is small (∼30%), suggesting that both the estimations are reasonable. However, considering that the analysis of the NIRSpec spectrum was only preliminary, the result of 4.0×10^{18} cm^{-2} was not adopted, instead it was used as a reference to constrain the uncertainty of our fitting results to the 9.74 µm band (σ=1.0×10^{18} cm^{-2}).

4.K Additional figures for the JWST fitting results

Here we provide supplementary figures for the overall fittings of the COM fingerprint range (6.8–8.8 µm) in the JWST/MIRI-MRS spectrum of B1-c. Figure 4.K.1 shows the best-fit lab spectra as in Fig. 4.7 but in a step-by-step form. Figure 4.K.2 shows the corner plot of the MCMC fitting introduced in Sect. 4.4.2.4.

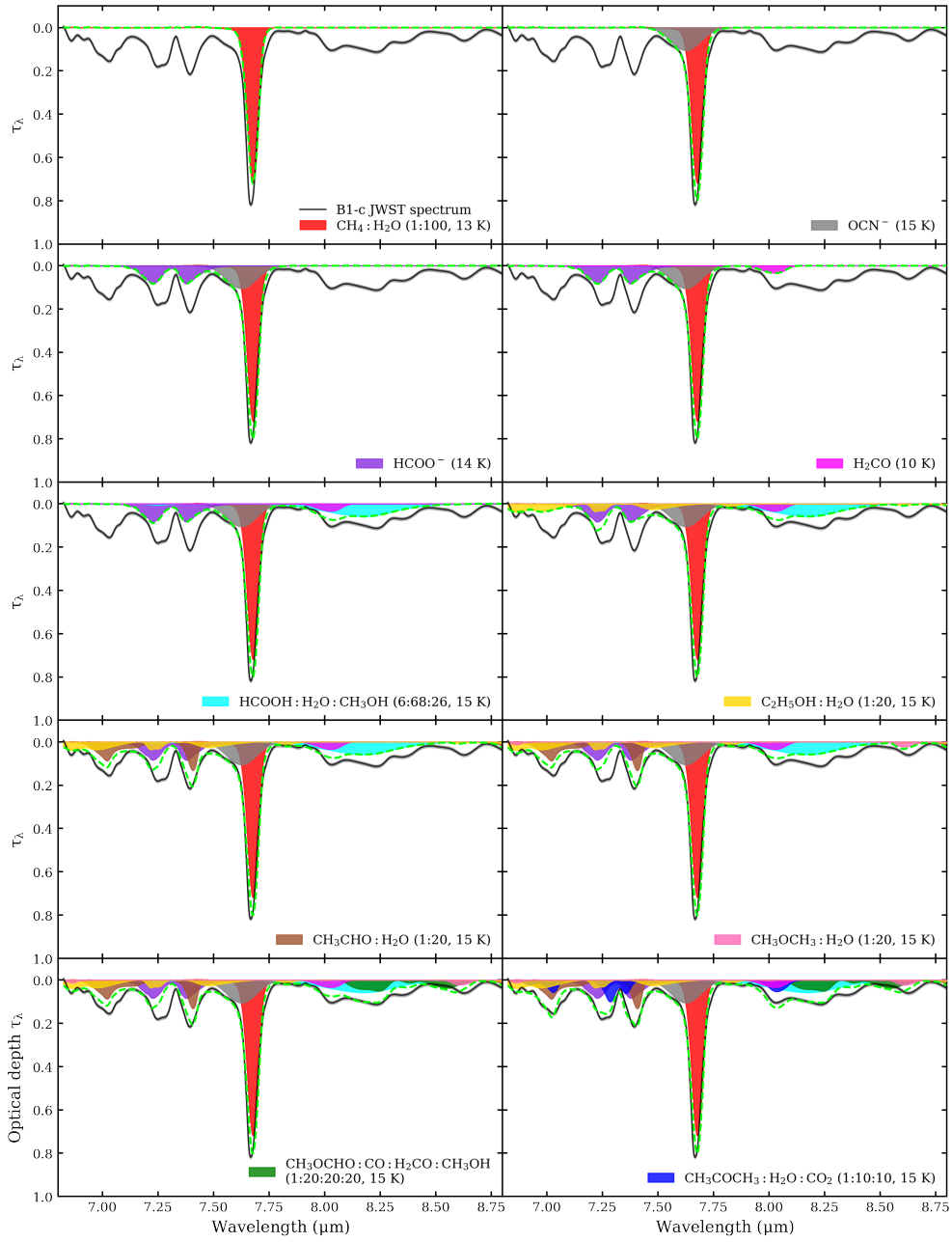


Figure 4.K.1: Same as Fig. 4.7 but displays the best fits species by species.

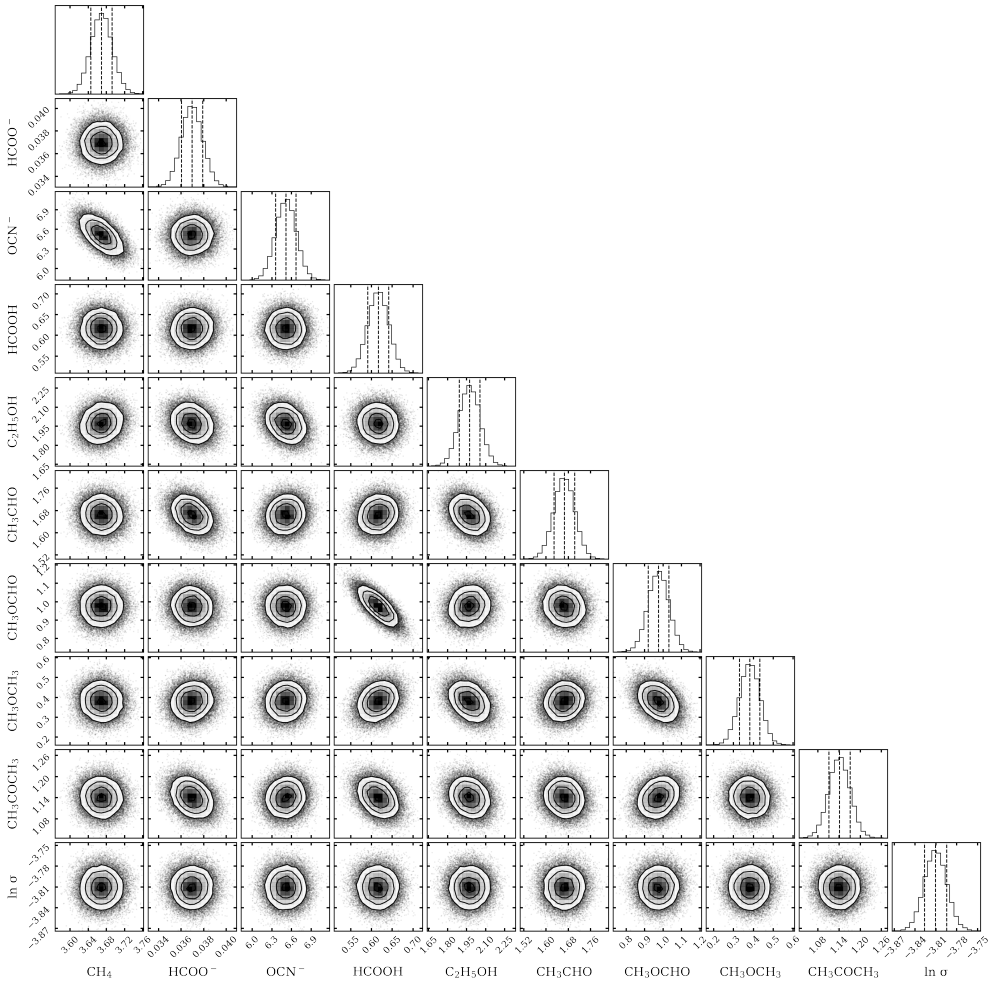


Figure 4.K.2: Corner plot of the MCMC fitting. Nine out of 12 candidate species are considered. The dashed lines in the histograms indicate the 16%, 50%, and 84% quantiles from left to right.

4.L Additional table

Table 4.L.1 lists the rotational transitions of the detected species reported in Table 4.1 covered in our ALMA Band 7 data (Sect. 4.2.1). Only those transitions that are used in the LTE fittings are listed.

Table 4.L.1: Rotational transitions of the detected gas-phase molecules in the ALMA spectra.

Species Database	Transition (J K_a K_c s — J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$	Note
CH_3OH	9 1 8 1 – 8 2 6 2	333864.722	125.52	8.04(-7)	⋆
(vt=0–2)	21 5 17 4 – 22 4 19 4	334326.177	964.39	1.99(-5)	⋆
CDMS	3 0 3 4 – 2 1 2 4	334426.571	314.47	5.55(-5)	†
	22 3 19 4 – 22 2 20 4	334626.849	1001.32	3.55(-5)	⋆
	25 3 23 5 – 24 2 23 5	334679.524	1073.85	1.46(-4)	†
	2 2 1 0 – 3 1 2 0	335133.570	44.67	2.69(-5)	†
	28 8 21 4 – 28 7 21 4	335205.794	1534.70	2.14(-4)	⋆
	7 1 7 0 – 6 1 6 0	335582.017	78.97	1.63(-4)	†
	14 7 7 0 – 15 6 9 0	336438.224	488.22	3.61(-5)	†
	14 7 8 0 – 15 6 10 0	336438.224	488.22	3.61(-5)	†
	7 1 7 6 – 6 1 6 6	336605.889	747.41	1.64(-4)	†
	27 9 18 0 – 26 6 21 3	336646.204	1293.65	2.29(-5)	
	27 9 19 0 – 26 6 20 3	336646.533	1293.65	2.29(-5)	
	7 6 1 6 – 6 6 0 6	336970.183	1022.71	4.57(-5)	
	7 6 2 6 – 6 6 1 6	336970.183	1022.71	4.57(-5)	
	7 3 4 8 – 6 3 3 8	337021.917	979.71	1.36(-4)	†
	7 2 6 6 – 6 2 5 6	337029.573	941.38	1.55(-4)	†
	7 2 5 6 – 6 2 4 6	337029.662	941.38	1.55(-4)	
	27 9 19 2 – 28 8 21 2	345318.303	1278.18	5.97(-5)	
	16 1 15 0 – 15 2 14 0	345903.916	332.65	1.04(-4)	†
	18 3 15 2 – 17 4 14 2	345919.260	459.43	7.29(-5)	†
	19 3 16 4 – 19 2 17 4	347445.285	856.25	4.04(-5)	
$^{13}CH_3OH$	12 1 11 -0 – 12 0 12 +0	335560.207	192.66	4.04(-4)	†
(vt=0–1)	14 1 13 -0 – 14 0 14 +0	347188.283	254.25	4.36(-4)	†
CDMS					
$CH_3^{18}OH$	4 0 4 1 – 3 1 3 1	336100.324	35.09	7.42(-5)	∗
(v=0–2)	11 6 6 2 – 12 5 8 2	345858.184	326.94	3.54(-5)	
CDMS					
CH_2DOH	19 8 11 2 – 20 7 14 2	333835.968	672.13	1.06(-5)	
JPL	19 8 12 2 – 20 7 13 2	333835.968	672.13	1.06(-5)	
	16 2 15 0 – 16 1 16 0	333966.537	306.68	1.59(-4)	†
	20 3 17 1 – 20 2 18 1	334615.482	494.74	1.40(-5)	
	16 2 15 2 – 16 1 15 1	334683.954	326.58	1.09(-4)	†
	13 0 13 1 – 12 1 11 2	334900.931	207.26	2.69(-5)	
	23 0 23 0 – 22 1 21 2	335102.203	583.84	3.71(-5)	

Table 4.L.1: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$	Note
	18 5 13 0 – 18 4 15 2	335103.631	464.26	2.80(-5)	
	18 5 14 0 – 18 4 14 2	335116.209	464.26	2.80(-5)	
	19 2 17 0 – 18 3 16 1	335578.757	427.23	1.50(-5)	
	20 3 18 2 – 19 4 15 2	336131.672	504.96	2.04(-5)	
	21 8 13 2 – 22 7 15 1	336973.378	759.76	2.43(-5)	
	21 8 14 2 – 22 7 16 1	336973.383	759.76	2.43(-5)	
	16 2 14 0 – 15 3 13 0	345398.904	310.21	5.83(-5)	
	23 4 19 1 – 23 3 21 2	345701.510	662.81	1.34(-4)	⋆
	3 2 1 1 – 2 1 2 1	345718.718	39.44	4.23(-5)	
	19 1 19 1 – 18 2 17 2	345820.793	418.03	2.94(-5)	⋆
	22 4 19 1 – 22 3 19 2	345850.485	613.62	1.29(-4)	⋆
	8 2 6 2 – 7 3 4 0	346031.073	112.62	1.76(-6)	
	20 4 16 1 – 19 5 15 0	346042.562	521.63	1.32(-5)	
	20 4 17 1 – 19 5 14 0	346062.966	521.63	1.32(-5)	
	21 4 18 1 – 21 3 18 2	346419.064	566.56	1.30(-4)	⋆
	7 5 2 2 – 8 4 5 2	346434.692	175.80	7.74(-6)	
	7 5 3 2 – 8 4 4 2	346434.725	175.80	7.74(-6)	
	20 4 16 1 – 20 3 18 2	347222.992	521.63	1.32(-4)	
	19 4 16 1 – 19 3 16 2	347371.167	478.84	1.30(-4)	
	20 7 14 1 – 21 6 15 0	347467.621	651.60	2.02(-5)	
	20 7 13 1 – 21 6 16 0	347468.666	651.60	2.02(-5)	
CH_3CHO JPL	18 1 18 1 – 17 1 17 1	333941.356	155.22	1.28(-3)	B1-c†
	18 1 18 0 – 17 1 17 0	333959.754	155.15	1.28(-3)	B1-c†
	18 1 18 3 – 17 1 17 3	334341.281	361.72	1.30(-3)	⋆
	17 2 15 3 – 16 2 14 3	334637.404	357.95	1.28(-3)	
	17 4 14 7 – 16 4 13 7	334665.820	548.85	1.18(-3)	
	25 4 22 1 – 25 3 23 1	334756.255	337.33	1.30(-4)	
	17 2 15 2 – 16 2 14 2	334904.591	152.63	1.28(-3)	B1-c†
	17 2 15 0 – 16 2 14 0	334931.415	152.61	1.28(-3)	B1-c†
	18 1 18 4 – 17 1 17 4	334980.853	359.89	1.28(-3)	⋆
	18 0 18 2 – 17 0 17 2	335318.109	154.93	1.30(-3)	B1-c†
	18 0 18 0 – 17 0 17 0	335358.722	154.85	1.29(-3)	B1-c†
	18 0 18 3 – 17 0 17 3	335382.461	361.48	1.31(-3)	⋆
	17 6 12 7 – 16 6 11 7	335560.509	591.60	1.10(-3)	
	17 2 15 6 – 16 2 14 6	336127.216	529.83	1.25(-3)	
	18 1 17 8 – 17 1 16 8	345992.130	545.09	1.38(-3)	
	18 2 17 4 – 17 2 16 4	346065.350	371.35	1.40(-3)	
	18 7 11 5 – 17 7 10 5	346451.075	474.02	1.22(-3)	⋆
	18 6 12 5 – 17 6 11 5	346591.092	445.29	1.28(-3)	
	18 7 12 0 – 17 7 11 0	346957.556	268.61	1.22(-3)	B1-c†
	18 7 11 0 – 17 7 10 0	346957.558	268.61	1.22(-3)	B1-c†
	18 7 12 1 – 17 7 11 1	346995.532	268.57	1.22(-3)	
	18 7 11 6 – 17 7 10 6	346999.913	646.35	1.22(-3)	⋆

Table 4.L.1: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$	Note
	18 7 12 6 – 17 7 11 6	346999.941	646.35	1.22(-3)	⋆
	18 6 13 1 – 17 6 12 1	347132.686	239.32	1.28(-3)	
	18 4 14 8 – 17 4 13 8	347155.125	573.93	1.37(-3)	
	9 4 6 4 – 9 3 7 4	347155.775	282.78	1.20(-4)	
	19 0 19 0 – 18 1 18 0	347169.520	171.81	2.09(-4)	
	18 4 14 5 – 17 4 13 5	347182.413	400.38	1.37(-3)	⋆
	18 5 13 5 – 17 5 12 5	347216.798	420.44	1.33(-3)	
	19 0 19 2 – 18 1 18 1	347240.396	171.89	2.08(-4)	
	18 5 14 4 – 17 5 13 4	347251.822	419.67	1.33(-3)	
	18 5 14 0 – 17 5 13 0	347288.264	214.70	1.33(-3)	
	18 5 13 0 – 17 5 12 0	347294.873	214.70	1.33(-3)	
	18 5 13 2 – 17 5 12 2	347345.710	214.64	1.33(-3)	
	18 5 14 1 – 17 5 13 1	347349.278	214.61	1.33(-3)	
	18 7 12 3 – 17 7 11 3	347459.353	473.15	1.22(-3)	⋆
	18 7 11 3 – 17 7 10 3	347459.355	473.15	1.22(-3)	⋆
	18 3 16 0 – 17 3 15 0	347519.185	178.75	1.41(-3)	B1-c†
	18 3 16 1 – 17 3 15 1	347563.334	178.71	1.41(-3)	B1-c†
C_2H_5OH	25 7 19 2 – 25 6 20 2	334110.890	334.91	2.39(-4)	
(v=0)	44 7 38 2 – 44 6 39 2	334543.992	897.54	2.64(-4)	
CDMS	24 7 17 2 – 24 6 18 2	334602.657	313.84	2.37(-4)	
	12 5 7 0 – 11 4 7 1	334742.442	152.19	1.27(-4)	
	20 1 19 2 – 19 2 18 2	334769.072	178.68	2.75(-4)	⋆
	12 5 8 0 – 11 4 8 1	334811.021	152.19	1.27(-4)	
	24 7 18 2 – 24 6 19 2	334911.577	313.83	2.38(-4)	
	23 7 16 2 – 23 6 17 2	335440.539	293.61	2.36(-4)	
	23 7 17 2 – 23 6 18 2	335630.619	293.61	2.36(-4)	
	13 3 11 2 – 12 2 10 2	335949.651	87.87	1.43(-4)	
	19 3 16 1 – 18 3 15 1	336572.411	232.34	3.14(-4)	
	19 2 18 2 – 18 1 17 2	336626.401	162.61	2.71(-4)	
	21 1 21 1 – 20 1 20 1	345295.355	246.22	3.67(-4)	
	21 0 21 0 – 20 0 20 0	345333.442	241.54	3.72(-4)	
	20 13 7 1 – 19 13 6 1	345966.686	443.28	2.10(-4)	
	20 13 8 1 – 19 13 7 1	345966.686	443.28	2.10(-4)	
	20 14 6 1 – 19 14 5 1	345970.743	476.27	1.85(-4)	
	20 14 7 1 – 19 14 6 1	345970.743	476.27	1.85(-4)	
	20 12 8 1 – 19 12 7 1	345981.237	412.70	2.33(-4)	
	20 12 9 1 – 19 12 8 1	345981.237	412.70	2.33(-4)	
	20 11 9 1 – 19 11 8 1	346017.212	384.56	2.55(-4)	⋆
	20 11 10 1 – 19 11 9 1	346017.212	384.56	2.55(-4)	⋆
	20 10 10 1 – 19 10 9 1	346085.563	358.85	2.74(-4)	⋆
	20 10 11 1 – 19 10 10 1	346085.563	358.85	2.74(-4)	⋆
	20 10 10 0 – 19 10 9 0	346424.583	353.45	2.91(-4)	⋆
	20 10 11 0 – 19 10 10 0	346424.583	353.45	2.91(-4)	⋆

Table 4.L.1: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$	Note
	20 9 12 0 – 19 9 11 0	346505.347	330.25	3.09(-4)	
	20 9 11 0 – 19 9 10 0	346505.347	330.25	3.09(-4)	
	20 7 14 1 – 19 7 13 1	346565.082	296.37	3.22(-4)	
	20 7 13 1 – 19 7 12 1	346565.398	296.37	3.22(-4)	
	20 17 3 0 – 19 17 2 0	346610.251	584.43	1.12(-4)	
	20 17 4 0 – 19 17 3 0	346610.251	584.43	1.12(-4)	
	20 8 13 0 – 19 8 12 0	346620.325	309.51	3.26(-4)	⋆
	20 8 12 0 – 19 8 11 0	346620.333	309.51	3.26(-4)	⋆
	20 6 14 1 – 19 6 13 1	346938.874	280.48	3.34(-4)	
	21 0 21 2 – 20 1 20 2	346962.603	185.84	4.45(-4)	⋆
	20 6 15 0 – 19 6 14 0	347147.209	275.46	3.55(-4)	
	20 6 14 0 – 19 6 13 0	347157.997	275.46	3.55(-4)	
	14 3 12 2 – 13 2 11 2	347351.089	99.66	1.56(-4)	
	21 1 21 2 – 20 0 20 2	347445.568	185.85	4.47(-4)	
	20 5 16 1 – 19 5 15 1	347473.561	267.09	3.45(-4)	
CH_3OCH_3	37 7 30 0 – 37 6 31 0	347340.892	710.39	2.38(-4)	∗
(v=0)	37 7 30 1 – 37 6 31 1	347342.977	710.39	2.38(-4)	∗
CDMS	37 7 30 5 – 37 6 31 5	347345.059	710.39	2.38(-4)	∗
	37 7 30 3 – 37 6 31 3	347345.066	710.39	2.38(-4)	∗
CH_3OCHO	15 6 10 1 – 14 5 9 2	333972.824	94.90	2.42(-5)	
JPL	27 11 16 2 – 26 11 15 2	334017.031	303.76	4.74(-4)	⋆
	27 11 17 0 – 26 11 16 0	334031.781	303.76	4.74(-4)	†
	27 11 16 0 – 26 11 15 0	334031.781	303.76	4.74(-4)	
	27 11 17 1 – 26 11 16 1	334044.362	303.76	4.74(-4)	
	27 8 20 3 – 26 8 19 3	334074.073	453.00	5.18(-4)	
	15 6 10 0 – 14 5 9 0	334109.114	94.90	3.79(-5)	
	27 6 22 4 – 26 6 21 4	334138.972	435.17	5.39(-4)	
	29 5 24 2 – 28 6 23 1	334179.438	282.10	2.99(-5)	
	26 5 21 5 – 25 5 20 5	334228.598	416.22	5.48(-4)	⋆
	29 5 24 0 – 28 6 23 0	334235.866	282.10	2.99(-5)	⋆
	27 8 19 5 – 26 8 18 5	334660.939	453.20	5.11(-4)	
	27 8 19 3 – 26 8 18 3	334700.891	453.07	5.21(-4)	⋆
	27 10 17 2 – 26 10 16 2	334850.947	290.09	4.93(-4)	
	27 10 18 0 – 26 10 17 0	334867.014	290.09	4.94(-4)	⋆
	27 10 17 0 – 26 10 16 0	334872.791	290.09	4.94(-4)	⋆
	27 10 18 1 – 26 10 17 1	334877.566	290.08	4.93(-4)	⋆
	27 7 21 3 – 26 7 20 3	335015.896	443.53	5.33(-4)	⋆
	27 8 20 4 – 26 8 19 4	335391.536	453.05	5.14(-4)	
	27 9 19 0 – 26 9 18 0	336028.165	277.85	5.14(-4)	
	27 9 19 1 – 26 9 18 1	336032.357	277.85	4.90(-4)	
	27 9 18 2 – 26 9 17 2	336086.182	277.86	4.90(-4)	
	27 9 18 0 – 26 9 17 0	336111.324	277.86	5.15(-4)	
	9 9 0 0 – 8 8 1 0	345718.662	80.32	9.86(-5)	

Table 4.L.1: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$	Note
	9 9 1 0 – 8 8 0 0	345718.662	80.32	9.86(-5)	
	28 6 23 4 – 27 6 22 4	345828.557	451.76	5.99(-4)	
	28 12 16 2 – 27 12 15 2	345974.664	335.44	5.16(-4)	
	28 12 17 0 – 27 12 16 0	345985.381	335.44	5.16(-4)	
	28 12 16 0 – 27 12 15 0	345985.381	335.44	5.16(-4)	
	28 12 17 1 – 27 12 16 1	346001.616	335.43	5.16(-4)	⋆
	28 9 20 4 – 27 9 19 4	346564.463	480.56	5.69(-4)	
	27 6 21 3 – 26 6 20 3	346580.244	437.64	6.04(-4)	
	27 5 22 2 – 26 5 21 2	347478.251	247.25	6.14(-4)	†
	27 5 22 0 – 26 5 21 0	347493.965	247.26	6.14(-4)	†
CH_2OHCHO	37 8 29 0 – 36 9 28 0	334382.061	439.41	1.83(-4)	∗
(v=0)	35 14 21 0 – 35 13 22 0	334587.394	469.56	5.85(-4)	
CDMS	35 14 22 0 – 35 13 23 0	334588.479	469.56	5.85(-4)	
	50 7 43 0 – 50 6 44 0	334859.254	750.32	4.85(-4)	
	17 6 11 0 – 16 5 12 0	335097.924	107.27	3.89(-4)	
	34 14 20 0 – 34 13 21 0	335441.062	450.00	5.81(-4)	
	34 14 21 0 – 34 13 22 0	335441.575	450.00	5.81(-4)	
	30 5 25 0 – 29 6 24 0	335456.548	282.11	4.31(-4)	
	47 15 32 0 – 47 14 33 0	346572.448	765.89	7.26(-4)	∗
	18 6 12 0 – 17 5 13 0	347445.973	117.44	4.02(-4)	
***a*-$(CH_2OH)_2$**	32 18 14 1 – 31 18 13 0	334043.836	418.76	6.24(-4)	
CDMS	32 18 15 1 – 31 18 14 0	334043.836	418.76	6.24(-4)	
	32 17 15 1 – 31 17 14 0	334058.781	401.61	6.55(-4)	
	32 17 16 1 – 31 17 15 0	334058.781	401.61	6.56(-4)	
	32 19 13 1 – 31 19 12 0	334059.121	436.89	5.91(-4)	
	32 19 14 1 – 31 19 13 0	334059.121	436.89	5.91(-4)	
	32 20 12 1 – 31 20 11 0	334100.228	456.00	5.57(-4)	
	32 20 13 1 – 31 20 12 0	334100.228	456.00	5.57(-4)	
	32 16 16 1 – 31 16 15 0	334109.885	385.45	6.85(-4)	
	32 16 17 1 – 31 16 16 0	334109.885	385.45	6.85(-4)	
	36 2 35 0 – 35 2 34 1	334141.377	310.88	9.15(-4)	†
	36 1 35 0 – 35 1 34 1	334141.397	310.88	9.15(-4)	†
	32 21 11 1 – 31 21 10 0	334163.829	476.08	5.20(-4)	
	32 21 12 1 – 31 21 11 0	334163.829	476.08	5.21(-4)	
	33 10 24 0 – 32 10 23 1	334193.102	325.62	8.29(-4)	
	32 15 17 1 – 31 15 16 0	334205.280	370.28	7.13(-4)	⋆
	32 15 18 1 – 31 15 17 0	334205.280	370.28	7.14(-4)	⋆
	33 10 23 0 – 32 10 22 1	334211.959	325.62	8.30(-4)	
	32 22 10 1 – 31 22 9 0	334247.377	497.14	4.82(-4)	
	32 22 11 1 – 31 22 10 0	334247.377	497.14	4.82(-4)	
	32 23 9 1 – 31 23 8 0	334348.904	519.17	4.43(-4)	
	32 23 10 1 – 31 23 9 0	334348.904	519.17	4.43(-4)	
	32 14 18 1 – 31 14 17 0	334356.392	356.11	7.40(-4)	⋆

Table 4.L.1: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$	Note
	32 14 19 1 – 31 14 18 0	334356.392	356.11	7.40(-4)	⋆
	32 24 8 1 – 31 24 7 0	334466.874	542.17	4.01(-4)	
	32 24 9 1 – 31 24 8 0	334466.874	542.17	4.01(-4)	
	32 6 26 0 – 31 6 25 1	334555.253	282.34	8.95(-4)	
	32 13 20 1 – 31 13 19 0	334579.759	342.94	7.66(-4)	
	32 13 19 1 – 31 13 18 0	334579.760	342.94	7.66(-4)	
	36 0 36 0 – 35 1 35 0	334804.512	300.16	1.85(-4)	
	36 1 36 0 – 35 0 35 0	334804.513	300.16	1.85(-4)	
	36 0 36 1 – 35 1 35 1	334821.007	300.51	1.85(-4)	
	36 1 36 1 – 35 0 35 1	334821.008	300.51	1.85(-4)	
	32 12 21 1 – 31 12 20 0	334900.249	330.78	7.91(-4)	
	32 12 20 1 – 31 12 19 0	334900.284	330.78	7.91(-4)	
	32 6 27 1 – 31 6 26 0	335030.116	279.36	9.34(-4)	
	33 9 25 0 – 32 9 24 1	335179.801	316.68	8.51(-4)	
	32 11 22 1 – 31 11 21 0	335356.975	319.65	8.15(-4)	
	32 11 21 1 – 31 11 20 0	335357.726	319.65	8.15(-4)	
	33 9 24 0 – 32 9 23 1	335396.713	316.71	8.53(-4)	
	34 4 31 1 – 33 3 30 1	336053.816	297.13	1.33(-4)	
	34 4 31 0 – 33 3 30 0	336054.127	296.80	1.12(-4)	
	34 5 29 1 – 33 6 28 1	336608.832	311.78	3.62(-4)	
	33 11 23 1 – 32 11 22 0	345737.011	335.95	9.00(-4)	
	33 11 22 1 – 32 11 21 0	345738.443	335.95	9.01(-4)	
	34 9 26 0 – 33 9 25 1	345812.719	333.57	9.36(-4)	
	34 8 27 0 – 33 8 26 1	345832.886	325.79	5.43(-4)	
	38 1 38 0 – 37 1 37 1	345912.524	333.62	1.02(-3)	
	38 0 38 0 – 37 0 37 1	345912.524	333.62	1.02(-3)	
	34 7 28 0 – 33 7 27 1	346462.657	318.83	9.43(-4)	⋆
	33 10 24 1 – 32 10 23 0	346468.981	325.91	9.26(-4)	⋆
	33 10 23 1 – 32 10 22 0	346490.976	325.91	9.26(-4)	
	34 5 29 0 – 33 5 28 1	346630.908	311.52	1.07(-3)	
	36 4 33 0 – 35 4 32 1	347361.377	330.36	1.02(-3)	
	36 3 33 0 – 35 3 32 1	347377.705	330.36	1.01(-3)	⋆
	32 6 26 1 – 31 6 25 0	347387.171	282.65	9.81(-4)	
	33 9 25 1 – 32 9 24 0	347486.902	316.98	9.51(-4)	
	33 8 25 1 – 32 8 25 1	347596.647	309.41	1.23(-4)	
***g*-$(CH_2OH)_2$** CDMS	23 5 19 0 – 22 4 19 1	333807.874	147.69	1.88(-4)	
	31 17 14 1 – 31 16 16 0	333830.012	382.39	1.47(-4)	
	31 17 15 1 – 31 16 15 0	333830.012	382.39	1.47(-4)	
	30 17 13 1 – 30 16 15 0	333926.319	367.26	1.43(-4)	
	30 17 14 1 – 30 16 14 0	333926.319	367.26	1.43(-4)	
	32 5 27 1 – 31 5 26 0	334079.683	276.07	3.19(-4)	
	28 17 11 1 – 28 16 13 0	334088.087	338.47	1.35(-4)	
	28 17 12 1 – 28 16 12 0	334088.087	338.47	1.35(-4)	

Table 4.L.1: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$	Note
	32 7 25 0 – 31 7 24 1	334098.077	284.34	3.31(-4)	
	34 3 31 1 – 33 4 30 1	334112.202	295.02	1.78(-4)	
	34 3 31 0 – 33 4 30 0	334120.029	294.96	3.08(-4)	
	27 17 10 1 – 27 16 12 0	334155.000	324.80	1.30(-4)	
	27 17 11 1 – 27 16 11 0	334155.000	324.80	1.30(-4)	
	20 7 13 1 – 19 6 13 0	334167.497	126.59	2.43(-4)	
	26 17 9 1 – 26 16 11 0	334213.515	311.63	1.25(-4)	
	26 17 10 1 – 26 16 10 0	334213.515	311.63	1.25(-4)	
	34 4 31 1 – 33 3 30 1	334253.644	295.03	3.07(-4)	
	34 4 31 0 – 33 3 30 0	334264.186	294.97	1.79(-4)	
	25 17 8 1 – 25 16 10 0	334264.285	298.94	1.19(-4)	
	25 17 9 1 – 25 16 9 0	334264.285	298.94	1.19(-4)	
	24 17 7 1 – 24 16 9 0	334307.936	286.74	1.12(-4)	
	24 17 8 1 – 24 16 8 0	334307.936	286.74	1.12(-4)	
	18 8 11 1 – 17 7 10 1	334320.081	114.67	1.86(-4)	
	18 8 10 1 – 17 7 11 1	334325.869	114.67	1.86(-4)	
	23 17 6 1 – 23 16 8 0	334345.068	275.03	1.04(-4)	
	23 17 7 1 – 23 16 7 0	334345.068	275.03	1.04(-4)	
	18 8 11 0 – 17 7 10 0	334354.965	114.62	1.88(-4)	
	18 8 10 0 – 17 7 11 0	334360.695	114.62	1.89(-4)	
	12 11 1 1 – 11 10 2 1	334506.538	96.91	3.56(-4)	⋆
	12 11 2 1 – 11 10 1 1	334506.538	96.91	3.56(-4)	⋆
	12 11 1 0 – 11 10 2 0	334538.884	96.86	3.56(-4)	⋆
	12 11 2 0 – 11 10 1 0	334538.884	96.86	3.56(-4)	⋆
	20 7 14 1 – 19 6 14 0	334562.201	126.58	2.05(-4)	
	16 9 8 1 – 15 8 7 1	334599.304	105.79	2.28(-4)	
	16 9 7 1 – 15 8 8 1	334599.324	105.79	2.28(-4)	
	14 10 4 1 – 13 9 5 1	334604.473	99.88	2.81(-4)	
	14 10 5 1 – 13 9 4 1	334604.473	99.88	2.81(-4)	
	16 9 8 0 – 15 8 7 0	334633.280	105.74	2.30(-4)	
	16 9 7 0 – 15 8 8 0	334633.300	105.74	2.30(-4)	
	14 10 4 0 – 13 9 5 0	334637.793	99.83	2.82(-4)	
	14 10 5 0 – 13 9 4 0	334637.793	99.83	2.82(-4)	
	36 0 36 1 – 35 0 35 0	334709.849	298.92	3.61(-4)	⋆
	36 1 36 1 – 35 1 35 0	334709.849	298.92	3.61(-4)	⋆
	33 20 13 0 – 32 20 12 1	334838.561	467.86	2.18(-4)	
	33 20 14 0 – 32 20 13 1	334838.561	467.86	2.18(-4)	
	33 21 12 0 – 32 21 11 1	334840.150	487.72	2.05(-4)	
	33 21 13 0 – 32 21 12 1	334840.150	487.72	2.05(-4)	
	33 19 14 0 – 32 19 13 1	334851.819	448.96	2.31(-4)	
	33 19 15 0 – 32 19 14 1	334851.819	448.96	2.31(-4)	
	33 22 11 0 – 32 22 10 1	334852.879	508.55	1.91(-4)	
	33 22 12 0 – 32 22 11 1	334852.879	508.55	1.91(-4)	

Table 4.L.1: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$	Note
	33 18 15 0 – 32 18 14 1	334884.208	431.04	2.43(-4)	
	33 18 16 0 – 32 18 15 1	334884.208	431.04	2.43(-4)	
	33 17 16 0 – 32 17 15 1	334940.915	414.08	2.55(-4)	
	33 17 17 0 – 32 17 16 1	334940.915	414.08	2.55(-4)	
	22 6 17 1 – 21 5 17 0	334954.517	141.61	2.37(-4)	
	33 5 29 0 – 32 4 28 0	334987.470	286.45	3.98(-4)	
	33 16 17 0 – 32 16 16 1	335028.496	398.10	2.66(-4)	
	33 16 18 0 – 32 16 17 1	335028.496	398.10	2.66(-4)	
	33 15 18 0 – 32 15 17 1	335155.577	383.11	2.76(-4)	
	33 15 19 0 – 32 15 18 1	335155.577	383.11	2.76(-4)	
	23 4 20 0 – 22 3 20 1	335226.462	142.78	1.46(-4)	
	33 14 20 0 – 32 14 19 1	335333.954	369.09	2.86(-4)	
	33 14 19 0 – 32 14 18 1	335333.954	369.09	2.86(-4)	
	34 4 31 1 – 33 4 30 0	335422.139	295.03	2.84(-4)	⋆
	34 3 31 1 – 33 3 30 0	335459.569	295.02	4.14(-4)	
	18 8 10 1 – 17 7 10 0	335462.789	114.67	2.38(-4)	
	18 8 11 1 – 17 7 11 0	335467.703	114.67	2.37(-4)	
	36 6 30 1 – 35 7 29 1	335488.573	349.86	1.07(-4)	
	33 13 21 0 – 32 13 20 1	335580.432	356.08	2.96(-4)	
	33 13 20 0 – 32 13 19 1	335580.435	356.08	2.96(-4)	
	35 3 33 0 – 34 3 32 1	335588.334	302.57	3.60(-4)	
	35 2 33 0 – 34 2 32 1	335589.204	302.57	3.55(-4)	
	32 7 25 1 – 31 7 24 0	336072.528	284.39	3.32(-4)	
	33 22 11 1 – 32 22 10 0	336092.641	508.57	1.93(-4)	
	33 22 12 1 – 32 22 11 0	336092.641	508.57	1.93(-4)	
	33 20 13 1 – 32 20 12 0	336099.583	467.89	2.20(-4)	
	33 20 14 1 – 32 20 13 0	336099.583	467.89	2.20(-4)	
	33 23 10 1 – 32 23 9 0	336116.773	530.36	1.78(-4)	
	33 23 11 1 – 32 23 10 0	336116.773	530.36	1.78(-4)	
	33 15 18 1 – 32 15 17 0	336558.515	383.14	2.79(-4)	
	33 15 19 1 – 32 15 18 0	336558.515	383.14	2.79(-4)	
	23 4 20 1 – 22 2 20 1	337017.819	142.83	1.10(-4)	
	25 6 20 0 – 24 5 19 0	337038.379	177.09	1.09(-4)	
	34 15 19 0 – 33 15 18 1	345385.355	399.71	3.07(-4)	
	34 15 20 0 – 33 15 19 1	345385.355	399.71	3.07(-4)	
	13 11 2 1 – 12 10 2 0	345836.731	103.25	3.87(-4)	⋆
	13 11 3 1 – 12 10 3 0	345836.731	103.25	3.87(-4)	⋆
	17 9 8 1 – 16 8 8 0	345841.998	114.10	2.83(-4)	⋆
	17 9 9 1 – 16 8 9 0	345841.998	114.10	2.83(-4)	⋆
	34 13 22 0 – 33 13 21 1	345862.205	372.71	3.27(-4)	⋆
	34 13 21 0 – 33 13 20 1	345862.211	372.71	3.27(-4)	⋆
	15 10 5 1 – 14 9 5 0	345906.318	107.20	3.25(-4)	
	15 10 6 1 – 14 9 6 0	345906.318	107.20	3.25(-4)	

Table 4.L.1: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) $a(b) = a \times 10^b$	Note
	36 2 34 1 – 35 3 33 1	345985.032	319.24	3.20(-4)	
	36 3 34 1 – 35 2 33 1	345986.658	319.24	3.23(-4)	
	36 2 34 0 – 35 3 33 0	345987.306	319.18	3.24(-4)	
	36 3 34 0 – 35 2 33 0	345988.992	319.18	3.21(-4)	
	34 17 17 1 – 33 17 16 0	346444.360	430.71	2.87(-4)	
	34 17 18 1 – 33 17 17 0	346444.360	430.71	2.87(-4)	
	24 3 21 1 – 23 2 21 0	346551.132	154.12	1.52(-4)	
	35 5 30 0 – 34 6 29 0	346551.443	326.48	1.29(-4)	
	35 5 30 1 – 34 6 29 1	346568.765	326.51	2.15(-4)	
	34 16 18 1 – 33 16 17 0	346577.753	414.74	2.98(-4)	
	34 16 19 1 – 33 16 18 0	346577.753	414.74	2.98(-4)	
	29 6 24 1 – 28 5 23 1	346986.538	231.40	1.13(-4)	
	34 14 21 1 – 33 14 20 0	346993.104	385.75	3.20(-4)	
	34 14 20 1 – 33 14 19 0	346993.104	385.75	3.20(-4)	
	36 3 34 1 – 35 3 33 0	347254.922	319.24	3.93(-4)	
	36 2 34 1 – 35 2 33 0	347255.462	319.24	3.97(-4)	
	34 13 22 1 – 33 13 21 0	347306.553	372.75	3.30(-4)	
	34 13 21 1 – 33 13 20 0	347306.559	372.75	3.30(-4)	
	34 10 25 0 – 33 10 24 1	347502.019	339.86	3.56(-4)	
	37 2 36 0 – 36 2 35 1	347507.368	326.07	4.01(-4)	
	37 1 36 0 – 36 1 35 1	347507.368	326.07	4.01(-4)	
	34 10 24 0 – 33 10 23 1	347535.163	339.86	3.56(-4)	
t-HCOOH	35 4 31 – 35 3 32	334247.840	739.34	7.49(-6)	
CDMS	15 2 14 – 14 2 13	334265.833	141.58	4.17(-4)	†
H_2CCO	17 1 16 – 16 1 15	346600.451	162.79	4.74(-4)	
CDMS					
CH_2DCN	20 1 20 – 19 1 19	345685.375	179.63	3.60(-3)	
CDMS	20 7 13 – 19 7 12	346968.947	438.92	3.20(-3)	
	20 7 14 – 19 7 13	346968.947	438.92	3.20(-3)	
	20 0 20 – 19 0 19	347043.458	174.96	3.65(-3)	⋆
	20 6 14 – 19 6 13	347044.649	368.95	3.32(-3)	⋆
	20 6 15 – 19 6 14	347044.649	368.95	3.32(-3)	⋆
	20 5 15 – 19 5 14	347110.032	309.72	3.42(-3)	⋆
	20 5 16 – 19 5 15	347110.032	309.72	3.42(-3)	⋆
	20 4 17 – 19 4 16	347166.409	261.24	3.50(-3)	⋆
	20 4 16 – 19 4 15	347166.421	261.24	3.50(-3)	⋆
	20 2 19 – 19 2 18	347188.253	196.56	3.61(-3)	
	20 3 18 – 19 3 17	347216.880	223.53	3.57(-3)	
	20 3 17 – 19 3 16	347219.365	223.53	3.57(-3)	
	20 2 18 – 19 2 17	347388.249	196.62	3.62(-3)	
C_2H_5CN	37 5 32 – 36 5 31	333921.554	330.94	3.11(-3)	
(v=0)	39 3 37 – 38 3 36	345921.198	344.46	3.50(-3)	

Table 4.L.1: continued.

Species Database	Transition (J K_a K_c s – J K_a K_c s)	Frequency (MHz)	E_{up} (K)	A_{ij} (s^{-1}) Note $a(b) = a \times 10^b$
CDMS	38 3 35 – 37 3 34	346983.834	333.42	3.53(-3)
NH_2CHO	16 2 15 – 15 2 14	336136.877	149.67	2.76(-3)
CDMS	16 1 15 – 15 1 14	345326.688	145.16	3.02(-3)

Note:
⋆: for the robustly detected species, key transitions that are used to constrain the column densities and excitation temperatures; must be optically thin and unblended.
∗: for the tentatively detected species, key transitions that are used to estimate upper limits of column densities.
†: transitions that are (expected to be) optically thick and not considered in the fitting.

5 JOYS+: An overview of the ice features in the COM fingerprint range (6.8–8.8 μm)

Y. Chen, E. F. van Dishoeck, L. Francis, K. Slavicinska, Ł. Tychoniec, M. L. van Gelder, A. C. A. Boogert, C. A. F. Schravendijk, H. Beuther, A. Caratti o Garatti, E. K. H. Eenhorst, C. Gieser, Th. Henning, M. Güdel, V. J. M. Le Gouellec, M. K. McClure, M. Narang, P. Nazari, S. D. Reyes, M. Tolman, and Y.-L. Yang

Submitted to Astronomy & Astrophysics

Abstract

Context. Complex organic molecules (COMs) are the key to understanding the chemical evolution from simple interstellar species to potential prebiotic material. Observations with the *James Webb* Space Telescope (JWST) have recently enabled direct detection of multiple oxygen-bearing COMs in ices toward protostars. The characterization is mainly made with the absorption features in the 6.8–8.8 µm range, also known as the COM fingerprint range.

Aims. This study aims to characterize the ice absorption features in the COM ice fingerprint range across a large sample of protostars (mainly low-mass), to explore their prevalence, spectral diversity, and relation to more abundant ice species like H_2O and CH_4. We aim to identify the species responsible for the observed ice bands through a systematic analysis of their band properties, and to assess spectral fitting results from previous case studies.

Methods. We analyzed JWST/MIRI-MRS spectra of ∼20 low-mass protostars from the JWST Observations of Young protoStars (JOYS+) program, with additional five sources from the Investigating Protostellar Accretion (IPA) program. We implemented an interactive workflow that isolates the COM fingerprint region by removing gas-phase lines of SO_2 and CH_4 (when necessary), determining a global continuum, subtracting the silicate and the H_2O libration bands, and fitting a local continuum. Key properties (peak, width, and area) were measured for four absorption bands at 7.02, 7.24, 7.4, and 7.67 µm using Gaussian fitting. These band properties were compared with the same quantities derived from laboratory spectra of COM ice mixtures that may contribute to the observed bands.

Results. A strong correlation is observed between the peak optical depths of the silicate band at ∼9.8 µm and the H_2O ice column densities. Positive correlations are found between the ice column densities of H_2O and CH_4, as well as between the integrated areas of the 7.24, 7.4, and 7.67 µm (CH_4) bands. These correlations suggest that the absorption features in the COM ice fingerprint range are indeed tracing the ice reservoir in a protostellar system. The bands at 7.24, 7.4, and 7.67 µm are commonly detected in Class 0 and borderline 0/I sources, whereas those at 7.02 µm and beyond 8.0 µm are less constrained due to overlapping with the strong band of NH_4^+ and silicates, respectively. Class I sources barely show absorption features in the COM ice fingerprint range, likely due to the loss of their envelopes. On one hand, the 7.67 µm band shows consistent properties across the sample, which strongly supports the sole contribution from CH_4 ice. On the other hand, the bands at 7.02, 7.24 and 7.40 µm, which are attributed to ice mixtures of $HCOO^-$, CH_3CHO, and C_2H_5OH, exhibit considerable variation in their peak positions and FHWMs. Comparisons with laboratory spectra suggest that multiple ice components are likely to contribute to these bands.

Conclusions. This work presents the first large-sample survey of COM ice features in the 6.8–8.8 µm range using JWST/MIRI-MRS spectra. The several absorption bands detected in the COM ice fingerprint range along with the silicate and the liberation H_2O bands are well correlated with each other in terms of optical depth or band area, suggesting that they are good tracers of ice reservoirs in protostellar envelopes.

The detection of those ice bands also suggest that salts like $HCOO^-$ and COMs have already formed in the solid phase by the protostellar stage, which complements the gas-phase observations of interstellar molecules in star-forming regions.

5.1 Introduction

Understanding the chemical evolution of the universe from simple atoms in diffuse clouds to the complex chemistry necessary for life is one of the overarching goals of astrochemistry. Complex organic molecules (COMs), typically defined as carbon-containing species with six or more atoms (Herbst & van Dishoeck 2009), are a key piece in this puzzle due to their potential links to prebiotic chemistry. COMs have been detected in various astrophysical environments, but mostly in the warm (> 100 K) inner regions of low- and high-mass protostars, often referred to as hot corinos and hot cores, respectively (Jørgensen et al. 2020; Ceccarelli et al. 2023). Observations with (sub)millimeter facilities such as the Atacama Large Millimeter/submillimeter Array (ALMA) have revealed rich gas-phase COMs in protostellar sources, indicating a high degree of chemical complexity at early stages of star formation (see the review by Nazari 2025 and references therein).

The formation history of COMs remains a topic of active research. A growing body of observational, theoretical, and experimental work supports a primarily grain-surface origin for these species. In cold prestellar cores ($T \lesssim 10$ K), CO gas freezes out onto dust grains and undergoes successive hydrogenation, forming CH_3OH and more complex species within the ice mantles (e.g., Watanabe & Kouchi 2002; Fuchs et al. 2009; Simons et al. 2020; Fedoseev et al. 2022). As the protostar forms and heats its surroundings, these COMs are thermally desorbed into the gas phase, becoming observable at millimeter (mm) wavelengths through their rotational transitions. However, gas-phase formation routes may also play a role in COM formation (Balucani et al. 2015), especially in the warm-up phase or in regions where energetic processing (e.g., UV irradiation, cosmic rays) is significant (Garrod et al. 2022). The relative contributions of solid-phase inheritance and gas-phase reprocessing are still debated and likely vary among sources and molecules, as implied from the scattered gas-phase ratios of COMs with respect to CH_3OH in large samples of protostars (Nazari et al. 2022a; Chapter 2).

Recently, the advent of the *James Webb* Space Telescope (JWST) has significantly enhanced our ability to probe the solid-phase inventory of COMs. Before JWST, solid-phase detections were largely limited to CH_3OH. Now, thanks to the Medium Resolution Spectroscopy (MRS) mode of JWST's Mid-InfraRed Instrument (MIRI), absorption features of multiple oxygen-bearing COMs (CH_3CHO, C_2H_5OH, CH_3OCHO, and CH_3COCH_3) have been firmly or tentatively detected in mid-infrared (mid-IR) spectra of protostars by focusing on the COM ice fingerprint range in 6.8–8.8 µm (Rocha et al. 2024; Chapter 4; van Dishoeck et al. 2025). The simplest nitrogen-bearing COM, CH_3CH, has also been tentatively detected in the near-IR using the Near Infrared Spectrograph (NIRSpec; Nazari et al. 2024a). These observations provide, for the first time, the opportunity to directly compare the abundances of COMs in both the solid and the gas phases in the same (low-mass) protostellar sources (i.e., B1-c and NGC 1333 IRAS 2A, hereafter IRAS 2A; Chapter 4). Such comparisons suggest that both inheritance and reprocessing are plausible when COMs sublimate from icy mantles to the gas phase.

Conclusions drawn from case studies, however, need to be validated in a larger sample of sources, and the methodology of isolating and decomposing ice features needs to be standardized for application in large-sample studies as well. In this work, we

investigated the COM fingerprint range (6.8–8.8 µm) along with the silicate (9.8 µm) and the water liberation (∼13 µm) bands in a combined sample of 25 protostars observed by three JWST Cycle 1 programs (Sect. 5.2). To facilitate our analysis of a large amount of MIRI-MRS spectra of protostellar sources, we developed an interactive workflow that optimizes the fitting with visual inspection (Sect. 5.3). Given the fact that most of the sources do not have as rich ice features as B1-c and IRAS 2A show in the COM range, we did not fit the absorption bands with laboratory spectra and derive the ice column density of candidate species. Instead, we analyzed the band properties (e.g., peak wavelength, optical depth, width, and area) to reveal systematic trends and correlations between different bands and species (Sect. 5.4). The conclusions are summarized in Sect. 5.5.

5.2 Observations

The JWST/MIRI-MRS spectra used in this work were collected from three Cycle 1 programs focusing on (primarily low-mass) protostars, including two Guaranteed Time Observation (GTO) programs 1290 (P.I.: E. F. van Dishoeck; van Dishoeck et al. 2025) and 1236 (P.I.: M. E. Ressler), and one General Observers (GO) program 1802, Investigating Protostellar Accretion (IPA; P.I.: T. Megeath; Federman et al. 2024). GTO 1290 is also known as the JWST Observations of Young protoStars (JOYS) program, and GTO 1236 is part of the Europe-US collaboration based on JOYS (collectively referred to as JOYS+).

The data reduction of GTO 1290, GTO 1236, and GO 1802 has been described in van Gelder et al. (2024a), van Gelder et al. (2024b), and Federman et al. (2024), respectively. All the spectra were extracted from the IR continuum peak positions to maximize the signal-to-noise ratios (SNRs) of absorption features. The extraction aperture (in diameter) was usually set as three to four times of the point spread function (PSF) of MIRI-MRS (Law et al. 2023), which is given by

$$\mathrm{FWHM_{PSF}} = 0.033'' \times (\lambda/\mu\mathrm{m}) + 0.106''. \tag{5.1}$$

A slightly different aperture was used to include extended emission in some sources (e.g., the binary of B1-a-NS and the scattered light of L1527). The coordinates and apertures of spectral extraction along with other source information are provided in Table 5.1.

A total of 25 sources were selected from the three programs (Table 5.1), of which 16 are from GTO 1290 (JOYS), four from GTO 1236 (JOYS+), and five from GO 1802 (IPA). We excluded a few sources in GTO 1290 and more than half of the sources in GTO 1236. Most of these excluded sources are less embedded later-stage (Class I/II) sources that show barely detected ice features or a flat SED due to their lower envelope masses; others are too faint (embedded) and have very low SNRs. We refer to van Dishoeck et al. (2025) and Federman et al. (2024) for more detailed information of the sources from JOYS and IPA.

Except for one intermediate-mass (HOPS 370) and one high-mass (IRAS 20126) source, all the other sources considered in this work are low-mass. Although the JOYS program contains several high-mass sources, those with high-SNR COM fingerprint range have been or will be presented in other studies (Rocha et al. 2024; van Dishoeck

et al. 2025; Gieser et al. submitted), therefore we only include the low-mass sample. Both the two higher-mass sources are from the IPA program, which was designed to target protostars with a wide range of masses and luminosities. Considering its small sample size, we include all the five IPA sources in this study. Although no comparisons between low-mass and high-mass sources will be made in this work, studies focusing exclusively on high-mass sources are expected to be carried out in the future based on JOYS+ and the recently approved JWST Cycle 4 program GO 8887 (P.I.: Y. Chen).

5.3 Methods

The case studies on COM ices by Rocha et al. (2024) and Chapter 4 have proposed to conduct several steps to isolate the COM fingerprint range from the original spec-

Table 5.1: Source information

Source	Other name	Program*	RA† (°:′:″)	Dec† (hh:mm:ss)	Aperture‡	D (pc)	Class
B1-a-NS	Per-emb 40	JOYS	03:33:16.6850	+31:07:55.1011	5×PSF	293	I
B1-b	Per-emb 41	JOYS	03:33:20.3580	+31:07:21.3600	4×PSF	293	I
B1-c	Per-emb 29	JOYS	03:33:17.8850	+31:09:31.7050	4×PSF	293	0
NGC 1333 IRAS 1A	Per-emb 35A	JOYS+	03:28:37.0880	+31:13:30.6890	3×PSF	293	I
NGC 1333 IRAS 2A	Per-emb 27	JOYS+	03:28:55.5560	+31:14:36.7620	4×PSF	293	0
L1448-mm	Per-emb 26	JOYS	03:25:38.8700	+30:44:05.5200	4×PSF	293	0
L1448-IRS2B	Per-emb 22B	JOYS+	03:25:22.3480	+30:45:13.1540	4×PSF	293	0
L1448-IRS3B1	Per-emb 33B	JOYS+	03:25:36.3090	+30:45:14.8660	3×PSF	293	0
Per-emb 8		JOYS	03:44:43.9897	+32:01:35.5634	4×PSF	321	0
L1527	IRAS 04368+2557	JOYS	04:39:53.8600	+26:03:09.8500	3″ (fixed)	142	0/I
TMC1-E	IRAS 04381+2540	JOYS	04:41:12.7341	+25:46:34.6375	2×PSF	142	I
TMC1-W	IRAS 04381+2540	JOYS	04:41:12.6886	+25:46:34.6288	2×PSF	142	I
TMC1A	IRAS 04365+2535	JOYS	04:39:35.2100	+25:41:43.7600	4×PSF	142	I
Ser-S68N-N		JOYS	18:29:48.1358	+01:16:44.5310	3×PSF	436	0
Ser-SMM1A		JOYS	18:29:49.7780	+01:15:20.4300	3×PSF	436	0
Ser-SMM1B		JOYS	18:29:49.6650	+01:15:21.1560	3×PSF	436	0
Ser-SMM3		JOYS	18:29:59.3000	+01:14:00.7000	4×PSF	436	0
SVS4-5		JOYS	18:29:57.5980	+01:13:00.1900	4×PSF	436	I/II
(Ser-SMM4⋆)							(0⋆)
HOPS 153		IPA	05:37:57.0317	–07:06:56.2270	4×PSF	390	0
HOPS 370		IPA	05:35:27.6501	–05:09:34.9820	7×PSF	390	0/I
BHR71-IRS1		JOYS	12:01:36.4700	–65:08:49.4000	4×PSF	200	0
BHR71-IRS2		JOYS	12:01:35.6000	–65:08:47.9000	4×PSF	200	0
B335	IRAS 19345+0727	IPA	19:37:00.9145	+07:34:09.4620	4×PSF	165	0
IRAS 16253-2429		IPA	16:28:21.6167	–24:36:24.1180	4×PSF	140	0
IRAS 20126+4104		IPA	20:14:26.0757	+41:13:32.2240	4×PSF	1550	0

* JOYS = GTO 1290, JOYS+ = GTO 1236, IPA = GO 1802.
† The coordinates are those of the mid-IR continuum peak where the MIRI-MRS spectra were extracted. The coordinates are given in J2000.
‡ The extraction aperture in diameter. PSF is an abbreviation of $\mathrm{FWHM_{PSF}} = 0.033'' \times (\lambda/\mu\mathrm{m}) + 0.106''$, where λ is the median wavelength of each MIRI-MRS Channel (i.e., the extraction aperture increases with wavelength).
⋆ SVS4-5 is a Class I/II source located behind the envelope of the Class 0 source Ser-SMM4 in the line of sight.

trum for quantitative analysis (e.g., fitting with laboratory spectra). In this work, we followed the same procedure but optimized some fitting steps for a large sample in a way that they can be performed interactively and efficiently.

5.3.1 Removal of gas-phase lines

For some sources, the MIRI-MRS spectra extracted at IR continuum peaks are rich in gas-phase lines of simple molecules like H_2O, CO_2, and CH_4 (as shown in the overview study on gas-phase lines in protostars by van Gelder et al. 2024a). These lines can be present either in emission or in absorption, superposed on broader ice absorption bands. Extracting the spectrum at an offset position can be a workaround when studying simple molecules such as H_2O and CO_2, which are distributed in more extended regions and have strong ice bands, but this method does not work well for more complex and less abundant species with weak bands. Therefore, we kept our extraction positions at continuum peaks where both the ice column densities and the SNR of absorption features are higher.

To facilitate our studies on absorption bands, we traced a baseline of gas-phase lines by fitting a univariate spline function to the guiding points selected at line free channels. This baseline was used as the "gas-line-subtracted spectrum" in the following analysis instead of the original spectrum. For the sources in GTO 1290, these baselines have been traced by van Gelder et al. (2024a) for the purpose of gas-line fitting, and we repeated the same procedure for sources in the other two programs (GTO 1236 and GO 1802).

In this step, special attention should be paid to the ro-vibrational lines of SO_2 or CH_4 gas. The P/Q/R-branches of SO_2 gas at ∼7.35 µm and the Q-branch of CH_4 gas at 7.66 µm, whether in emission or in absorption, can create a local pseudo-continuum and significantly alter the shape of the absorption bands at 7.24, 7.40, and 7.67 µm (as shown in Fig. 5.1). A detailed case study on gas-phase SO_2 emission in IRAS 2A by van Gelder et al. (2024b) has shown the necessity of removing gas lines to restore the absorption features between 7.25 and 7.45 µm. In our combined sample, there are four sources with distinct SO_2 and CH_4 gas lines in emission (L1448-mm, IRAS 1A, IRAS 2A, IRAS 20126), and two in absorption (B1-c and L1448-IRS2B).

Although it is difficult to properly restore the original band profile without a least-squares fitting of the gas lines, the ro-vibrational transitions of SO_2 and CH_4 are quite sensitive to temperature (e.g., van Gelder et al. 2024a), hence it is relatively easy to obtain a plausible fit by visual inspection, which was done for sources in GTO 1236 and GO 1802 using slab models in `radexpy` (e.g., Francis et al. 2024; Tabone et al. 2023). For the sources in GTO 1290, we took the fitting results from van Gelder et al. (2024a). In practice, the by-eye fitting was performed species by species (SO_2 followed by CH_4), and the baseline was determined based on the spectrum with both SO_2 and CH_4 lines subtracted. We did not model the ro-vibrational lines of H_2O, which are abundant but do not show typical P/Q/R-branches that produce pseudo-continuum like symmetric molecules do, thus it is trivial to trace a baseline for them by eye.

In Chapter 4, the step of removing gas lines was done after the spectrum was converted into an optical depth scale. In this work, we brought this step to the front and performed it in the flux scale, in order to facilitate the modeling of gas lines in emission when necessary. We emphasize the importance of removing gas lines

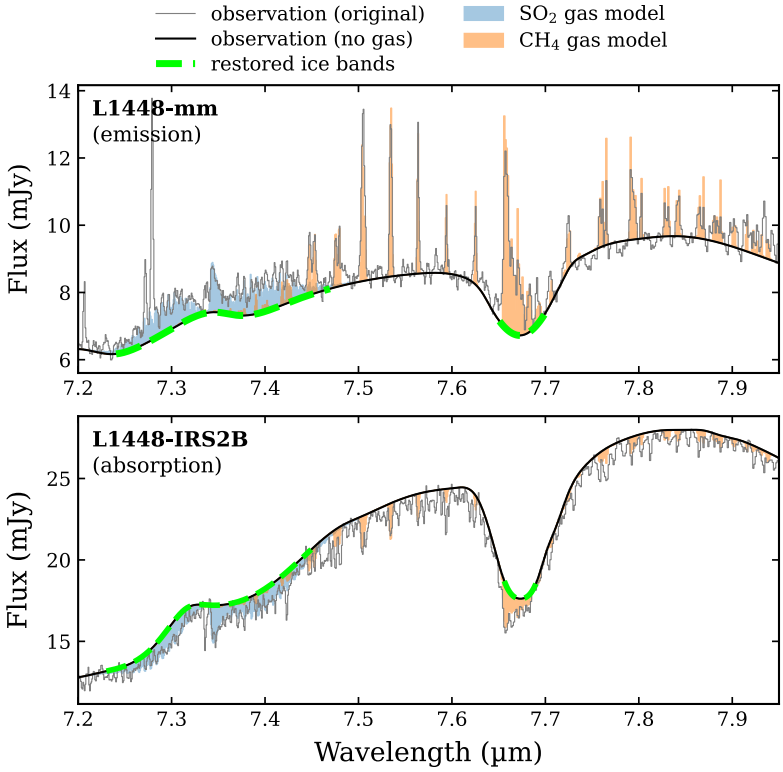


Figure 5.1: Two examples of how the gas-phase lines interfere with the characterization of absorption bands in the COM ice fingerprint range. The original observed spectra and the manually traced baselines of gas-phase lines are plotted in gray and black, respectively. The regions shaded in blue and orange indicate the modeled ro-vibrational lines of SO_2 and CH_4, respectively. The absorption features contaminated by gas-phase lines are highlighted as dashed green lines. We note that L1448-mm is a source with one of the strongest gas-phase emission lines in our sample (van Gelder et al. 2024a).

properly, especially for those that can create pseudo-continuum and hence mislead our identification and measurements of the underlying absorption bands.

Tracing the baseline of gas lines also allows us to smooth the low-SNR part of the spectrum (similar to binning) and clean leftover artifacts that were not fully removed during data reduction (e.g., very strong cosmic-ray showers; Regan 2023). We noticed that some artifacts may create fake emission or absorption features that are wavelength- or pixel-dependent. For example, there are two "absorption bands" present at 8.13 and 8.17 μm in the spectrum of L1527 and both are persistent around the continuum. Although they look like weak ice bands, similar features were not observed in any other source. After inspecting the data files in stages 1 and 2, we confirmed that they are artifacts caused by cosmic-rays, and filtered them out when removing the gas lines. This example shows a caveat that not every absorption feature present in the science-ready spectrum is real; we should check the calibrated images in different data reduction stages when seeing a suspicious case.

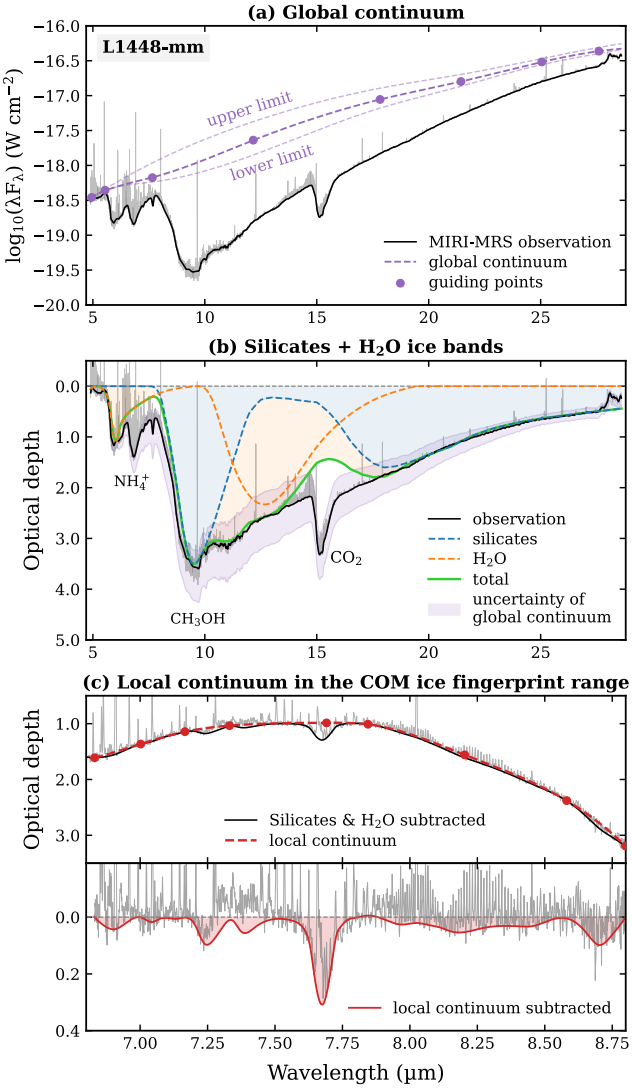


Figure 5.2: The routine of fitting (a) the global continuum, (b) the silicates and H_2O ice bands, and (c) the local continuum to isolate the COM ice features between ∼6.8–8.8 μm, taking L1448-mm as an example. In each panel, the spectrum in gray is the original spectrum without subtracting the ro-vibrational lines of gas-phase molecules (i.e., the "spiky" features). The thick black line is the spectrum with gas-phase lines removed (i.e., the baseline of gas-phase lines). In panel (a), the guiding points and the global continuum determined using the interactive fitting tool are shown in purple. Panel (b) shows the spectrum in optical depth converted by dividing the global continuum in panel (a). The purple shaded area correspond to the upper and lower limits of the global continuum in panel (a). The optical depth components of silicates and H_2O are shaded in blue and orange, respectively, and the green line indicates the sum of them. Panel (c) zooms into the COM ice fingerprint range between 6.8 and 8.8 μm; both the local continuum (upper sub-panel) and the observed spectrum with local continuum subtracted (lower sub-panel) are shown in red. The same plot for other sources are shown in Figs. 5.B.1–5.B.6.

5.3.2 Interactive fitting of global continuum, silicates, and H_2O

After removing the gas-phase lines, the observed spectrum was converted into optical depth using

$$\tau_{\rm obs} = \ln(F_{\rm cont}/F_{\rm obs}), \tag{5.2}$$

where $F_{\rm obs}$ is the observed spectrum in flux and $F_{\rm cont}$ is the global continuum to be determined. For the global continuum, we usually need to "guess" its shape based on the two ends of the MIRI-MRS spectrum (∼5.0–5.5 µm and ∼26–28 µm), where the absorption is supposed to be minimum. However, the broad and deep absorption of silicates bands at 9.8 and 18 µm and the H_2O libration band at ∼13 µm, which are usually fitted after the global continuum subtraction, makes it hard to uniquely determine the global continuum in between. In other words, the three steps of fitting the global continuum, the silicate bands, and the H_2O libration band are linked to each other; if one of the last two steps does not work well, we have to go back to the previous one(s) to adjust.

This routine of determining the global continuum of an IR protostellar spectrum was initially introduced and applied to ∼40 low-mass protostars observed by *Spitzer* in Boogert et al. (2008), who described the work as "not trivial". Unfortunately, after entering the era of JWST, this is still an inevitable and tricky problem to ice studies. Most of the strategies employed so far are tailored to individual sources (e.g., Rocha et al. 2024; Chapter 4; Slavicinska et al. 2025b; Brunken et al. 2025; Rayalacheruvu et al. 2025) and become very labor-intensive when dealing with a large sample with different SED shapes. The main obstacle is that it is very challenging to fully automate the fitting routine by models, and fitting by eye seems the most pragmatic and practical way.

To make this subjective fitting routine more straightforward and efficient, we developed an interactive tool that allows us to fit the global continuum along with the silicate and the H_2O liberation bands in one step. We refer to Appendix 5.A for a detailed introduction to this tool. In general, its interactive interface can update changes in real time when the users adjust any of the fitting parameters (e.g., the shape of the global continuum, the scaling factor of the silicate or the H_2O spectrum). In this way, we no longer need to go back and forth among separate steps to check and tune our fitting; instead, we integrate these steps into one where all the adjustments can be made interactively, allowing us to explore and compare different possibilities efficiently. Apart from finding the best fit for the silicate and the H_2O bands, we can also constrain the upper and lower limits of the global continuum (the dashed purple lines shown in Fig. 5.2a), that is, the extent to which the continuum can be lifted or lowered while still producing a plausible fit for silicates and H_2O, as shown in Fig. 5.2b.

In this step, we can measure the peak optical depths of the silicate band at 9.8 µm and the ice column densities ($N_{\rm ice}$) of H_2O using

$$N_{\rm ice} = \frac{1}{A}\int_{\tilde{\nu}_1}^{\tilde{\nu}_2} a\tau_{\tilde{\nu},\rm lab}\ {\rm d}\tilde{\nu}, \tag{5.3}$$

where A is the band strength, $\tilde{\nu}$ is the wavenumber, $\tilde{\nu}_1$ and $\tilde{\nu}_2$ are the two ends of the band in wavenumber, $\tau_{\tilde{\nu},\rm lab}$ is the laboratory spectrum in optical depth (the H_2O spectra were measured by Slavicinska et al. 2024), and a is the scaling factor of this spectrum fitted to the observations. The upper and lower limits of the global

continuum estimated using the interactive fitting tool can be further translated to the upper and lower limits of the silicate optical depths and the H_2O ice column densities.

The spectra of synthetic silicates used in the fitting were generated by `OpTool` (Dominik et al. 2021) for olivine ($MgFeSiO_4$; Dorschner et al. 1995) and pyroxene ($Mg_{0.7}Fe_{0.3}SiO_3$; Dorschner et al. 1995), each mixed with 13% mass of carbon (Zubko et al. 1996). Here 13% is the default carbon mass fraction used in `OpTool`, following the DIANA standard dust model (Woitke et al. 2016). The grain size distribution is a default powerlaw ($a^{-3.5}$) with the grain size (a) varies between 0.005 and 1 µm. For some sources such as B1-c, the 18 µm silicate band will be fitted better with larger grain sizes (up to 3 µm). The relative contribution of olivine and pyroxene also differs from source to source (e.g., sometimes only pyroxene is needed). There remains a substantial room for further research and discussion on the absorption features of silicates and H_2O, which is beyond the scope of this paper and will be pursued in future studies. The following analysis in this work will only involve the silicate optical depths at 9.8 µm and the H_2O ice column densities.

5.3.3 Interactive fitting of local continuum

After subtracting the global continuum and the broad absorption features of silicates and H_2O, we still need to fit and subtract a local continuum to isolate the weak features present in the COM ice fingerprint range (as shown in Fig. 5.2c). In Rocha et al. (2024) and Chapter 4, a polynomial was fitted to a few manually selected guiding points at absorption-free channels. However, after inspecting a larger sample in this work, we found that polynomials are not always able to properly resemble the local continuum, even up to the tenth order. Therefore, we again applied the spline function used for removing gas-phase lines (Sect. 5.3.1), which is more flexible and appropriate than polynomials in terms of fitting a local continuum.

The selection and adjustment of guiding points along with the spline fitting were also done in an interactive interface, similar to fitting the global continuum. Most of the guiding points were put in between the major ice bands, but we still need some anchor points in the middle of the CH_4 band at 7.67 µm (usually at the top of the fingerprint range) and the broad 8.2 µm band. We deliberately set the local continuum as close to the observed spectrum as possible (as done in Chapter 4), thus strictly speaking, the derived optical depths of all isolated ice bands are lower limits.

Another difference in doing this step between this work and previous studies is whether the local continuum should be traced on the spectrum before or after silicate subtraction. In Rocha et al. (2024) and Chapter 4, it is suggested to fit a local continuum after the silicate band at 9.8 µm is subtracted, so that the weak features beyond 8 µm can be lifted to a similar level to the short-wavelength part, and the local continuum can be fitted by lower-order polynomials. In this work, however, we noticed that the synthetic silicate spectra generated by `OpTool` may introduce subtle features to the observed ice bands beyond 7.8 µm. These features are likely artifacts due to the low spectral resolution of the silicate spectra measured in laboratories, on which the `OpTool` calculation is based. Those features are usually very small but may become non-negligible when the real absorption features are weak or the SNR is low (especially in the 8.0–8.8 µm range). Also when the silicate band at 9.8 µm is saturated in some sources, it will be over-subtracted and the shape of the silicate-subtracted spectrum

beyond ∼8 µm will be distorted. Taking all factors into account, we propose that it is more appropriate to trace the local continuum before subtracting silicate features.

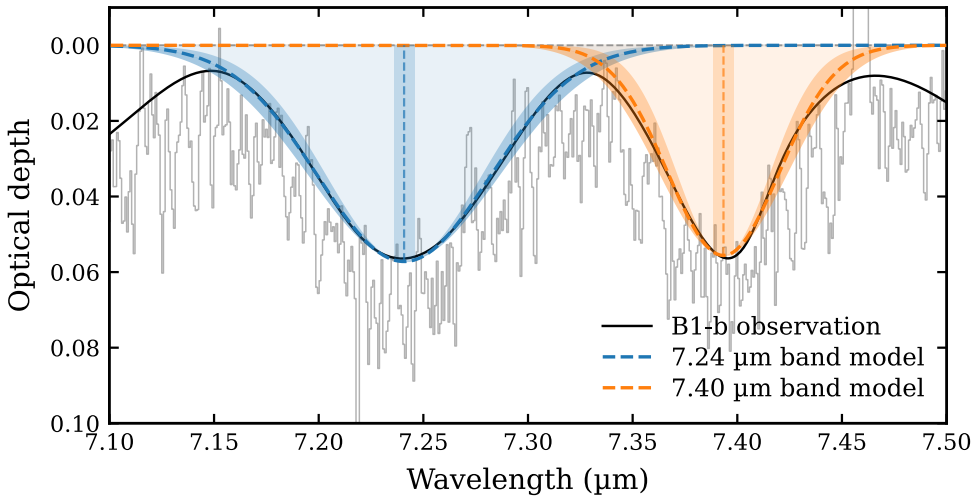


Figure 5.3: Gaussian fitting of the 7.24 and 7.4 µm bands, taking B1-b as an example. The observations before and after gas-line removal are shown as thin gray and thick blue lines, respectively. The shaded regions in blue and orange represent the two Gaussian functions jointly fitted to the 7.24 and 7.40 µm bands. The darker shaded regions indicate the uncertainties in the peak wavelength (0.005 µm) and FWHM (0.01 µm); see Appendix 5.D for more description about uncertainty estimation.

5.3.4 Determination of band properties

After obtaining an isolated spectrum in the COM ice fingerprint range, we measured four properties: peak wavelength ($\lambda_{\rm peak}$), peak optical depth ($\tau_{\rm peak}$), full width half maximum (FWHM), and integrated area of the bands at 7.02, 7.24, 7.4, and 7.67 µm, which concern the absorption features of CH_3CHO, C_2H_5OH, $HCOO^-$, and CH_4 ices. In most cases, the four bands could be well fitted by Gaussians, and the 7.24 and 7.4 µm bands were jointly fitted with two Gaussians when feasible (the joint fitting made use of `lmfit`; Newville et al. 2014), as shown in Fig. 5.3. If the band shape deviates too much from a Gaussian function, we measured the $\lambda_{\rm peak}$ and FWHM manually by definition (i.e., the wavelength at which the spectral feature reaches its maximum optical depth, and the width of the feature measured at half of its peak value). The fitting results of band properties and the fitting strategies (i.e., either using a Gaussian function or fitting by definition) are listed in Table 5.C.1. The uncertainty of the Gaussian fitting itself is usually small (e.g., on the order of 10^{-4} µm for $\lambda_{\rm peak}$ and FWHM). However, there are larger uncertainties resulting from previous steps, such as the gas-line removal (Sect. 5.3.1) and the local continuum subtraction (Sect. 5.3.3), of which the uncertainties are difficult to quantify. Here we adopt a conservative estimation for the uncertainties of $\lambda_{\rm peak}$ (0.005 µm), FWHM (0.01 µm), and integrated area (20%). More details on the uncertainty estimation are provided

in Appendix 5.D.

5.4 Results and discussion

5.4.1 Silicates and simple ices

In the step of fitting the global continuum and the bands of silicates and H_2O ice (Sect. 5.3.2), we obtained the peak optical depths of the silicate band at 9.8 µm ($\tau_{9.8,\,\mathrm{si}}$) and the ice column densities (N_{ice}) of H_2O derived from its libration band at ∼13 µm (integrated over 9.6–19.9 µm). The values are provided in Table 5.2. We remark that the $\tau_{9.8,\,\mathrm{si}}$ here is not the total optical depth at 9.8 µm, which is usually determined from the H_2 (0-0) S(3) line at 9.7 µm and denoted as $\tau_{9.7}$ or $\tau_{\mathrm{S(3)}}$ (e.g., van Dishoeck et al. 2025; Francis et al. 2025). The difference between $\tau_{9.8,\,\mathrm{si}}$ and $\tau_{9.7}$ depends on the ice bands of NH_3 at ∼9.0 µm and CH_3OH at 9.74 µm (excluded in $\tau_{9.8,\,\mathrm{si}}$ but included in $\tau_{9.7}$), which are often saturated on top of the silicate band (e.g., the example shown in Fig. 5.2b), but can be deep in some sources.

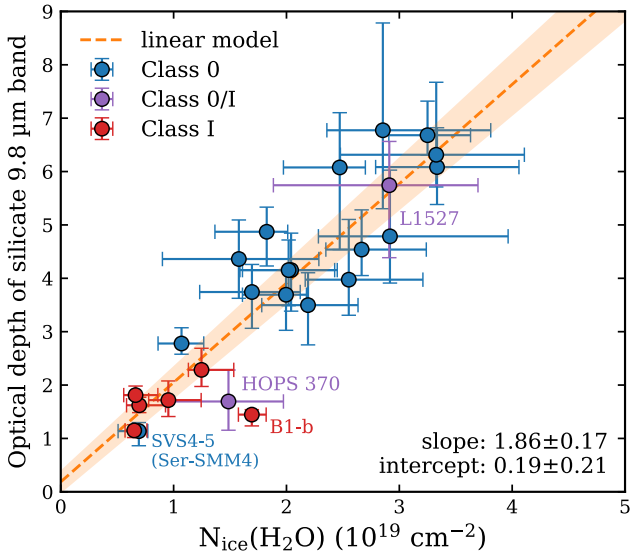


Figure 5.4: Peak optical depths of the silicate band at ∼9.8 µm versus H_2O ice column densities. Sources in different evolutionary stages are indicated by different colors (Class 0–blue, Class 0/I–purple, Class I–red). The best linear fit and its uncertainty are indicated by the dashed lines and the shaded region in orange, respectively. The best-fit value and the 1σ error of the slope and intercept are indicated in the lower right corner.

We fit a linear model to the relation between $\tau_{9.8,\,\mathrm{si}}$ and $N_{\mathrm{ice}}(H_2O)$ using Markov chain Monte Carlo (MCMC; Foreman-Mackey et al. 2013), with a likelihood function customized to properly account for asymmetric uncertainties and upper or lower limits. There is a strong correlation between $\tau_{9.8,\,\mathrm{si}}$ and $N_{\mathrm{ice}}(H_2O)$ in our sample of 25 protostars, as shown in Fig. 5.4. This is expected given that both quantities are considered representative of dust mass probed along the line of sight, which can be further

related to ice reservoirs in protostellar envelopes (e.g., Crapsi et al. 2008; Boogert et al. 2015). The small but (tentatively) positive intercept at $\tau_{9.8,\,\mathrm{si}} = 0.19 \pm 0.21$ implies an extinction threshold for H_2O ice formation on dust grains; that is, the environment should be dense and shielded enough to allow solid-phase formation of H_2O). Although our sample is biased toward dense environments (i.e., protostars), the presence of this threshold is consistent with previous studies, as summarized in Fig. 7 of Boogert et al. (2015). There is also a clear trend in evolutionary stage that the absorption features of silicates and H_2O are shallower in Class I sources, which is consistent with other theoretical and observational results that protostars in later stages are more exposed and have less massive envelopes (e.g., Young & Evans 2005; Enoch et al. 2009). There are two outliers, one is the borderline Class 0/I source L1527, which is an edge-on disk where the dust and ices are accumulated along the observer's line of sight; the other is SVS4-5, which is a Class I/II source hiding behind the envelope of the Class 0 source Ser-SMM4 (the data point is labeled as "Class 0" based on the source of the envelope being probed). Aside from these two sources, other sources are clearly separated into two groups by their evolutionary classifications.

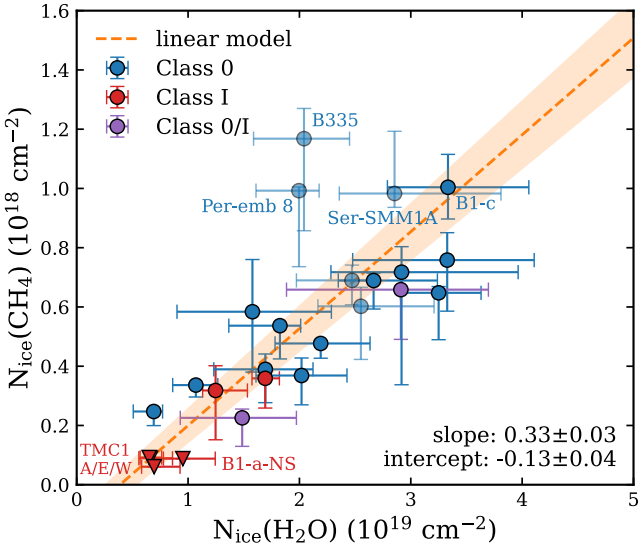


Figure 5.5: Ice column densities of H_2O versus CH_4. The legend is the same as in Fig. 5.4. In some Class I sources, the CH_4 band is detected in low SNR and thus considered as upper limit (denoted by downward triangles). Data points for sources with low SNR in their CH_4 band at 7.67 µm are plotted more transparently.

We also checked the correlation between the ice column densities of H_2O and CH_4. CH_4 ice has the strongest band within the COM fingerprint range and is supposed to form in H_2O-rich layers as the fully hydrogenated product of carbon atoms. Its precursor or dissociated product, the CH_3 radical, is suggested to be an important ingredient in the formation of many COMs that are abundant in the gas phase (e.g., CH_3OCH_3, CH_3OCHO, and CH_3COCH_3; see Sect. 3.5.1). A similarly strong and positive correlation was found between the ice column densities of H_2O and CH_4, as shown in Fig. 5.5. The two evident outliers, B335 and Per-emb 8, are sources with

Table 5.2: Optical depths of the silicate 9.8 µm band, ice column densities of H_2O and CH_4, and detection conditions of absorption bands in the COM fingerprint range.

Source	Si. 9.8 µm	H_2O 13 µm	CH_4	7.67 µm	7.24 µm	7.40 µm	7.02 µm	>7.8 µm
	τ_{peak}	N_{ice} (cm^{-2})	N_{ice}* (cm^{-2})	σ†	σ†	σ†	σ†	σ†
B1-a-NS	$1.7^{+0.4}_{-0.3}$	$9.5^{+2.9}_{-1.7} \times 10^{18}$	$< 8.8 \times 10^{16}$	13 ✔	– ✘	– ✘	– ✘	– ✘
B1-b	$1.4^{+0.1}_{-0.2}$	$1.7^{+0.1}_{-0.1} \times 10^{19}$	$3.6^{+0.3}_{-1.0} \times 10^{17}$	36 ✔	6.9 ✔	6.8 ✔	4.5 ✔!	1.8 !
B1-c	$6.1^{+0.7}_{-0.7}$	$3.3^{+0.7}_{-0.5} \times 10^{19}$	$1.0^{+0.1}_{-0.1} \times 10^{18}$	258 ✔	51 ✔	54 ✔	32 ✔	91 ✔
NGC 1333 IRAS 1A	$2.3^{+0.4}_{-0.3}$	$1.2^{+0.3}_{-0.1} \times 10^{19}$	$3.2^{+0.8}_{-1.7} \times 10^{17}$	67 ✔	18 ✔	11 ✔	6.6 ✔	3.6 ✔!
NGC 1333 IRAS 2A	$4.8^{+1.2}_{-0.9}$	$2.9^{+1.0}_{-0.6} \times 10^{19}$	$7.2^{+0.9}_{-3.8} \times 10^{17}$	57 ✔	12 ✔	12 ✔	4.6 ✔!	8.5 ✔
L1448-mm	$3.5^{+0.6}_{-0.7}$	$2.2^{+0.4}_{-0.4} \times 10^{19}$	$4.8^{+0.1}_{-0.5} \times 10^{17}$	29 ✔	9.0 ✔	5.2 ✔	– ✘	5.1 ✔
L1448-IRS2B	$4.9^{+0.5}_{-0.6}$	$1.8^{+0.2}_{-0.5} \times 10^{19}$	$5.4^{+0.1}_{-1.1} \times 10^{17}$	90 ✔	24 ✔	20 ✔	10 ✔	10 ✔
L1448-IRS3B1	$4.5^{+0.7}_{-0.5}$	$2.7^{+0.6}_{-0.3} \times 10^{19}$	$6.9^{+0.3}_{-1.0} \times 10^{17}$	64 ✔	11 ✔	12 ✔	5.1 ✔	4.5 ✔!
Per-emb 8	$3.7^{+0.5}_{-0.7}$	$2.0^{+0.2}_{-0.4} \times 10^{19}$	$9.9^{+0.1}_{-2.6} \times 10^{17}$	7.3 ✔	– ✘	– ✘	– ✘	– ✘
L1527	$5.7^{+0.8}_{-1.4}$	$2.9^{+0.8}_{-1.0} \times 10^{19}$	$6.6^{+0.5}_{-1.7} \times 10^{17}$	36 ✔	4.3 ✔!	4.9 ✔!	8.2 ✔	– ?
TMC1-E	$1.6^{+0.2}_{-0.1}$	$7.0^{+2.3}_{-1.1} \times 10^{18}$	$< 6.1 \times 10^{16}$	10 ✔	3.2 ✔!	1.9 !	– ✘	– ✘
TMC1-W	$1.2^{+0.1}_{-0.1}$	$6.5^{+1.1}_{-0.8} \times 10^{18}$	$< 9.2 \times 10^{16}$	13 ✔	– ✘	– ✘	– ✘	– ✘
TMC1A	$1.8^{+0.2}_{-0.1}$	$6.6^{+2.0}_{-1.0} \times 10^{18}$	$< 8.9 \times 10^{16}$	10 ✔	– ✘	– ✘	– ✘	– ✘
Ser-S68N-N	$3.7^{+0.5}_{-0.7}$	$1.7^{+0.4}_{-0.5} \times 10^{19}$	$3.9^{+0.5}_{-1.1} \times 10^{17}$	24 ✔	7.6 ✔	4.8 ✔!	2.8 !	2.5 !
Ser-SMM1A	$6.8^{+2.0}_{-1.5}$	$2.9^{+1.0}_{-0.5} \times 10^{19}$	$9.8^{+2.1}_{-0.5} \times 10^{17}$	21 ✔	2.3 !	3.3 ✔!	– ?	4.3 ✔!
Ser-SMM1B	$6.3^{+1.4}_{-0.6}$	$3.3^{+0.8}_{-0.8} \times 10^{19}$	$7.6^{+0.9}_{-1.7} \times 10^{17}$	175 ✔	24 ✔	22 ✔	11 ✔	7.0 ✔
Ser-SMM3	$2.8^{+0.3}_{-0.2}$	$1.1^{+0.2}_{-0.2} \times 10^{19}$	$3.4^{+0.1}_{-0.4} \times 10^{17}$	12 ✔	2.3 !	2.8 !	1.4 !	– ?
SVS4-5 (Ser-SMM4)	$1.1^{+0.2}_{-0.3}$	$6.9^{+0.8}_{-1.8} \times 10^{18}$	$2.5^{+0.1}_{-0.5} \times 10^{17}$	34 ✔	17 ✔	16 ✔	13 ✔	3.3 ✔!
HOPS 153	$4.0^{+1.1}_{-0.7}$	$2.6^{+0.7}_{-0.4} \times 10^{19}$	$6.0^{+0.6}_{-1.8} \times 10^{17}$	45 ✔	6.3 ✔	6.2 ✔	7.2 ✔	– ?
HOPS 370	$1.7^{+0.6}_{-0.5}$	$1.5^{+0.5}_{-0.6} \times 10^{19}$	$2.3^{+0.3}_{-1.0} \times 10^{17}$	129 ✔	122 ✔	86 ✔	40 ✔	12 ✔
BHR71-IRS1	$6.7^{+0.6}_{-0.4}$	$3.2^{+0.4}_{-0.3} \times 10^{19}$	$6.5^{+0.2}_{-1.6} \times 10^{17}$	166 ✔	34 ✔	28 ✔	23 ✔	22 ✔
BHR71-IRS2	$6.1^{+1.0}_{-1.5}$	$2.5^{+0.2}_{-0.5} \times 10^{19}$	$6.9^{+0.5}_{-0.8} \times 10^{17}$	15 ✔	2.4 !	2.1 !	– ?	– ✘
B335	$4.2^{+0.7}_{-0.8}$	$2.0^{+0.4}_{-0.5} \times 10^{19}$	$1.2^{+0.1}_{-0.3} \times 10^{18}$	53 ✔	7.9 ✔	6.3 ✔	– ?	– ?
IRAS 16253	$4.2^{+0.6}_{-0.7}$	$2.0^{+0.4}_{-0.4} \times 10^{19}$	$3.7^{+0.6}_{-1.0} \times 10^{17}$	67 ✔	7.1 ✔	5.2 ✔	4.4 ✔!	2.3 !
IRAS 20126	$4.4^{+0.7}_{-0.7}$	$1.6^{+0.7}_{-0.7} \times 10^{19}$	$5.8^{+1.8}_{-2.0} \times 10^{17}$	186 ✔	38 ✔	24 ✔	18 ✔	97 ✔

* We consider those $N_{ice}(CH_4)$ smaller than 1×10^{17} cm^{-2} as upper limits.
† Indication of different symbols: ✔ ($\geq 5\sigma$; robust detection), ✔! ($< 5\sigma$ and $\geq 3\sigma$; detection), ! ($< 3\sigma$ and $\geq 1\sigma$; tentative detection), ? (this band is visible but has a irregular shape), ✘ (this band is not detected).

very low SNRs between ∼8 and 14 µm, where the bands of silicates and H_2O tend to be underestimated in the global fitting (see Fig. 5.B.5). In contrast, the CH_4 band is located in a wavelength range of higher flux and SNR, and is therefore better detected in both sources. The strong correlation between H_2O and CH_4 ices and the positive x-intercept is consistent with our current understanding of their formation: both H_2O and CH_4 ices form through the hydrogenation of O and C atoms in polar layers (see Sect. 1.2.1 and Fig. 1.1), but CH_4 ice accumulates to a detectable abundance later than H_2O.

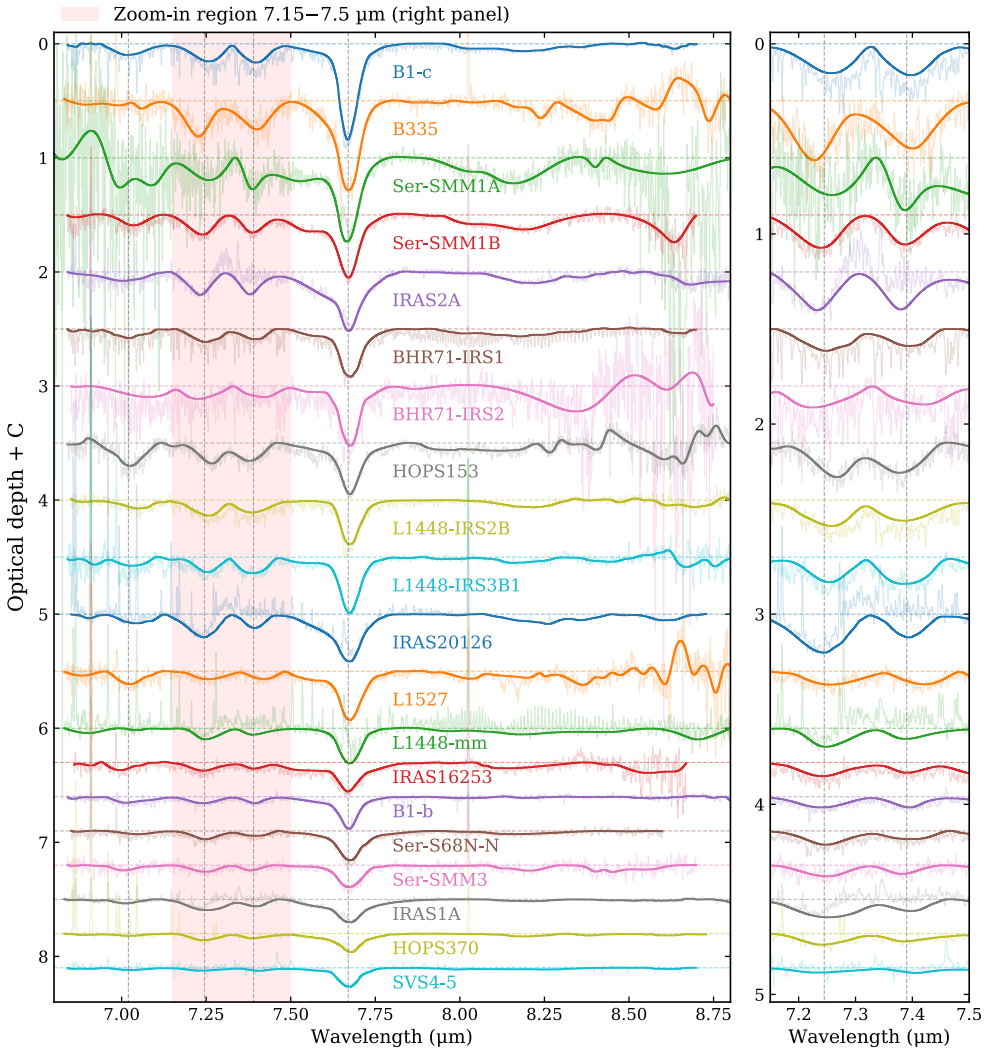


Figure 5.6: Overview of the isolated COM ice fingerprint range of 20 protostars, with a zoom-in panel of the 7.15–7.5 μm range. The spectra are displayed from top to bottom in a rough order of the intensity of the CH_4 band. The spectra plotted in thinner and lighter lines are the original spectra before removing the gas-phase lines (Sect. 5.3.1). Four vertical dotted lines in gray are plotted at 7.02, 7.245, 7.39, and 7.67 μm, respectively, for comparison with the peak wavelengths of observed absorption bands.

5.4.2 COM ice fingerprint range (6.8–8.8 µm)

5.4.2.1 Detection of ice bands

After going through the multiple steps introduced in Sect. 5.3, we obtained the isolated spectra in the COM ice fingerprint range (6.8–8.8 µm). Noise levels in optical depth (τ_σ) were estimated from the original observed spectra (i.e., those without removing the gas lines) at around 7.0, 7.2, 7.4, 7.7, and 8.2 µm, using channels free from emission or absorption features (listed in Table 5.C.1). The SNR in multiples of σ was then calculated for each absorption band by dividing their $\tau_{\rm peak}$ by τ_σ (listed in Table 5.2).

Figure 5.6 shows the isolated spectra in the COM ice fingerprint range of 20 selected protostars. We excluded five sources from the full sample of 25 sources: four of them are Class I sources (B1-a-NS, TMC1A, TMC1-E, and TMC1-W) which show a weak CH_4 band (i.e., $\tau_{\rm peak} < 0.1$ and $N_{\rm ice}(CH_4) < 10^{17}$ cm^{-2}), and their other absorption bands in this range are barely detected (see Fig. 5.B.6 for their spectra in the 6.8–8.8 µm range). The other source, Per-emb 8, is a deeply embedded Class 0 source with a faint IR continuum and therefore a very low SNR (the lowest for the CH_4 band), although its $\lambda_{\rm peak}$ and $N_{\rm ice}$ of CH_4 are not low.

In the selected sample displayed in Fig. 5.6, the CH_4 band at 7.67 µm is robustly detected in all of them; the two bands at 7.24 and 7.4 µm, which can be attributed to $HCOO^-$, C_2H_5OH, and CH_3CHO, are detected in most of them, except for Ser-SMM1A, Ser-SMM3, and BHR71-IRS2, where the detection is considered tentative due to low SNRs. Qualitatively speaking, these two bands appear together in regular shapes, which will be analyzed more quantitatively in Sect. 5.4.2.2.

The 7.02 µm band is located in a fainter range of the spectrum, on top of the red wing of the strong NH_4^+ band at 6.85 µm; therefore, its shape and intensity are more susceptible to the local continuum than the other three bands mentioned above. The 7.02 µm band can be contributed by CH_3OH, CH_3CHO, C_2H_5OH, and CH_3COCH_3. Its detection is robust for high-SNR sources such as B1-c, Ser-SMM1B, BHR71-IRS1, HOPS 370, and IRAS 20126 where $\tau_{\rm peak} \geq 10\,\tau_\sigma$, but in general, its shape is less regular and its SNRs are lower than those of the bands at 7.24 and 7.40 µm.

The absorption features in the "red part" of the COM fingerprint range (i.e., redward of the CH_4 band, between 7.8 and 8.8 µm) are generally broad and weak, and they are highly dependent on where the local continuum is set. The main influencing factor is the silicate band at 9.8 µm, which lowers the flux, hence the SNR in this range (see the third row in Figs. 5.B.1–5.B.5). Even if the SNR is high, there could be prominent ro-vibrational lines of SiO present between 7.8 and 8.6 µm (e.g., in BHR71-IRS1 and L1448-mm; van Gelder et al. 2024a). In general, the characterization of absorption features in 7.8–8.8 µm is much less reliable than that in 6.8–7.8 µm. In case studies on B1-c (Chapter 4) and IRAS 2A (Rocha et al. 2024), these features are attributed to several ice mixtures, including H_2CO (8.02 µm), HCOOH (8.23 µm), CH_3OCHO (8.26 µm), CH_3OCH_3 (8.63 µm), and potentially CH_3COCH_3 (8.03 µm). In our sample, only several sources in addition to B1-c and IRAS 2A show high-SNR ($> 10\,\tau_\sigma$) and relatively regular absorption bands in this range (IRAS 20126, BHR71-IRS1, HOPS 370, and L1448-IRS2B). A few more show SNRs of $> 5\,\tau_\sigma$ (IRAS 2A, Ser-SMM1B, and L1448-mm), and others mostly show irregular features or have low SNRs in this wavelength range.

5.4.2.2 Band properties I: area

According to Eq. (5.3), the area of an absorption band (i.e., the integral over wavenumber) can be related to the ice column densities of contributing species. Figure 5.7 displays the correlations between the areas of the three bands in the COM ice fingerprint range (i.e., at 7.24, 7.40, and 7.67 µm). We applied the same MCMC fitting used in Sect. 5.4.1, but excluded those sources with low SNRs at each band. The positive correlations between the areas of the three bands echo that seen between the ice column densities of H_2O and CH_4 (Fig. 5.5). Taken together, these results suggest a coherent set of correlations among the presence and strength of different ice bands.

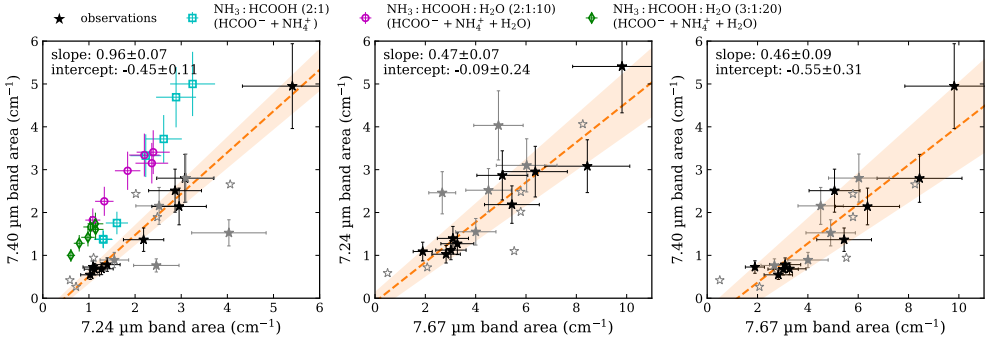


Figure 5.7: Integrated areas of the three bands at 7.24, 7.4, and 7.67 µm. Values measured from observations are denoted by stellar markers in black or gray; markers in hollow gray and solid gray indicate the observed 7.24 or 7.40 µm bands are in low SNRs and contaminated by the gas-phase lines, respectively. In each panel, the best linear fit and its uncertainty are indicated by the dashed line and shaded region in orange, respectively, with the value of the best-fit slope and intercept annotated in the upper left corner. Only the solid points (i.e., those with a robust SNR) were taken into account in the fitting. In the left panel, areas of two $HCOO^-$ bands at 7.24 and 7.39 µm measured from laboratory spectra at different temperatures (between 15 and ∼170 K, see Table 5.E.2) are added for comparison with the observations. Blue, purple, and green correspond to three $HCOO^-$ ices: the pure one and two H_2O-mixed ones with different mixing ratios (see the common legend at the top). The band area uncertainties for the observed and laboratory spectra are set as 20% and 15%, respectively (see explanations in Appendix 5.D).

Since $HCOO^-$ ice has long been suggested as a major contributor to the observed bands at 7.24 and 7.4 µm (Schutte et al. 1999), we took a step further and compared the observations with the laboratory-measured areas of the two $HCOO^-$ bands at ∼7.24 and 7.39 µm (see the left panel of Fig. 5.7). In laboratories, $HCOO^-$ ice is produced through acid-base reactions in HCOOH:NH_3 mixtures. Here we considered pure and two H_2O-mixtures of $HCOO^-$ and measured the peak wavelengths and FWHMs of their 7.24 and 7.39 µm bands from IR spectra (see more details in Appendix 5.E). Figure 5.7 shows that the 7.24 µm band of $HCOO^-$ ice in laboratory spectra is slightly weaker than its 7.39 µm band, but in observations, the 7.24 µm band is the stronger one. This difference in band ratios implies that $HCOO^-$ may not be the only contributor to the two observed bands, and there are likely other species contributing to the 7.24 µm band.

5.4.2.3 Band properties II: peak and width

In order to better analyze the possible compositions of the observed absorption features in the COM ice fingerprint range, we measured the peak wavelengths and FHWMs of four bands observed at 7.02, 7.24, 7.4, and 7.67 μm, and compared them with the same quantities derived from laboratory spectra of different ice mixtures (as displayed in Fig. 5.8). We first discuss the 7.24, 7.4, and 7.67 μm bands, which have been characterized in previous work (e.g., Schutte et al. 1999; Yang et al. 2022), and return to the 7.02 μm band at the end. We skip the analysis of the absorption features in the 7.8–8.8 μm range, which contains multiple ice components that can only be fitted jointly with laboratory spectra. As mentioned at the end of Sect. 5.1, dedicated fitting with laboratory spectra was not performed in this work and will be reserved for future studies.

The 7.24 μm band (top-left panel of Fig. 5.8). The observed peak wavelengths and FWHMs scatter within 7.225–7.265 μm and 0.07–0.14 μm, respectively. The two candidate species contributing to this band are $HCOO^-$ and C_2H_5OH. Both pure and H_2O-mixed $HCOO^-$ ices match the observations in terms of peak wavelength, but the FWHM of the H_2O-mixed one fits better. Here we only plotted the more diluted mixture of $HCOO^-$ (NH_3:HCOOH:H_2O = 3:1:20), given that the values are similar for the other mixture (NH_3:HCOOH:H_2O = 3:1:10).

In addition to $HCOO^-$, C_2H_5OH ice at low temperatures, either in pure form or mixed with other common interstellar ices (e.g., H_2O and CH_3OH), match the observations well in terms of peak wavelength and FWHM. Only the C_2H_5OH:CO mixture and the crystalline C_2H_5OH (at high temperatures) are excluded. This is consistent with the fitting results of the COM ice fingerprint range of B1-c (Chapter 4), which prefers H_2O and CH_3OH mixtures of O-COMs, whereas CO mixtures usually have narrower band widths (see Fig. 4.4). However, it should be noted that C_2H_5OH has several other bands with similar strengths in 6.8–7.1 μm and at ∼7.5 μm, which would constrain its contribution to the 7.24 μm band in quantitative fitting with laboratory spectra.

The 7.4 μm band (top-right panel of Fig. 5.8). The properties of this band are also quite scattered in observations. The peak wavelengths vary between ∼7.38 and ∼7.42 μm (or 7.4 μm if not considering low-SNR sources), and the FWHMs vary between 0.05 and 0.12 μm (or between 0.06 and 0.11 μm excluding low-SNR sources). The $HCOO^-$:H_2O mixture (labeled as HCOOH:NH_3:H_2O) fits the observations well in both peak wavelength and FWHM, and the band of pure $HCOO^-$ is too redshifted, though with an appropriate FWHM. As for the other candidate species CH_3CHO, there is an evident difference in peak wavelength and FWHM compared with the observations. The CH_3CHO band is generally narrower and more redshifted; only the H_2O mixture at low temperatures and the CO+CH_3OH mixture tend to fit the observations. However, even though CH_3CHO is not likely to be the dominant contributor to the 7.4 μm band, it could still be present in tandem with $HCOO^-$, given its potential contribution to another band at 7.02 μm.

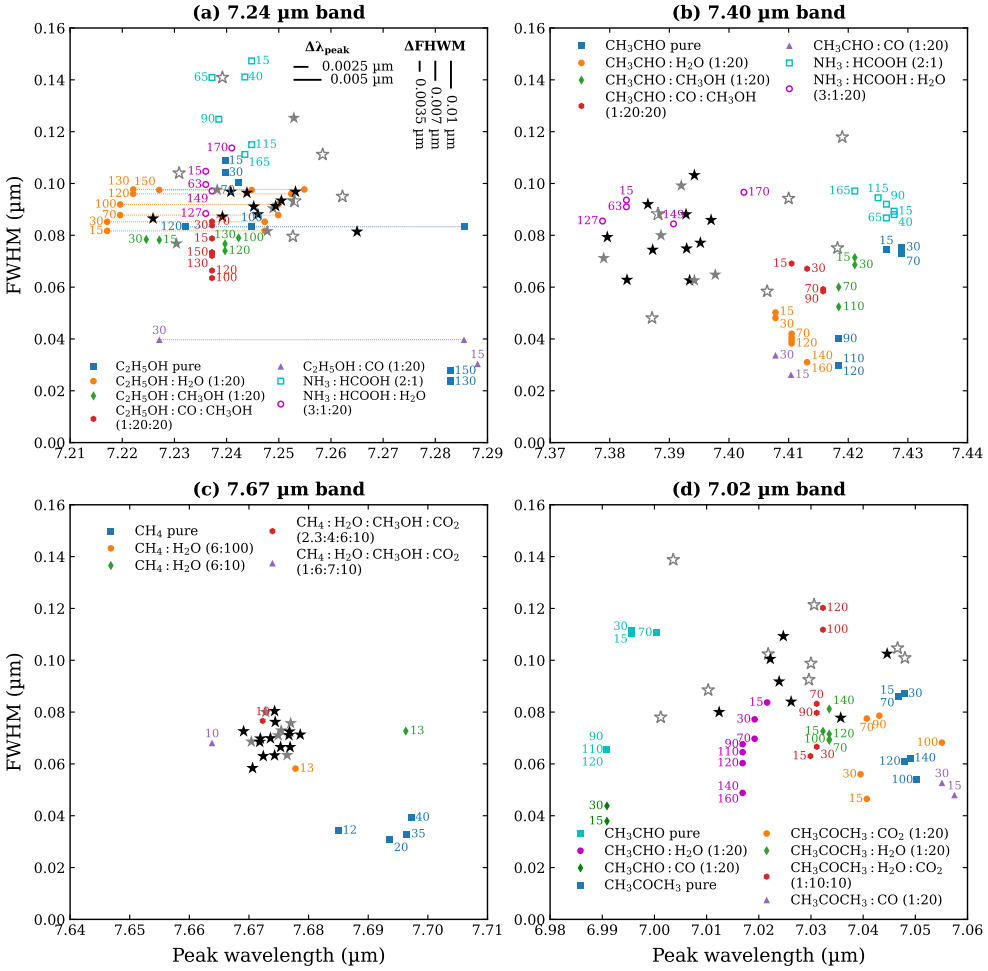


Figure 5.8: Peak wavelength versus FWHM of the absorption bands at (a) 7.24 μm, (b) 7.42 μm, (c) 7.67 μm, and (d) 7.02 μm in the JWST/MIRI-MRS spectra and the laboratory spectra of C_2H_5OH, CH_3CHO, and $HCOO^-$ ices (labeled as NH_3:HCOOH mixtures). Values measured from laboratory spectra are denoted by multi-shape markers in different colors (see legend in each panel), and the number next to each marker indicates the temperature in Kelvin. The laboratory data are taken from Terwisscha van Scheltinga et al. (2018) for CH_3CHO and C_2H_5OH, Rachid et al. (2020) for CH_3COCH_3, and this work (Appendix 5.E) for $HCOO^-$. The legend convention for observational data points is the same as in Fig. 5.7. In the laboratory spectra of some C_2H_5OH ice mixtures, the band around 7.24 μm is split into two, which are indicated by two data points linked with dotted lines. To avoid clutter, representative uncertainty bars in peak wavelength (0.0025 and 0.005 μm) and FWHM (0.0035, 0.007, and 0.01 μm) are indicated in the upper right corner of panel (a). The observational data points have $\Delta\lambda_{peak} = 0.005$ μm and $\Delta FWHM = 0.01$ μm; the experimental data points have $\Delta\lambda_{peak} = 0.0025$ μm and $\Delta FWHM = 0.0035$ μm for $HCOO^-$ (panels a, b) and CH_3COCH_3 ices (panel d), and $\Delta\lambda_{peak} = 0.005$ μm and $\Delta FWHM = 0.007$ μm for C_2H_5OH (panel a), CH_3CHO (panels b, d), and CH_4 ices (panel c). See Appendix 5.D for discussion of uncertainty estimation.

The 7.67 µm band (bottom-left panel of Fig. 5.8). This band shows very consistent peak wavelengths (mostly 7.67–7.68 µm) and FWHMs (mostly 0.06–0.08 µm) in our sample, suggesting a simple contribution from CH_4 ice. Although only a few mixtures of CH_4 have been studied in laboratories, the comparisons already show clear evidence that CH_4 ice in protostellar envelopes is not pure but mixed with (or trapped within) other common interstellar molecules such as H_2O, CO_2, and CH_3OH. This also shows a good example that systematic analyses of band properties are very helpful for identifying the contributing ice species and their mixing environments.

The 7.02 µm band (bottom-right panel of Fig. 5.8). This band is less constrained than the other three bands and shows more scattered peak wavelengths in observations (varying from 7.0 to 7.05 µm). The FHWMs are generally larger (0.08–0.14 µm), consistent with the broad profiles displayed in Fig. 5.6. There are several potential contributors to this band, including CH_3OH, CH_3CHO, C_2H_5OH, and CH_3COCH_3. Only CH_3CHO (the CH_3 deformation mode at ∼6.995 µm) and CH_3COCH_3 (the CH_3 asymmetric deformation mode at ∼7.05 µm) have reported peak wavelengths and FWHMs in the literature (Terwisscha van Scheltinga et al. 2018; Rachid et al. 2020). This is largely because the absorption bands of CH_3OH and C_2H_5OH at around 7.0 µm are wide and have multiple peaks, making it tricky to determine their peak positions and widths. In reality, the CH_3OH band at ∼6.9 µm is much stronger than the nearby bands of other COMs given its dominant ice abundance. For a similar reason, we did not measure the laboratory band properties at ∼7.0 µm for the CH_3OH-mixed CH_3CHO, C_2H_5OH, and CH_3COCH_3 ices, since their bands at ∼7.0 µm are blended with the much stronger CH_3OH band and are therefore very difficult to isolate; it is not impossible, but it would be very tricky, and the results would not be reliable. Nevertheless, the band properties of their CH_3OH mixtures should not differ too much from those of their other mixtures (e.g., with H_2O, CO_2, and CO). The comparison with observations shows that the H_2O mixtures of CH_3CHO and CH_3COCH_3 are promising contributors to this band. Although the FWHM of each of them is slightly smaller than the observations, they can create broader band profiles by mixing with one another.

In general, only the 7.67 µm band is contributed by one ice species and shows very consistent peak wavelengths and FWHMs. The other three bands at 7.02, 7.24, and 7.4 µm are more scattered in their properties, which is a reasonable result of being composed of more than one ice band. The favored contributing species of each band are consistent with the decomposition results in case studies (Rocha et al. 2024; Chapter 4). Although no spectral fitting was performed in this work, there is still a lot of important information and validation that can be learned from these systematic analyses.

5.5 Conclusions

In this paper, we present the first overview study of ice absorption features in the COM fingerprint range (6.8–8.8 µm) along with the silicate 9.8 µm band and the H_2O liberation band using the JWST/MIRI-MRS spectra of 25 protostellar sources taken from the JOYS+ and IPA programs. To facilitate large-sample analyses, we developed

an interactive tool that integrates the steps of fitting the global continuum and the absorption bands of silicates and H_2O. After isolating the COM ice fingerprint range, we measured the band properties of four absorption bands at 7.02, 7.24, 7.40, and 7.67 µm, and compared them with those measured from laboratory spectra of ice mixtures with plausible contributions. We summarize our conclusions as follows:

- The H_2O ice column densities and optical depths of the silicate 9.8 µm band show a strong positive correlation, which reinforces their role in tracing dust and ice reservoirs in protostellar systems.

- A strong positive correlation is also found between the ice column densities of H_2O and CH_4, indicating that the CH_4 band is also a reliable tracer of interstellar ice and can serve as a reference band in future analyses.

- There is a clear separation between Class 0 and Class I objects in the silicate optical depths and the ice column densities of H_2O and CH_4, which supports previous observational and theoretical results that protostars are losing their envelopes and becoming more exposed as they evolve.

- The 7.67 µm band is the strongest and most ubiquitous band in the COM ice fingerprint range. The two bands at 7.24 and 7.4 µm are also well recognized in most Class 0 and Class 0/I sources, and they tend to appear together. The 7.02 µm band and absorption features beyond 7.8 µm are sensitive to the local continuum subtraction and are only considered robustly detected in high-SNR sources.

- The integrated areas of the three bands at 7.24, 7.40, and 7.67 µm are positively correlated with each other, suggesting that these absorption features arise from ice mantles on dust grains and can reflect the total ice reservoir to some extent.

- The 7.67 µm band has remarkably consistent peak wavelengths and FWHMs, suggesting that it is primarily attributed to CH_4 ice. In contrast, the three absorption bands at 7.02, 7.24, and 7.4 µm, which can be attributed to species such as $HCOO^-$, C_2H_5OH, CH_3CHO, and CH_3COCH_3, show considerable variation in their band properties, implying more than one composition.

- By comparing the peak wavelengths and FWHMs of the absorption bands measured from observations and laboratory spectra, we found that $HCOO^-$ mixed with H_2O reproduces the 7.24 and 7.4 µm bands best. However, the difference in band area ratios suggests additional contributors to the 7.24 µm band (e.g., C_2H_5OH). This highlights the spectral complexity of COM ice features and the need for more laboratory data, particularly for ice mixtures under astrophysical conditions.

Overall, this is the first study that applies a systematic analysis to the COM ice fingerprint range in a large sample of protostellar sources. Our results will provide an essential benchmark for interpreting the weak ice features in protostars, and help unravel the chemical inventory in the early stages of star and planet formation.

Acknowledgements

This work is based on observations made with the NASA/ESA/CSA James Webb Space Telescope. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. Astrochemistry in Leiden is supported by the Netherlands Research School for Astronomy (NOVA), by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 101019751 MOLDISK), by the Dutch Research Council (NWO) grants TOP-1 614.001.751 and 618.000.001, and by the Danish National Research Foundation through the Center of Excellence "InterCat" (Grant agreement no.: DNRF150). The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. A.C.G. acknowledges support from PRIN-MUR 2022 20228JPA3A "The path to star and planet formation in the JWST era (PATH)" funded by NextGeneration EU and by INAF-GoG 2022 "NIR-dark Accretion Outbursts in Massive Young stellar objects (NAOMY)" and Large Grant INAF 2022 "YSOs Outflows, Disks and Accretion: towards a global framework for the evolution of planet forming systems (YODA)". This work benefited from the use of GitHub Copilot, an AI-based tool, in the development of the interactive fitting tool introduced in Appendix 5.A.

Appendix

5.A Interactive fitting tool

Figure 5.A.1 shows the interactive interface of the fitting tool mentioned in Sect. 5.3.2. The interface is composed of three display panels and a few control buttons and sliding bars in the bottom. In the Panel 1, the observed spectrum is plotted as an SED (in solid gray and black lines; the gray one shows the original spectrum, and the black one shows the spectrum with gas-phase lines removed). The SED spectra shown in blue and red are those with the opacity of silicates and H_2O added back to the observed spectrum (spectra with gas-phase lines retained are plotted in lighter colors). This is the core of this fitting tool, that is, we use the "opacity-corrected" spectrum to provide a clearer guidance of where to set (the guiding points of) the global continuum. Users can adjust the scaling of the silicate and the H_2O spectrum in Panel 2 and 3, respectively, using the sliding bars in the slider panel. Both the silicate and the H_2O bands can be fitted with two selected components (two compositions of silicates and two temperatures of H_2O). The adjustment of the silicate and/or the H_2O fitting will be updated in real time to the opacity-corrected spectra shown in Panel 1.

After finding a plausible fitting for silicates and H_2O, users can manually add (or remove) guiding points to (or from) the canvas in Panel 1 by pressing specific keys on the keyboard. The existing guiding points can also be moved to a different location by dragging with a pressed mouse. After clicking the "Start Cont Fit" button in the control panel (which will become "Stop Cont Fit" after clicking), a spline function will be fitted to the guiding points to create a global continuum (dashed green line). The smoothing factor of the spline function can be tuned using the sliding bar named

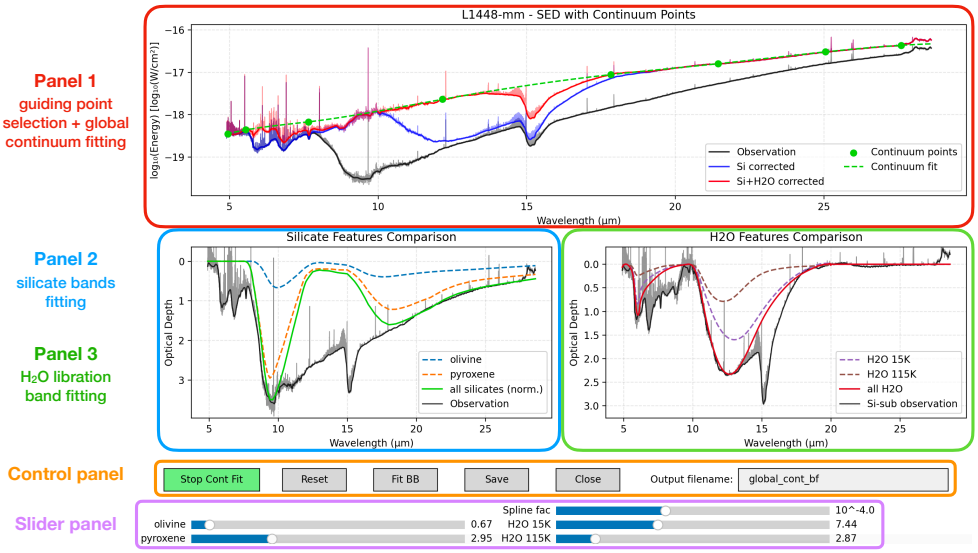


Figure 5.A.1: Interface of the interactive fitting tool.

"Spline fac" in the slider panel. In the meantime, the observed spectrum in optical depth will be displayed in Panel 2, and the silicate-subtracted spectrum in Panel 3 (in solid gray and black lines). The users will know how their preliminary fitting of silicates and H_2O looks like with the current global continuum, and make further adjustments accordingly to achieve the best fit. All the relevant elements shown in the three display panels will be updated in real time if any adjustments are made.

If the users want to start from scratch, they can click "Reset" or close and rerun the interface by clicking "Close" in the control panel. If a satisfactory fit is reached, the users can save the fitting results including all the elements apart from the input observed spectrum into an output file with a customized filename (using the text box in the control panel). The users can also obtain a blackbody fitting for the global continuum (the dashed green line) by clicking the "Fit BB" button.

5.B Additional figures

Figures 5.B.1–5.B.6 show the same content as Fig. 5.2 but for other sources studied in this Chapter. In particular, Fig. 5.B.6 shows the sources without prominent ice features in the COM ice fingerprint range (see the third row therein).

5.C Table of band properties in observations

Table 5.C.1 lists the properties of the four observed absorption bands in the COM ice fingerprint range at 7.02, 7.24, 7.40, and 7.67 μm. There are six columns for each band: the peak wavelength ($\lambda_{\rm peak}$), the FWHM, the peak optical depth ($\tau_{\rm peak}$), the noise level in optical depth (τ_σ), the integrated area, and the fitting method (i.e., fitted with a Gaussian function or by definition).

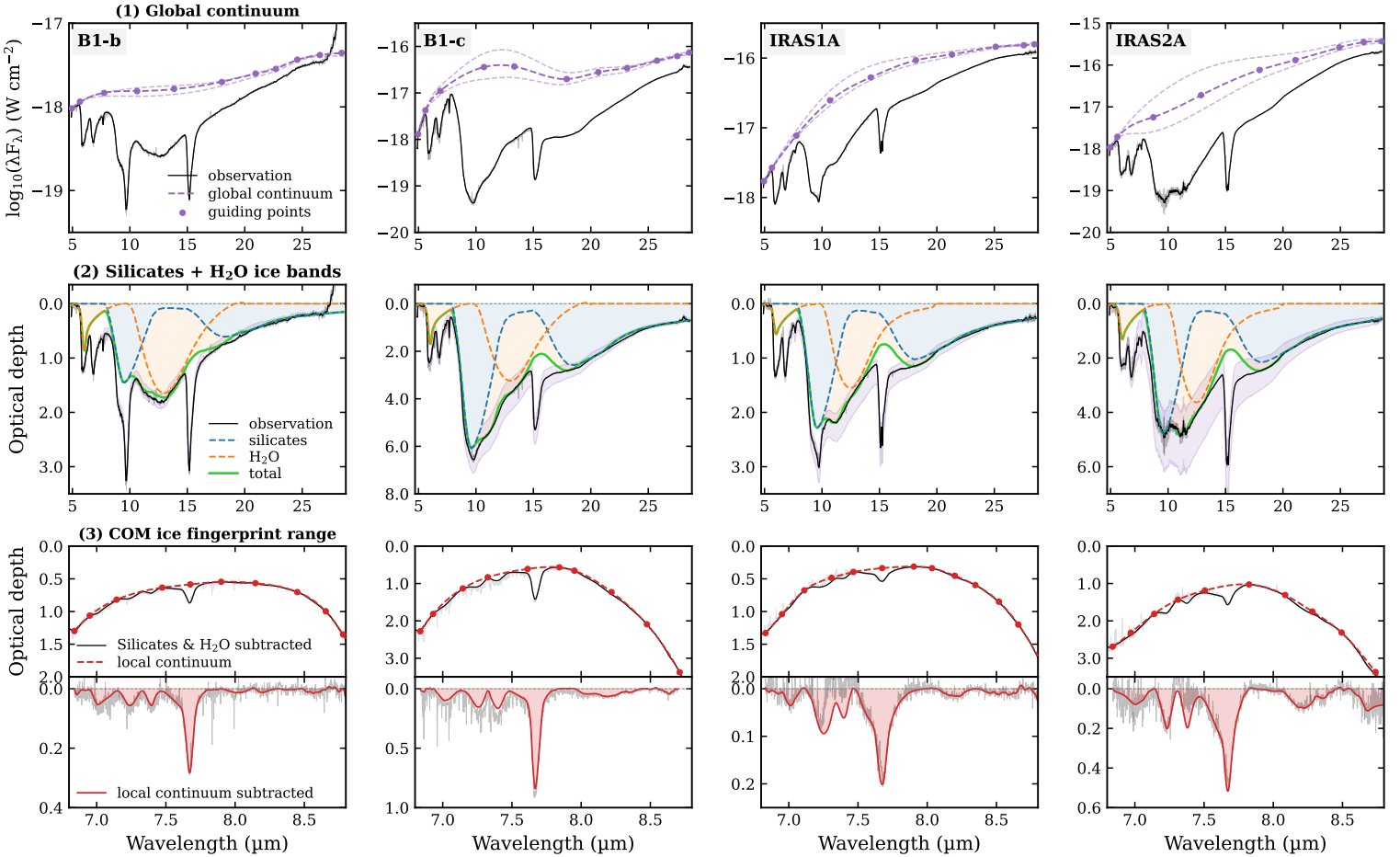


Figure 5.B.1: Same figure as Fig. 5.2 but for B1-b, B1-c, NGC 1333 IRAS 1A and NGC 1333 IRAS 2A.

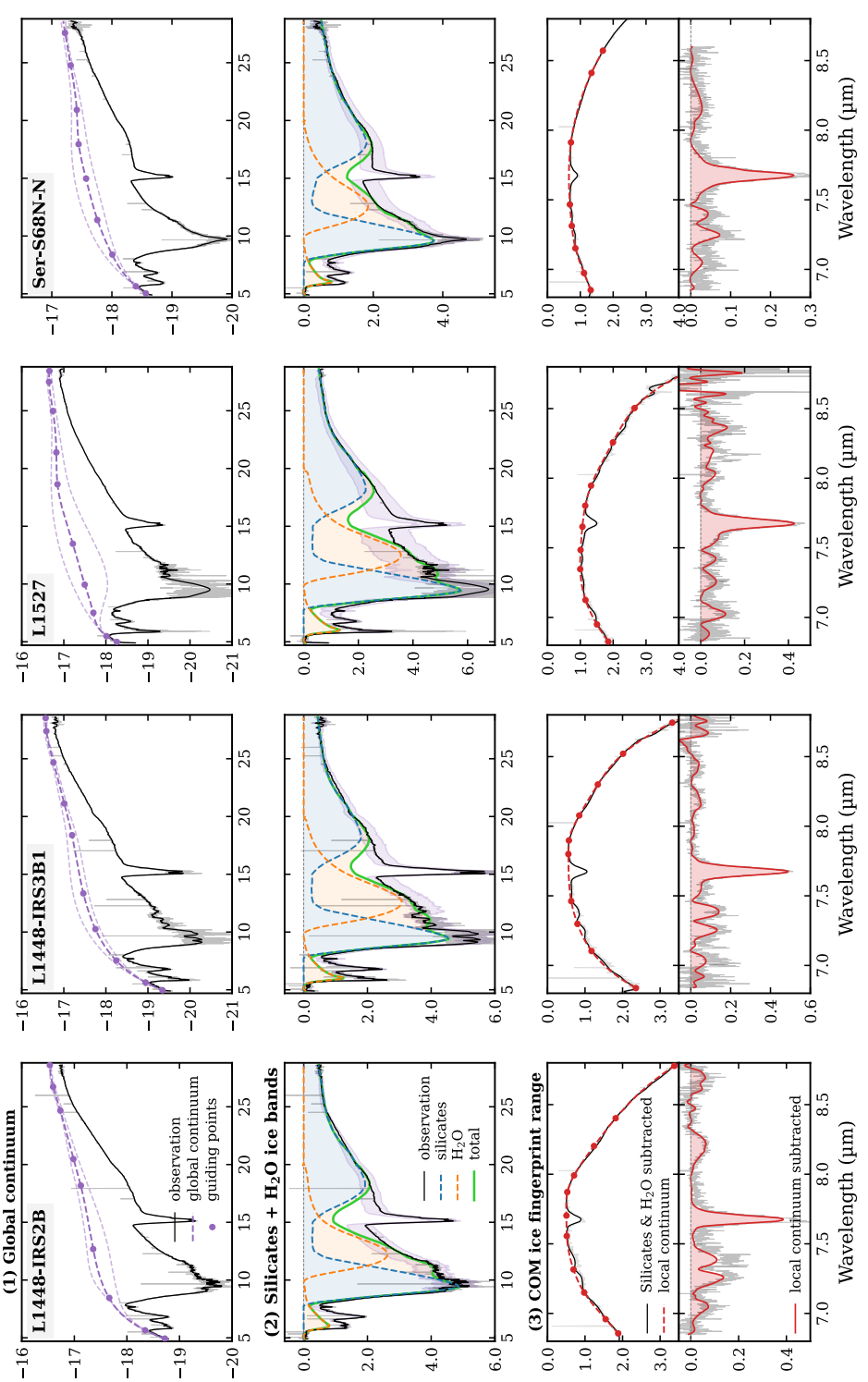


Figure 5.B.2: Same figure as Fig. 5.2 but for L1448-IRS2B, L1448-IRS3B1, L1527, and Ser-S68N-N.

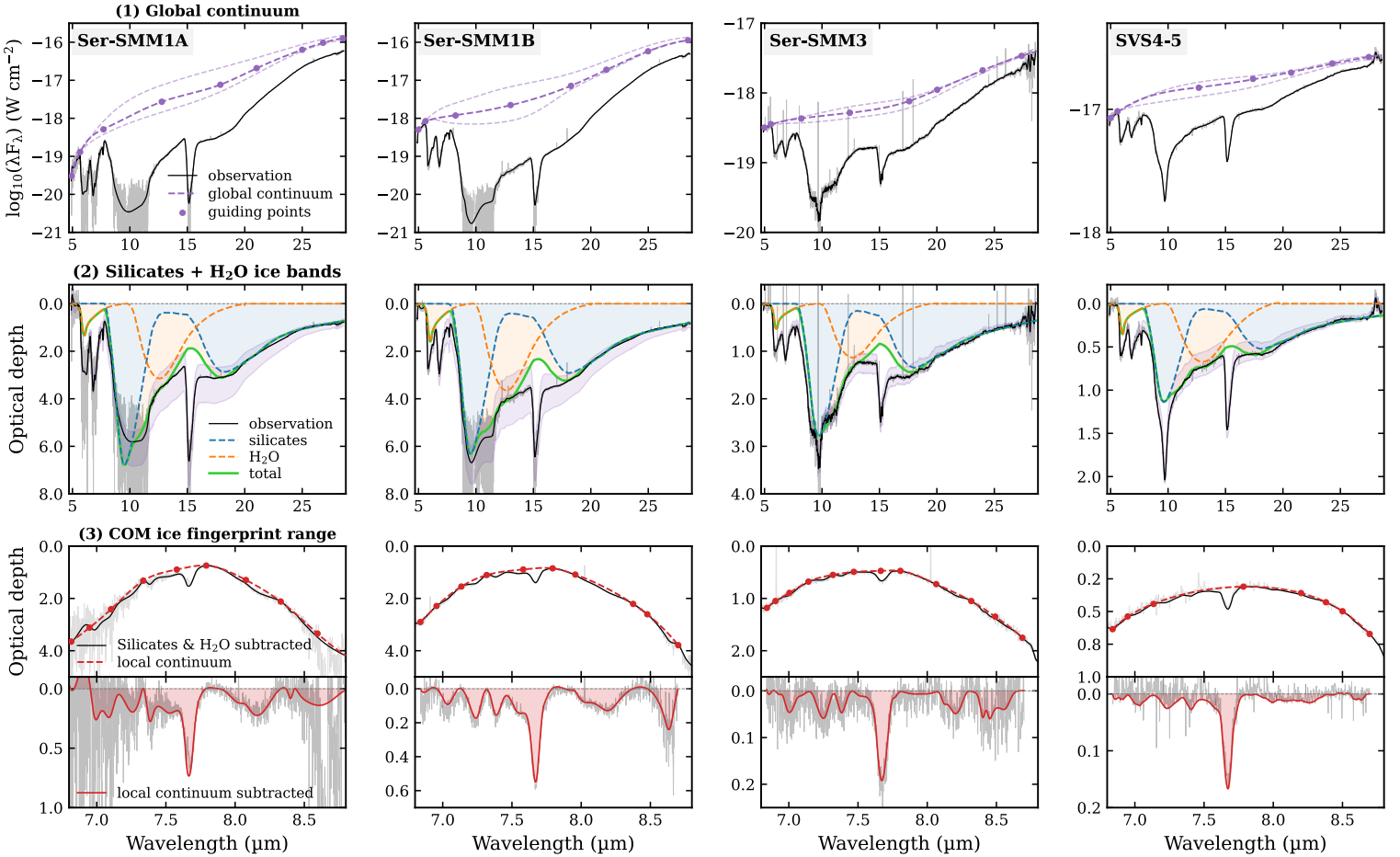


Figure 5.B.3: Same figure as Fig. 5.2 but for Ser-SMM1A, Ser-SMM1B, Ser-SMM3, and SVS4-5.

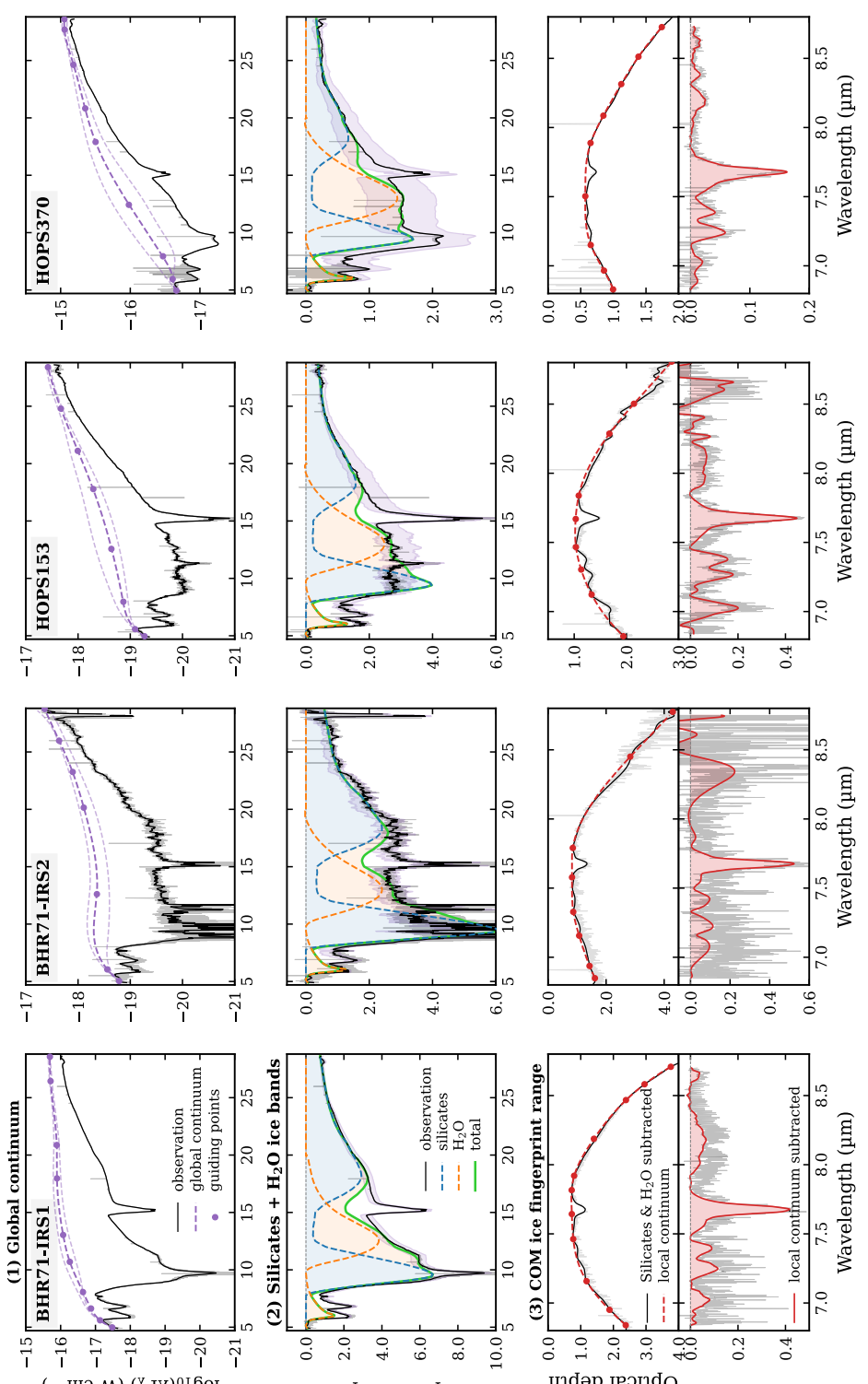


Figure 5.B.4: Same figure as Fig. 5.2 but for BHR71-IRS1, BHR71-IRS2, HOPS 153, and HOPS 370.

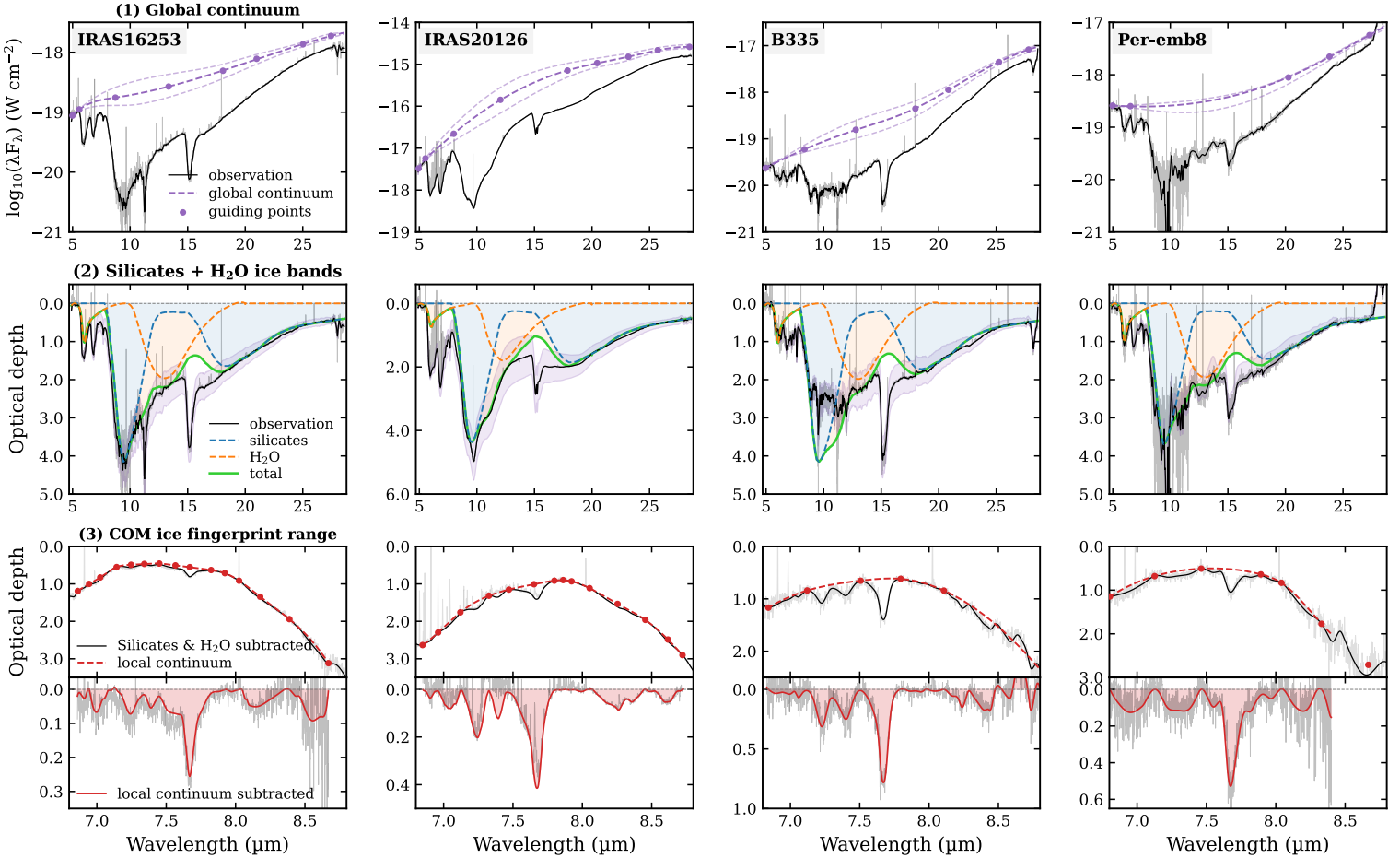


Figure 5.B.5: Same figure as Fig. 5.2 but for IRAS 16253, IRAS 20126, B335, and Per-emb 8.

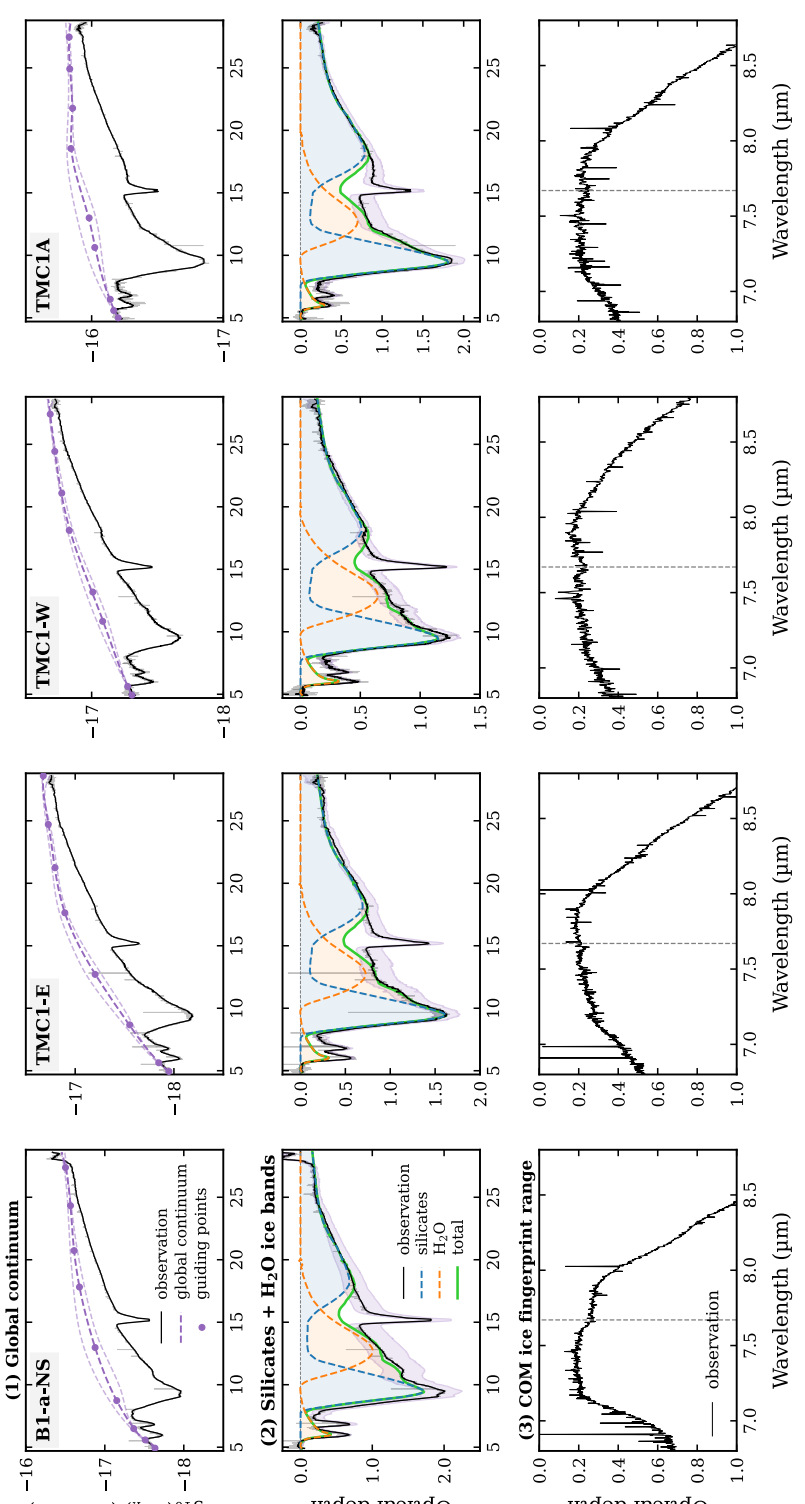


Figure 5.B.6: Same figure as Fig. 5.2 but for B1-a-NS, TMC1-E, TMC1-W, and TMC1A, The third row shows that these four sources do not have prominent CH_4 ice band at 7.67 μm (position indicated by vertical dashed gray lines) and other absorption features in the COM ice fingerprint range.

Table 5.C.1: Band properties of the observed absorption bands at 7.02, 7.24, 7.40, and 7.67 µm.

Source	7.02 µm band						7.24 µm band						7.40 µm band						7.67 µm band					
	λ_{peak} (µm)	FWHM (µm)	τ_{peak}	τ_σ *	area (cm^{-1})	fit†	λ_{peak} (µm)	FWHM (µm)	τ_{peak}	τ_σ	area (cm^{-1})	fit	λ_{peak} (µm)	FWHM (µm)	τ_{peak}	τ_σ	area (cm^{-1})	fit	λ_{peak} (µm)	FWHM (µm)	τ_{peak}	τ_σ	area (cm^{-1})	fit
B1-a-NS	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	7.6743	0.0856	<0.05	3.6(-3)	<0.74	G
B1-b	7.0306	0.1214	0.05	1.1(-2)	1.29	G	7.2408	0.0968	0.06	8.3(-3)	1.12	G	7.3934	0.0626	0.06	8.1(-3)	0.68	G	7.6724	0.0630	0.26	7.3(-3)	3.02	G
B1-c	7.0247	0.1093	0.10	3.1(-3)	2.22	G	7.2532	0.0968	0.16	3.1(-3)	3.08	G	7.3970	0.0859	0.17	3.1(-3)	2.80	G	7.6706	0.0584	0.80	3.1(-3)	8.44	G
IRAS 1A	7.0124	0.0800	0.04	5.4(-3)	0.61	G	7.2529	0.1253	0.10	5.4(-3)	2.46	G	7.3977	0.0648	0.06	5.4(-3)	0.77	G	7.6728	0.0800	0.18	2.8(-3)	2.67	G
IRAS 2A	7.0036	0.1388	0.08	1.7(-2)	2.36	G	7.2304	0.0768	0.20	1.7(-2)	3.10	G	7.3790	0.0712	0.20	1.7(-2)	2.80	D	7.6754	0.0729	0.46	8.1(-3)	6.03	G
L1448-mm	–	–	–	–	–	–	7.2478	0.0816	0.10	1.1(-2)	1.55	D	7.3886	0.0800	0.06	1.1(-2)	0.89	D	7.6749	0.0710	0.31	1.1(-2)	4.01	G
L1448-IRS2B	7.0466	0.1047	0.07	6.9(-3)	1.62	G	7.2489	0.0904	0.14	5.7(-3)	2.52	G	7.3920	0.0992	0.11	5.7(-3)	2.15	G	7.6764	0.0634	0.39	4.4(-3)	4.51	G
L1448-IRS3B1	7.0296	0.0925	0.08	1.6(-2)	1.48	G	7.2527	0.0796	0.13	1.2(-2)	2.02	G	7.3881	0.0882	0.15	1.2(-2)	2.43	G	7.6753	0.0665	0.48	7.5(-3)	5.79	G
Per-emb 8	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	7.6779	0.0875	0.53	7.2(-2)	8.34	G
L1527	7.0262	0.0840	0.11	1.4(-2)	2.07	G	7.2622	0.0950	0.06	1.4(-2)	1.10	G	7.4182	0.0751	0.07	1.4(-2)	0.94	G	7.6768	0.0725	0.42	1.2(-2)	5.53	G
TMC1-E	–	–	–	–	–	–	7.2530	0.0932	0.03	9.9(-3)	0.59	G	7.4190	0.1179	0.02	9.9(-3)	0.42	G	7.6500	0.0641	0.04	4.3(-3)	0.50	G
TMC1-W	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	7.6664	0.0688	0.06	4.7(-3)	0.77	G
TMC1A	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	7.6608	0.0899	0.05	4.5(-3)	0.75	G
Ser-S68N-N	7.0480	0.1009	0.03	1.0(-2)	0.62	G	7.2494	0.0913	0.07	9.4(-3)	1.27	G	7.3928	0.0881	0.05	9.4(-3)	0.69	G	7.6744	0.0762	0.24	9.7(-3)	3.27	G
Ser-SMM1A	–	–	–	–	–	–	7.2584	0.1112	0.20	8.5(-2)	4.06	D	7.3871	0.0481	0.28	8.5(-2)	2.66	G	7.6704	0.0686	0.67	3.2(-2)	8.26	G
Ser-SMM1B	7.0357	0.0778	0.09	8.5(-3)	1.57	G	7.2392	0.0872	0.17	7.2(-3)	2.95	D	7.3872	0.0744	0.16	7.2(-3)	2.14	D	7.6718	0.0685	0.51	2.9(-3)	6.37	G
Ser-SMM3	7.0103	0.0885	0.04	3.2(-2)	0.85	G	7.2460	0.0881	0.06	2.5(-2)	1.03	G	7.3829	0.0628	0.05	1.7(-2)	0.55	G	7.6743	0.0804	0.19	1.6(-2)	2.83	G
SVS4-5	7.0300	0.0988	0.02	1.5(-3)	0.43	G	7.2392	0.1409	0.03	1.5(-3)	0.72	G	7.4064	0.0584	0.02	1.5(-3)	0.27	G	7.6719	0.0698	0.16	4.8(-3)	2.08	G
HOPS 153	7.0239	0.0918	0.20	2.8(-2)	3.92	G	7.2650	0.0814	0.17	2.8(-2)	2.87	G	7.3796	0.0793	0.16	2.6(-2)	2.51	G	7.6769	0.0665	0.42	9.4(-3)	5.06	G
HOPS 370	7.0218	0.1024	0.02	4.8(-4)	0.42	G	7.2452	0.0911	0.06	4.8(-4)	1.09	G	7.3864	0.0920	0.04	4.8(-4)	0.73	D	7.6786	0.0713	0.15	1.1(-3)	1.90	G
BHR71-IRS1	7.0222	0.1005	0.08	3.4(-3)	1.66	G	7.2505	0.0934	0.12	3.4(-3)	2.18	G	7.3929	0.0749	0.09	3.4(-3)	1.37	G	7.6767	0.0711	0.42	2.6(-3)	5.44	G
BHR71-IRS2	–	–	–	–	–	–	7.2309	0.1040	0.12	5.0(-2)	2.48	G	7.4100	0.0942	0.10	5.0(-2)	1.89	G	7.6743	0.0633	0.51	3.4(-2)	5.79	G
B335	–	–	–	–	–	–	7.2259	0.0865	0.31	3.9(-2)	5.41	G	7.3942	0.1032	0.25	3.9(-2)	4.95	G	7.6736	0.0699	0.78	1.5(-2)	9.82	G
IRAS 16253	7.0012	0.0780	0.07	1.5(-2)	1.13	G	7.2439	0.0964	0.07	1.0(-2)	1.40	G	7.3952	0.0771	0.05	1.0(-2)	0.78	G	7.6691	0.0726	0.24	3.5(-3)	3.10	G
IRAS 20126	7.0446	0.1025	0.08	4.6(-3)	1.74	G	7.2382	0.0976	0.20	5.3(-3)	4.03	G	7.3942	0.0626	0.13	5.3(-3)	1.53	G	7.6770	0.0757	0.36	1.9(-3)	4.91	G

* τ_σ are shown in the form $a(b)$, which denotes $a \times 10^b$.

† The "fit" column lists how the band was fitted and the properties were derived: G = by Gaussian fitting, D = by definition of peak and FWHM.

5.D Uncertainty estimation of band properties

5.D.1 Observed spectra

As mentioned in Sect. 5.3.4, the uncertainties in the Gaussian-fitted peak wavelength and FWHM are relatively small, only on the order of 10^{-4} µm. For sources with decent SNRs, the peak wavelength uncertainty ($\Delta\lambda_{\rm peak}$) is about (1–2)×10^{-4} µm, and the FWHM uncertainty (ΔFWHM) is about (3–5)×10^{-4} µm. For those with low SNRs, $\Delta\lambda_{\rm peak}$ and ΔFWHM can reach up to ∼5×10^{-4} and ∼1×10^{-3} µm, respectively. However, larger uncertainties should have been introduced by removing the gas-phase lines (Sect. 5.3.1) and subtracting the local continuum (Sect. 5.3.3), which cannot be quantified analytically. Although the actual uncertainties differ from source to source, we adopt a uniform and conservative estimate for $\Delta\lambda_{\rm peak}$ and ΔFWHM, which are 0.005 µm and 0.01 µm (i.e., one order of magnitude larger than the typical uncertainties of the Gaussian fitting), respectively. The estimated $\Delta\lambda_{\rm peak}$ and ΔFWHM for observed bands are denoted as black bars in Fig. 5.8(a); they are also shown as shaded-regions in Fig. 5.3, which look reasonable.

The uncertainty in band area is even more complicated to estimate. Assuming a Gaussian profile, the band area can be calculated analytically from

$$\text{area} = \int_{\tilde{\nu}_1}^{\tilde{\nu}_2} \tau(\tilde{\nu})\,\mathrm{d}\tilde{\nu} = \sqrt{2\pi}\,\tau_{\rm peak}\sigma = \sqrt{\frac{\pi}{4\ln 2}}\,\tau_{\rm peak}\,\text{FWHM}, \tag{5.4}$$

where $\tau_{\rm peak}$ is the peak optical depth, and $\sigma = \text{FWHM}/2\sqrt{\ln 2}$. The error propagation then gives

$$\frac{\Delta\text{area}}{\text{area}} = \sqrt{\left(\frac{\Delta\tau_{\rm peak}}{\tau_{\rm peak}}\right)^2 + \left(\frac{\Delta\text{FWHM}}{\text{FWHM}}\right)^2}. \tag{5.5}$$

The relative uncertainty of FWHM, when considering the conservative estimate (0.01 µm), is around 10–20%. The relative uncertainty of $\Delta\tau_{\rm peak}$ obtained from Gaussian fitting itself is small (well below 1%). If we estimate $\Delta\tau_{\rm peak}/\tau_{\rm peak}$ to be 10%, the relative uncertainty of band area would be around 20%., which is adopted in Fig. 5.7 for data points in observations.

5.D.2 Laboratory spectra

We assume that the uncertainties in the peak wavelength and FWHM of laboratory spectra are dominated by measurement uncertainties (i.e., the spectral resolution of experiments). Conventionally, the laboratory spectra are measured in wavenumber, so we need to convert the uncertainties from wavenumber to wavelength via $\Delta\lambda = \lambda^2\Delta\tilde{\nu}$. For CH_3CHO and C_2H_5OH ices, the spectral resolution $\Delta\tilde{\nu}$ is 1 cm^{-1} (Terwisscha van Scheltinga et al. 2018), which corresponds to $\Delta\lambda \sim$ 0.005–0.006 µm in 7–8 µm. The spectral resolution for $HCOO^-$ is better (Appendix 5.E.1), with $\Delta\tilde{\nu} = 0.5$ cm^{-1} and the corresponding $\Delta\lambda$ ∼0.0025–0.003 µm. The absorption band widths of the laboratory spectra are measured based on the definition of FWHM (i.e., FWHM = $\tilde{\nu}_2 - \tilde{\nu}_1$) instead of fitting a Gaussian function, so the propagated uncertainty is given by $\Delta\text{FWHM}_{\tilde{\nu}} = \sqrt{\Delta\tilde{\nu}_1^2 + \Delta\tilde{\nu}_2^2} = \sqrt{2}\Delta\tilde{\nu}$. The factor of $\sqrt{2}$ (or 1.41) is the same between $\Delta\text{FWHM}_\lambda$ and $\Delta\lambda$. As a result, the FWHM uncertainty is 0.007–0.0084 for CH_3CHO

and C_2H_5OH spectra, and 0.0035–0.0042 µm for $HCOO^-$ ones. The uncertainty bars for laboratory spectra in Fig. 5.8(a) correspond to 0.003 and 0.006 for $\Delta\lambda$, and 0.0084 and 0.0042 for ΔFWHM.

According to Eq. (5.5), the relative uncertainty in band area depends on the relative uncertainties in peak optical depth ($\tau_{\rm peak}$) and FWHM, of which the latter is generally smaller than 10%. Usually, $\tau_{\rm peak}$ is not measured and reported by experimentalists, as it is proportional to the column densities of the deposited ices, but its relative uncertainty should be fairly small (e.g., smaller than observational uncertainties) given the high SNR of laboratory spectra. The main uncertainty of $\tau_{\rm peak}$ should come from the baseline correction (see Appendix 5.E.2), but would still not exceed 10–20%. As a result, a relative uncertainty of 15% should be a reasonable conservative estimate for band areas measured in laboratory spectra, and is adopted in the left panel of Fig. 5.7.

5.E Laboratory spectra of $HCOO^-$ ices

5.E.1 Experiments

Laboratory IR spectra of ice mixtures of NH_3:HCOOH (2:1) and NH_3:HCOOH:H_2O (2:1:10 and 3:1:20) were collected using the InfraRed Absorption Setup for Ice Spectroscopy (IRASIS) in the Leiden Laboratory for Astrophysics. The setup, calibration, and deposition methods are described in Slavicinska et al. (2025a) and references therein. Briefly, three dosing lines with independent leak valves were calibrated individually to deposit either pure H_2O, NH_3, or HCOOH on a KBr substrate held at 15 K. Ice deposition rates for each species were chosen so that, once all three leak valves were opened simultaneously, the deposited ice mixture would have the desired mixing ratio. Deposition rates during calibrations (described in Appendix A in Slavicinska et al. 2025a) were determined via laser interferometry, using the ice refractive indexes and densities presented in Table 5.E.1.

Table 5.E.1: Ice refractive indexes and densities used in the calibration.

Molecule	n	ρ (g cm^{-3})	Reference
H_2O	1.234	0.719	Yarnall & Hudson (2022)
NH_3	1.33	0.68	Hudson & Yarnall (2022)
HCOOH	1.291	0.979	Hudson & Yarnall (2022)

HCOOH was chosen as the limiting reactant in the acid-base reaction NH_3 + HCOOH → NH_4HCOO to minimize the contribution of neutral HCOOH to the CH bending mode at 7.24 µm, a shared feature between HCOOH and $HCOO^-$. The formation of some NH_4HCOO was already observed during all of the ice mixture depositions at 15 K, although strong features at 5.85 µm (HCOOH C=O stretch) and 8.22 µm (HCOOH C–O stretch) indicated that the reaction remained far from complete at this temperature. After deposition, the ice mixture was heated at a rate of 1 K min^{-1}, during which the peaks characteristic of NH_4HCOO grew while the peaks of the neutral reactants diminished. IR spectra (0.5 cm^{-1} resolution, 256 scans per spectrum) of the ice mixture were continuously collected throughout heating, resulting in an average ice temperature difference of ∼3.6 K between each spectrum.

5.E.2 Baseline correction and band property measurements

The laboratory IR spectra of the aforementioned ice mixtures were baseline corrected by fitting a 5th- to 7th-order polynomial to manually selected wavelength ranges free of absorption features. Figure 5.E.1 shows the original and baseline-corrected spectra of NH_3:HCOOH (2:1) and NH_3:HCOOH:H_2O (3:1:20). The baseline correction was applied only to the ∼6.0–8.0 µm range, where the two $HCOO^-$ bands at 7.24 and 7.39 µm are located. The peak wavelength and FWHM of the two $HCOO^-$ bands were measured from the baseline-corrected spectra of three ice mixtures and at different temperatures (see Table 5.E.2).

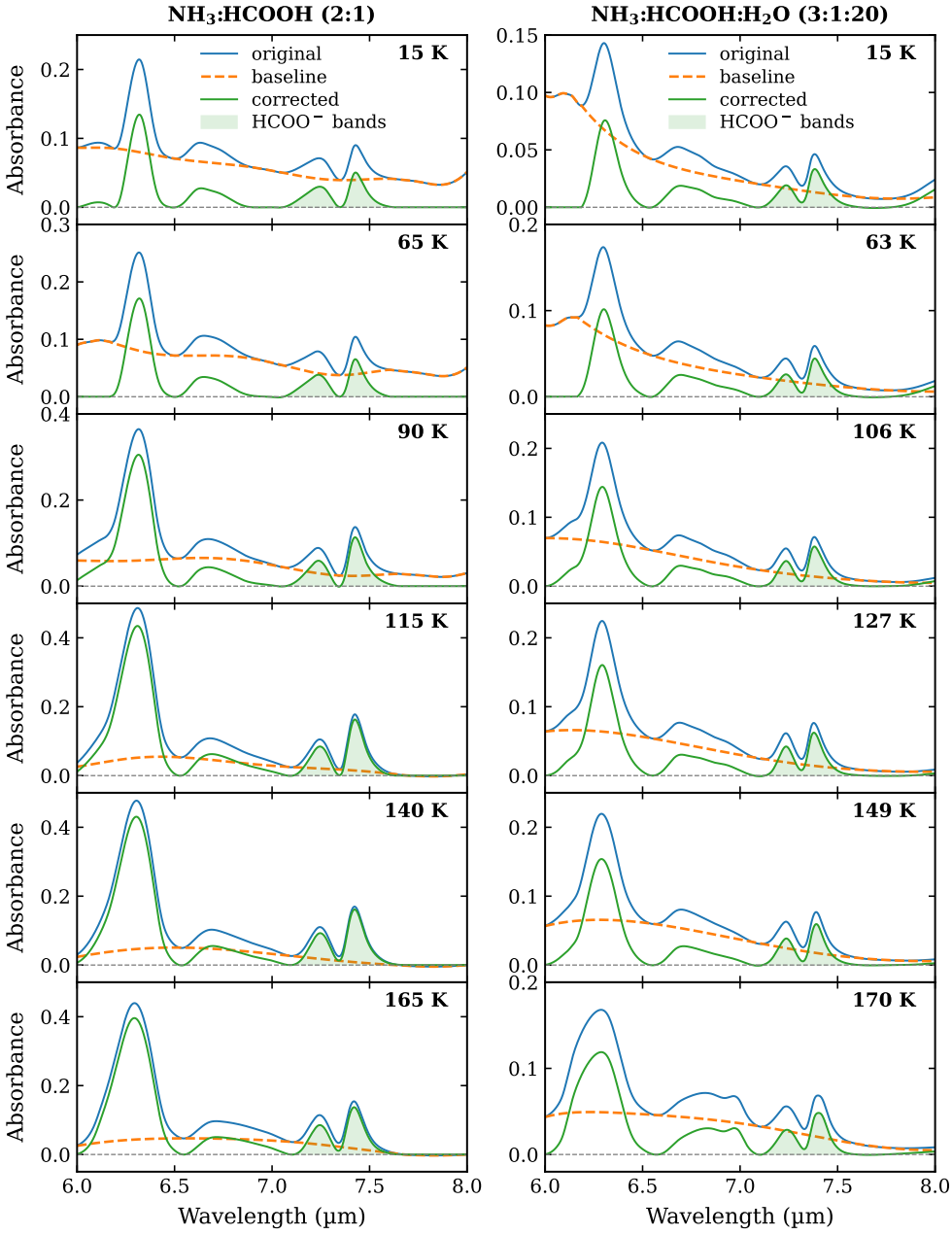


Figure 5.E.1: Baseline correction of the laboratory IR spectra of two ice mixtures, NH_3:HCOOH = 2:1 (left column) and NH_3:HCOOH:H_2O = 3:1:20 (right column). In each panel, the original and baseline-corrected spectra are plotted in blue and green, respectively. The polynomial baseline is shown as the dashed orange line. The two $HCOO^-$ bands of interest are shaded in green. The measurement temperature is indicated in the upper right corner.

Table 5.E.2: Band properties of $HCOO^-$ bands at 7.24 and 7.39 µm in three ice mixtures.

NH_3:HCOOH (2:1)					**NH_3:HCOOH:H_2O (2:1:20)**					**NH_3:HCOOH:H_2O (3:1:10)**				
T (K)	peak λ (µm)	FWHM (µm)	peak λ (µm)	FWHM (µm)	T (K)	peak λ (µm)	FWHM (µm)	peak λ (µm)	FWHM (µm)	T (K)	peak λ (µm)	FWHM (µm)	peak λ (µm)	FWHM (µm)
15	7.2448	0.1473	7.4277	0.0893	15	7.2360	0.1058	7.3867	0.0950	15	7.2360	0.1047	7.3828	0.0936
65	7.2372	0.1409	7.4264	0.0867	55	7.2347	0.1034	7.3854	0.0937	63	7.2360	0.0996	7.3828	0.0910
90	7.2385	0.1260	7.4264	0.0920	98	7.2360	0.0984	7.3841	0.0936	106	7.2360	0.0921	7.3814	0.0909
115	7.2448	0.1150	7.4250	0.0945	120	7.2372	0.0972	7.3841	0.0923	127	7.2360	0.0884	7.3788	0.0855
140	7.2473	0.1164	7.4237	0.0998	141	7.2372	0.0959	7.3867	0.0910	149	7.2372	0.0971	7.3907	0.0844
165	7.2435	0.1112	7.4211	0.0971	163	7.2435	0.1176	7.3999	0.0939	170	7.2410	0.1137	7.4025	0.0966

Bibliography

Allamandola, L. J., Sandford, S. A. & Valero, G. J., 1988, ICARUS, 76, 2, 225–252

Allodi, M. A., Baragiola, R. A., Baratta, G. A. et al., 2013, Space Sci. Rev., 180, 1-4, 101–175

André, P., Ward-Thompson, D. & Barsony, M., 1993, ApJ, 406, 122

André, P., Ward-Thompson, D. & Barsony, M., 2000, V. Mannings, A. P. Boss & S. S. Russell, editors, Protostars and Planets IV, 59

André, P., Di Francesco, J., Ward-Thompson, D. et al., 2014, H. Beuther, R. S. Klessen, C. P. Dullemond & T. Henning, editors, Protostars and Planets VI, 27–51

Argyriou, I., Glasse, A., Law, D. R. et al., 2023, A&A, 675, A111

Bacmann, A., Taquet, V., Faure, A. et al., 2012, A&A, 541, L12

Baek, G., Lee, J.-E., Hirota, T. et al., 2022, ApJ, 939, 2, 84

Balucani, N., Ceccarelli, C. & Taquet, V., 2015, MNRAS, 449, L16–L20

Belloche, A., Müller, H. S. P., Menten, K. M. et al., 2013, A&A, 559, A47

Belloche, A., Garrod, R. T., Müller, H. S. P. & Menten, K. M., 2014, Science, 345, 6204, 1584–1587

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

Belloche, A., Maury, A. J., Maret, S. et al., 2020, A&A, 635, A198

Belloche, A., Garrod, R. T., Zingsheim, O. et al., 2022, A&A, 662, A110

Belloche, A., Garrod, R. T., Müller, H. S. P. et al., 2025, A&A, 698, A143

Beltrán, M. T., Codella, C., Viti, S. et al., 2009, ApJ, 690, 2, L93–L96

Beltrán, M. T., Sánchez-Monge, Á., Cesaroni, R. et al., 2014, A&A, 571, A52

Beltrán, M. T., Cesaroni, R., Rivilla, V. M. et al., 2018, A&A, 615, A141

Bergantini, A., Abplanalp, M. J., Pokhilko, P. et al., 2018, ApJ, 860, 2, 108

Bergin, E. A. & Tafalla, M., 2007, ARA&A, 45, 1, 339–396

Bertin, M., Romanzin, C., Doronin, M. et al., 2016, ApJ, 817, 2, L12

Beuther, H., Churchwell, E. B., McKee, C. F. & Tan, J. C., 2007, B. Reipurth, D. Jewitt & K. Keil, editors, Protostars and Planets V, 165

Beuther, H., Kuiper, R. & Tafalla, M., 2025, arXiv e-prints, arXiv:2501.16866

Bianchi, E., Chandler, C. J., Ceccarelli, C. et al., 2020, MNRAS, 498, 1, L87–L92

Bisschop, S. E., Jørgensen, J. K., van Dishoeck, E. F. & de Wachter, E. B. M., 2007, A&A, 465, 3, 913–929

Biver, N., Bockelée-Morvan, D., Boissier, J. et al., 2021, A&A, 648, A49

Blake, G. A., Sutton, E. C., Masson, C. R. & Phillips, T. G., 1987, ApJ, 315, 621

Bonfand, M., Csengeri, T., Bontemps, S. et al., 2024, A&A, 687, A163

Bonnell, I. A. & Bate, M. R., 2006, MNRAS, 370, 1, 488–494

Bonnell, I. A., Bate, M. R., Clarke, C. J. & Pringle, J. E., 2001, MNRAS, 323, 4, 785–794

Boogert, A. C. A., Schutte, W. A., Tielens, A. G. G. M. et al., 1996, A&A, 315,

L377–L380
Boogert, A. C. A., Schutte, W. A., Helmich, F. P. et al., 1997, A&A, 317, 929–941
Boogert, A. C. A., Ehrenfreund, P., Gerakines, P. A. et al., 2000, A&A, 353, 349–362
Boogert, A. C. A., Pontoppidan, K. M., Knez, C. et al., 2008, ApJ, 678, 2, 985–1004
Boogert, A. C. A., Gerakines, P. A. & Whittet, D. C. B., 2015, ARA&A, 53, 541–581
Booth, A. S., Walsh, C., Terwisscha van Scheltinga, J. et al., 2021, Nature Astronomy, 5, 684–690
Booth, A. S., Law, C. J., Temmink, M. et al., 2023, A&A, 678, A146
Booth, A. S., Leemker, M., van Dishoeck, E. F. et al., 2024a, AJ, 167, 4, 164
Booth, A. S., Temmink, M., van Dishoeck, E. F. et al., 2024b, AJ, 167, 4, 165
Booth, A. S., Wölfer, L., Temmink, M. et al., 2025, ApJ, 986, 1, L9
Bottinelli, S., Ceccarelli, C., Lefloch, B. et al., 2004, ApJ, 615, 1, 354–358
Bottinelli, S., Boogert, A. C. A., Bouwman, J. et al., 2010, ApJ, 718, 2, 1100–1117
Boudin, N., Schutte, W. A. & Greenberg, J. M., 1998, A&A, 331, 749–759
Bouilloud, M., Fray, N., Bénilan, Y. et al., 2015, MNRAS, 451, 2, 2145–2160
Bouvier, M., Ceccarelli, C., López-Sepulcre, A. et al., 2022, ApJ, 929, 1, 10
Brogan, C. L., Hunter, T. R. & Fomalont, E. B., 2018, arXiv e-prints, arXiv:1805.05266
Brown, R. D., Crofts, J. G., Gardner, F. F. et al., 1975, ApJ, 197, L29–L31
Brown, R. D., Godfrey, P. D., McNaughton, D. et al., 1990, Journal of Molecular Spectroscopy, 140, 2, 340–352
Brunken, N. G. C., Booth, A. S., Leemker, M. et al., 2022, A&A, 659, A29
Brunken, N. G. C., Rocha, W. R. M., van Dishoeck, E. F. et al., 2024a, A&A, 685, A27
Brunken, N. G. C., van Dishoeck, E. F., Slavicinska, K. et al., 2024b, A&A, 692, A163
Brunken, N. G. C., van Dishoeck, E. F., Slavicinska, K. et al., 2024c, A&A, 692, A163
Brunken, N. G. C., Boogert, A. C. A., van Dishoeck, E. F. et al., 2025, ACS Earth and Space Chemistry, 9, 8, 1992–2003
Burbidge, E. M., Burbidge, G. R., Fowler, W. A. & Hoyle, F., 1957, Reviews of Modern Physics, 29, 4, 547–650
Busch, L. A., Belloche, A., Garrod, R. T. et al., 2024, A&A, 681, A104
Bushouse, H., Eisenhamer, J., Dencheva, N. et al., 2023, JWST Calibration Pipeline, doi: 10.5281/zenodo.10022973
Calcutt, H., Jørgensen, J. K., Müller, H. S. P. et al., 2018, A&A, 616, A90
Calcutt, H., Willis, E. R., Jørgensen, J. K. et al., 2019, A&A, 631, A137
Carr, J. S., 1989, ApJ, 345, 522–535
Caselli, P. & Ceccarelli, C., 2012, A&A Rev., 20, 56
Caselli, P., Walmsley, C. M., Tafalla, M. et al., 1999, ApJ, 523, 2, L165–L169
Cazaux, S., Tielens, A. G. G. M., Ceccarelli, C. et al., 2003, ApJ, 593, 1, L51–L55
Cazzoli, G. & Puzzarini, C., 2008, A&A, 487, 3, 1197–1202
Ceccarelli, C., Caselli, P., Bockelée-Morvan, D. et al., 2014, H. Beuther, R. S. Klessen, C. P. Dullemond & T. Henning, editors, Protostars and Planets VI, 859–882
Ceccarelli, C., Codella, C., Balucani, N. et al., 2023, S. Inutsuka, Y. Aikawa, T. Muto, K. Tomida & M. Tamura, editors, Protostars and Planets VII, volume 534 of *Astronomical Society of the Pacific Conference Series*, 379

Cernicharo, J., Agúndez, M., Cabezas, C. et al., 2021, A&A, 649, L15
Cernicharo, J., Agúndez, M., Cabezas, C. et al., 2022, European Physical Journal Web of Conferences, volume 265 of *European Physical Journal Web of Conferences*, 00041
Chahine, L., López-Sepulcre, A., Neri, R. et al., 2022, A&A, 657, A78
Charnley, S. B., Tielens, A. G. G. M. & Millar, T. J., 1992, ApJ, 399, L71
Charnley, S. B., Kress, M. E., Tielens, A. G. G. M. & Millar, T. J., 1995, ApJ, 448, 232
Chibueze, J. O., Omodaka, T., Handa, T. et al., 2014, ApJ, 784, 2, 114
Christiaens, V., Gonzalez, C., Farkas, R. et al., 2023, The Journal of Open Source Software, 8, 81, 4774
Chuang, K. J., Fedoseev, G., Ioppolo, S. et al., 2016, MNRAS, 455, 2, 1702–1712
Chuang, K. J., Fedoseev, G., Qasim, D. et al., 2017, MNRAS, 467, 3, 2552–2565
Chuang, K. J., Fedoseev, G., Qasim, D. et al., 2020, Astronomy and Astrophysics, 635, A199
Chuang, K. J., Fedoseev, G., Scirè, C. et al., 2021, A&A, 650, A85
Codella, C., López-Sepulcre, A., Ohashi, S. et al., 2022, MNRAS, 515, 1, 543–554
Coletta, A., Fontani, F., Rivilla, V. M. et al., 2020, A&A, 641, A54
Collings, M. P., Anderson, M. A., Chen, R. et al., 2004, MNRAS, 354, 4, 1133–1140
Colzi, L., Rivilla, V. M., Beltrán, M. T. et al., 2021, A&A, 653, A129
Coutens, A., Jørgensen, J. K., van der Wiel, M. H. D. et al., 2016, A&A, 590, L6
Crapsi, A., van Dishoeck, E. F., Hogerheijde, M. R. et al., 2008, A&A, 486, 1, 245–254
Crouzet, N., Mueller, M., Sargent, B. et al., 2025, A&A, 698, A77
Cruz-Diaz, G. A., Martín-Doménech, R., Muñoz Caro, G. M. & Chen, Y. J., 2016, A&A, 592, A68
Csengeri, T., Belloche, A., Bontemps, S. et al., 2019, A&A, 632, A57
Csengeri, T., Wyrowski, F., Menten, K. M. et al., 2022, A&A, 658, A193
Cummins, S. E., Linke, R. A. & Thaddeus, P., 1986, ApJS, 60, 819
Cuppen, H. M., van Dishoeck, E. F., Herbst, E. & Tielens, A. G. G. M., 2009, A&A, 508, 1, 275–287
Cuppen, H. M., Linnartz, H. & Ioppolo, S., 2024, ARA&A, 62, 1, 243–286
Cyganowski, C. J., Brogan, C. L., Hunter, T. R. et al., 2011, ApJ, 729, 2, 124
Danielson, R. E., Woolf, N. J. & Gaustad, J. E., 1965, ApJ, 141, 116
Dartois, E., Noble, J. A., Ysard, N. et al., 2022, A&A, 666, A153
Dartois, E., Noble, J. A., Caselli, P. et al., 2024, Nature Astronomy
de Graauw, T., Whittet, D. C. B., Gerakines, P. A. et al., 1996, A&A, 315, L345–L348
De Simone, M., Ceccarelli, C., Codella, C. et al., 2020, ApJ, 896, 1, L3
D'Hendecourt, L. B. & Jourdain de Muizon, M., 1989, A&A, 223, L5–L8
Dominik, C., Min, M. & Tazaki, R., 2021, OpTool: Command-line driven tool for creating complex dust opacities, Astrophysics Source Code Library, record ascl:2104.010
Dorschner, J., Begemann, B., Henning, T. et al., 1995, A&A, 300, 503
Drozdovskaya, M. N., van Dishoeck, E. F., Rubin, M. et al., 2019, MNRAS, 490, 1, 50–79

Drozdovskaya, M. N., Coudert, L. H., Margulès, L. et al., 2022, A&A, 659, A69
Dulieu, F., Amiaud, L., Congiu, E. et al., 2010, A&A, 512, A30
Ehrenfreund, P., Gerakines, P. A., Schutte, W. A. et al., 1996, A&A, 312, 263–274
Ehrenfreund, P., Boogert, A. C. A., Gerakines, P. A. et al., 1997, A&A, 328, 649–669
El-Abd, S. J., Brogan, C. L., Hunter, T. R. et al., 2019, ApJ, 883, 2, 129
Endres, C. P., Schlemmer, S., Schilke, P. et al., 2016, Journal of Molecular Spectroscopy, 327, 95–104
Enoch, M. L., Evans, N. J., II, Sargent, A. I. & Glenn, J., 2009, ApJ, 692, 2, 973–997
Evans, N. J., II, Allen, L. E., Blake, G. A. et al., 2003, PASP, 115, 810, 965–980
Evans, N. J., II, Dunham, M. M., Jørgensen, J. K. et al., 2009, ApJS, 181, 2, 321–350
Fabricant, B., Krieger, D. & Muenter, J. S., 1977, J. Chem. Phys., 67, 4, 1576–1586
Faúndez, S., Bronfman, L., Garay, G. et al., 2004, A&A, 426, 97–103
Fayolle, E. C., Öberg, K. I., Cuppen, H. M. et al., 2011, A&A, 529, A74
Fayolle, E. C., Öberg, K. I., Jørgensen, J. K. et al., 2017, Nature Astronomy, 1, 703–708
Fedele, D., van den Ancker, M. E., Henning, T. et al., 2010, A&A, 510, A72
Federman, S. A., Megeath, S. T., Rubinstein, A. E. et al., 2024, ApJ, 966, 1, 41
Fedoseev, G., Cuppen, H. M., Ioppolo, S. et al., 2015, MNRAS, 448, 2, 1288–1297
Fedoseev, G., Chuang, K. J., Ioppolo, S. et al., 2017, ApJ, 842, 1, 52
Fedoseev, G., Qasim, D., Chuang, K.-J. et al., 2022, ApJ, 924, 2, 110
Ferrer Asensio, J., Spezzano, S., Coudert, L. H. et al., 2023, A&A, 670, A177
Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J., 2013, PASP, 125, 925, 306
Francis, L., van Gelder, M. L., van Dishoeck, E. F. et al., 2024, A&A, 683, A249
Francis, L., van Dishoeck, E. F., Caratti o Garatti, A. et al., 2025, A&A, 694, A174
Fuchs, G. W., Cuppen, H. M., Ioppolo, S. et al., 2009, A&A, 505, 2, 629–639
Fuente, A., Cernicharo, J., Caselli, P. et al., 2014, A&A, 568, A65
Garrod, R. T., 2013, ApJ, 765, 1, 60
Garrod, R. T. & Herbst, E., 2006, A&A, 457, 3, 927–936
Garrod, R. T. & Pauly, T., 2011, ApJ, 735, 1, 15
Garrod, R. T., Jin, M., Matis, K. A. et al., 2022, The Astrophysical Journal Supplement Series, 259, 1, 1
Ge, J. X., He, J. H., Chen, X. & Takahashi, S., 2014, MNRAS, 445, 2, 1170–1185
Geballe, T. R., 1986, A&A, 162, 248–252
Gerakines, P., Yarnall, Y. & Hudson, R., 2022, American Astronomical Society Meeting Abstracts, volume 54 of *American Astronomical Society Meeting Abstracts*, 134.02
Gerakines, P. A., Schutte, W. A., Greenberg, J. M. & van Dishoeck, E. F., 1995, A&A, 296, 810
Gibb, E. L., Whittet, D. C. B., Boogert, A. C. A. & Tielens, A. G. G. M., 2004, ApJS, 151, 1, 35–73
Gieser, C., Semenov, D., Beuther, H. et al., 2019, A&A, 631, A142
Gillett, F. C. & Forrest, W. J., 1973, ApJ, 179, 483
Goldsmith, P. F., 2001, ApJ, 557, 2, 736–746

Grant, S. L., van Dishoeck, E. F., Tabone, B. et al., 2023, ApJ, 947, 1, L6
GRAVITY Collaboration, Caratti o Garatti, A., Fedriani, R. et al., 2020, A&A, 635, L12
Greene, T. P., Wilking, B. A., Andre, P. et al., 1994, ApJ, 434, 614
Greenfield, P. & Miller, T., 2016, Astronomy and Computing, 16, 41–53
Grim, R. J. A. & Greenberg, J. M., 1987, ApJ, 321, L91
Grim, R. J. A., Baas, F., Geballe, T. R. et al., 1991, A&A, 243, 473
Groner, P., Albert, S., Herbst, E. et al., 2002, ApJS, 142, 1, 145–151
Hagen, W., Tielens, A. G. G. M. & Greenberg, J. M., 1983, A&AS, 51, 389–416
Haisch, K. E., Jr., Lada, E. A. & Lada, C. J., 2001, ApJ, 553, 2, L153–L156
Henning, T. K. & Krasnokutski, S. A., 2019, Nature Astronomy, 3, 568–573
Herbst, E. & Klemperer, W., 1973, ApJ, 185, 505–534
Herbst, E. & van Dishoeck, E. F., 2009, ARA&A, 47, 1, 427–480
Hernández, J., Hartmann, L., Megeath, T. et al., 2007, ApJ, 662, 2, 1067–1081
Hidaka, H., Watanabe, M., Kouchi, A. & Watanabe, N., 2011, Physical Chemistry Chemical Physics (Incorporating Faraday Transactions), 13, 35, 15798
Hiraoka, K., Miyagoshi, T., Takayama, T. et al., 1998, ApJ, 498, 2, 710–715
Hollis, J. M., Lovas, F. J. & Jewell, P. R., 2000, ApJ, 540, 2, L107–L110
Hollis, J. M., Lovas, F. J., Jewell, P. R. & Coudert, L. H., 2002, ApJ, 571, 1, L59–L62
Horn, A., Møllendal, H., Sekiguchi, O. et al., 2004, ApJ, 611, 1, 605–614
Hsieh, T.-H., Murillo, N. M., Belloche, A. et al., 2019, ApJ, 884, 2, 149
Hsu, S.-Y., Liu, S.-Y., Liu, T. et al., 2022, ApJ, 927, 2, 218
Hudgins, D. M., Sandford, S. A., Allamandola, L. J. & Tielens, A. G. G. M., 1993, ApJS, 86, 713
Hudson, R. L. & Ferrante, R. F., 2020, MNRAS, 492, 1, 283–293
Hudson, R. L. & Gerakines, P. A., 2019, MNRAS, 485, 1, 861–871
Hudson, R. L. & Yarnall, Y. Y., 2022, ICARUS, 377, 114899
Hudson, R. L., Gerakines, P. A. & Moore, M. H., 2014, ICARUS, 243, 148–157
Hudson, R. L., Loeffler, M. J. & Yocum, K. M., 2017, ApJ, 835, 2, 225
Hudson, R. L., Gerakines, P. A. & Ferrante, R. F., 2018, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 193, 33–39
Hudson, R. L., Gerakines, P. A., Yarnall, Y. Y. & Coones, R. T., 2021, ICARUS, 354, 114033
Ilyushin, V. V., Müller, H. S. P., Jørgensen, J. K. et al., 2022, A&A, 658, A127
Imai, M., Oya, Y., Svoboda, B. et al., 2022, ApJ, 934, 1, 70
Ioppolo, S., van Boheemen, Y., Cuppen, H. M. et al., 2011, MNRAS, 413, 3, 2281–2287
Ioppolo, S., Fedoseev, G., Chuang, K. J. et al., 2021, Nature Astronomy, 5, 197–205
Ishibashi, A., Molpeceres, G., Hidaka, H. et al., 2024, ApJ, 976, 2, 162
Isokoski, K., Bottinelli, S. & van Dishoeck, E. F., 2013, A&A, 554, A100
Issac, N., Tej, A., Liu, T. et al., 2020, MNRAS, 497, 4, 5454–5472
Jeans, J. H., 1902, Philosophical Transactions of the Royal Society of London Series A, 199, 1–53
Jiménez-Serra, I., Vasyunin, A. I., Caselli, P. et al., 2016a, ApJ, 830, 1, L6
Jiménez-Serra, I., Vasyunin, A. I., Caselli, P. et al., 2016b, ApJ, 830, 1, L6

Jiménez-Serra, I., Rodríguez-Almeida, L. F., Martín-Pintado, J. et al., 2022, A&A, 663, A181
Jin, M. & Garrod, R. T., 2020, ApJS, 249, 2, 26
Johnson, H. R. & Strandberg, M. W. P., 1952, J. Chem. Phys., 20, 4, 687–695
Johnstone, D., Boonman, A. M. S. & van Dishoeck, E. F., 2003, A&A, 412, 157–174
Jørgensen, J. K., Bourke, T. L., Myers, P. C. et al., 2005, ApJ, 632, 2, 973–981
Jørgensen, J. K., van der Wiel, M. H. D., Coutens, A. et al., 2016, A&A, 595, A117
Jørgensen, J. K., Müller, H. S. P., Calcutt, H. et al., 2018, A&A, 620, A170
Jørgensen, J. K., Belloche, A. & Garrod, R. T., 2020, ARA&A, 58, 727–778
Keane, J. V., Boogert, A. C. A., Tielens, A. G. G. M. et al., 2001, A&A, 375, L43–L46
Kemper, F., Vriend, W. J. & Tielens, A. G. G. M., 2004, ApJ, 609, 2, 826–837
Kennicutt, R. C. & Evans, N. J., 2012, ARA&A, 50, 531–608
Kessler, M. F., Steinz, J. A., Anderegg, M. E. et al., 1996, A&A, 315, 2, L27–L31
Knacke, R. F., Cudaback, D. D. & Gaustad, J. E., 1969, ApJ, 158, 151
Knacke, R. F., McCorkle, S., Puetter, R. C. et al., 1982, ApJ, 260, 141–146
König, C., Urquhart, J. S., Csengeri, T. et al., 2017, A&A, 599, A139
Krasnokutski, S. A., Kuhn, M., Renzler, M. et al., 2016, ApJ, 818, 2, L31
Kristensen, L. E., van Dishoeck, E. F., Bergin, E. A. et al., 2012, A&A, 542, A8
Kruczkiewicz, F., Dulieu, F., Ivlev, A. V. et al., 2024, A&A, 686, A236
Labiano, A., Argyriou, I., Álvarez-Márquez, J. et al., 2021, A&A, 656, A57
Lacy, J. H., Baas, F., Allamandola, L. J. et al., 1984, ApJ, 276, 533–543
Lada, C. J., 1987, M. Peimbert & J. Jugaku, editors, Star Forming Regions, volume 115 of *IAU Symposium*, 1
Lamberts, T., Fedoseev, G., van Hemert, M. C. et al., 2022, ApJ, 928, 1, 48
Larson, R. B., 1969, MNRAS, 145, 271
Law, D. R., E. Morrison, J., Argyriou, I. et al., 2023, AJ, 166, 2, 45
Lee, C.-F., Codella, C., Li, Z.-Y. & Liu, S.-Y., 2019a, ApJ, 876, 1, 63
Lee, J.-E., Lee, S., Baek, G. et al., 2019b, Nature Astronomy, 3, 314–319
Leger, A., Gauthier, S., Defourneau, D. & Rouan, D., 1983, A&A, 117, 1, 164–169
Leung, C. M., Herbst, E. & Huebner, W. F., 1984, ApJS, 56, 231–256
Li, C., Qin, S.-L., Liu, T. et al., 2024, MNRAS, 533, 2, 1583–1617
Li, Z.-Y., Liu, X., Liu, T. et al., 2025, A&A, 697, A190
Ligterink, N. F. W., Tenenbaum, E. D. & van Dishoeck, E. F., 2015, A&A, 576, A35
Ligterink, N. F. W., El-Abd, S. J., Brogan, C. L. et al., 2020, ApJ, 901, 1, 37
Linnartz, H., Ioppolo, S. & Fedoseev, G., 2015, International Reviews in Physical Chemistry, 34, 2, 205–237
Liu, T., Evans, N. J., Kim, K.-T. et al., 2020, MNRAS, 496, 3, 2790–2820
Liu, X., Liu, T., Zhu, L. et al., 2024, Research in Astronomy and Astrophysics, 24, 2, 025009
Luna, R., Molpeceres, G., Ortigoso, J. et al., 2018, A&A, 617, A116
Lykke, J. M., Coutens, A., Jørgensen, J. K. et al., 2017, A&A, 597, A53
Léger, A., Klein, J., de Cheveigne, S. et al., 1979, A&A, 79, 1-2, 256–259
Maity, S., Kaiser, R. I. & Jones, B. M., 2015, Physical Chemistry Chemical Physics (Incorporating Faraday Transactions), 17, 5, 3081–3114

Manigand, S., Calcutt, H., Jørgensen, J. K. et al., 2019, A&A, 623, A69
Manigand, S., Jørgensen, J. K., Calcutt, H. et al., 2020, A&A, 635, A48
Mastrapa, R. M., Sandford, S. A., Roush, T. L. et al., 2009, ApJ, 701, 2, 1347–1356
Masunaga, H. & Inutsuka, S.-i., 2000, ApJ, 531, 1, 350–365
Masunaga, H., Miyama, S. M. & Inutsuka, S.-i., 1998, ApJ, 495, 1, 346–369
Maud, L. T., Moore, T. J. T., Lumsden, S. L. et al., 2015, MNRAS, 453, 1, 645–665
Maury, A. J., Belloche, A., André, P. et al., 2014, A&A, 563, L2
McClure, M. K., Rocha, W. R. M., Pontoppidan, K. M. et al., 2023, Nature Astronomy, 7, 431–443
McGuire, B. A., 2018, ApJS, 239, 2, 17
McGuire, B. A., 2022, ApJS, 259, 2, 30
McGuire, B. A., Burkhardt, A. M., Loomis, R. A. et al., 2020, ApJ, 900, 1, L10
McKee, C. F. & Tan, J. C., 2003, ApJ, 585, 2, 850–871
McLean, I. S., Becklin, E. E., Bendiksen, O. et al., 1998, A. M. Fowler, editor, Infrared Astronomical Instrumentation, volume 3354 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 566–578
McMullin, J. P., Waters, B., Schiebel, D. et al., 2007, R. A. Shaw, F. Hill & D. J. Bell, editors, Astronomical Data Analysis Software and Systems XVI, volume 376 of *Astronomical Society of the Pacific Conference Series*, 127
Merrill, K. M., Russell, R. W. & Soifer, B. T., 1976, ApJ, 207, 763–769
Milam, S. N., Savage, C., Brewster, M. A. et al., 2005, ApJ, 634, 2, 1126–1132
Mininni, C., Beltrán, M. T., Rivilla, V. M. et al., 2020, A&A, 644, A84
Mininni, C., Beltrán, M. T., Colzi, L. et al., 2023, A&A, 677, A15
Minissale, M., Congiu, E., Manicò, G. et al., 2013, A&A, 559, A49
Minissale, M., Aikawa, Y., Bergin, E. et al., 2022, ACS Earth and Space Chemistry, 6, 3, 597–630
Molinari, S., Schilke, P., Battersby, C. et al., 2025, A&A, 696, A149
Molpeceres, G., Enrique-Romero, J. & Aikawa, Y., 2023, A&A, 677, A39
Motte, F., Bontemps, S. & Louvet, F., 2018, ARA&A, 56, 41–82
Müller, H. S. P., Thorwirth, S., Roth, D. A. & Winnewisser, G., 2001, A&A, 370, L49–L52
Müller, H. S. P., Schlöder, F., Stutzki, J. & Winnewisser, G., 2005, Journal of Molecular Structure, 742, 1-3, 215–227
Murakami, H., Baba, H., Barthel, P. et al., 2007, PASJ, 59, S369–S376
Nazari, P., 2025, Life Sciences in Space Research
Nazari, P., van Gelder, M. L., van Dishoeck, E. F. et al., 2021, A&A, 650, A150
Nazari, P., Meijerhof, J. D., van Gelder, M. L. et al., 2022a, A&A, 668, A109
Nazari, P., Tabone, B., Rosotti, G. P. et al., 2022b, A&A, 663, A58
Nazari, P., Tabone, B. & Rosotti, G. P., 2023a, A&A, 671, A107
Nazari, P., Tabone, B., van't Hoff, M. L. R. et al., 2023b, ApJ, 951, 2, L38
Nazari, P., Rocha, W. R. M., Rubinstein, A. E. et al., 2024a, A&A, 686, A71
Nazari, P., Rocha, W. R. M., Rubinstein, A. E. et al., 2024b, A&A, 686, A71
Nazari, P., Tabone, B., Ahmadi, A. et al., 2024c, A&A, 686, A201
Neugebauer, G., Habing, H. J., van Duinen, R. et al., 1984, ApJ, 278, L1–L6

Newville, M., Stensitzki, T., Allen, D. B. & Ingargiola, A., 2014, LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python, doi: 10.5281/zenodo.11813

Noble, J. A., Fraser, H. J., Pontoppidan, K. M. & Craigon, A. M., 2017, MNRAS, 467, 4, 4753–4762

Noble, J. A., Fraser, H. J., Smith, Z. L. et al., 2024, Nature Astronomy, 8, 1169–1180

Öberg, K. I., Boogert, A. C. A., Pontoppidan, K. M. et al., 2008, ApJ, 678, 2, 1032–1041

Öberg, K. I., Fayolle, E. C., Cuppen, H. M. et al., 2009a, A&A, 505, 1, 183–194

Öberg, K. I., Garrod, R. T., van Dishoeck, E. F. & Linnartz, H., 2009b, A&A, 504, 3, 891–913

Öberg, K. I., Boogert, A. C. A., Pontoppidan, K. M. et al., 2011, ApJ, 740, 2, 109

Öberg, K. I., Guzmán, V. V., Furuya, K. et al., 2015, Nature, 520, 7546, 198–201

Öberg, K. I., Facchini, S. & Anderson, D. E., 2023, ARA&A, 61, 287–328

Oort, J. H. & van de Hulst, H. C., 1946, Bulletin of The Astronomical Institutes of the Netherlands, 10, 187

Paardekooper, D. M., Bossa, J. B. & Linnartz, H., 2016, A&A, 592, A67

Palumbo, M. E., Tielens, A. G. G. M. & Tokunaga, A. T., 1995, ApJ, 449, 674

Peebles, P. J. E., 1993, Principles of Physical Cosmology

Perotti, G., Rocha, W. R. M., Jørgensen, J. K. et al., 2020, A&A, 643, A48

Perotti, G., Jørgensen, J. K., Fraser, H. J. et al., 2021, A&A, 650, A168

Perotti, G., Jørgensen, J. K., Rocha, W. R. M. et al., 2023, A&A, 678, A78

Pickett, H. M., Poynter, R. L., Cohen, E. A. et al., 1998, J. Quant. Spectr. Rad. Transf., 60, 5, 883–890

Pilbratt, G. L., Riedinger, J. R., Passvogel, T. et al., 2010, A&A, 518, L1

Pontoppidan, K. M., 2006, A&A, 453, 3, L47–L50

Pontoppidan, K. M., Fraser, H. J., Dartois, E. et al., 2003, A&A, 408, 981–1007

Pontoppidan, K. M., Boogert, A. C. A., Fraser, H. J. et al., 2008, ApJ, 678, 2, 1005–1031

Potapov, A. & Garrod, R. T., 2024, A&A, 692, A252

Qasim, D., Fedoseev, G., Chuang, K. J. et al., 2019, A&A, 627, A1

Qasim, D., Fedoseev, G., Chuang, K. J. et al., 2020, Nature Astronomy, 4, 781–785

Qin, S.-L., Liu, T., Liu, X. et al., 2022, MNRAS, 511, 3, 3463–3476

Rachid, M. G., Terwisscha van Scheltinga, J., Koletzki, D. & Linnartz, H., 2020, A&A, 639, A4

Rachid, M. G., Brunken, N., de Boe, D. et al., 2021, A&A, 653, A116

Rachid, M. G., Rocha, W. R. M. & Linnartz, H., 2022, A&A, 665, A89

Raunier, S., Chiavassa, T., Marinelli, F. & Aycard, J.-P., 2004, Chemical Physics, 302, 1, 259–264

Rayalacheruvu, P., Majumdar, L., Rocha, W. R. M. et al., 2025, ApJS, 281, 2, 51

Regan, M., 2023, Detection and Flagging of Showers and Snowballs in JWST, Technical Report JWST-STScI-008545

Ribas, Á., Merín, B., Bouy, H. & Maud, L. T., 2014, A&A, 561, A54

Richard, C., Jørgensen, J. K., Margulès, L. et al., 2021, A&A, 651, A120

Rieke, G. H., Ressler, M. E., Morrison, J. E. et al., 2015, PASP, 127, 953, 665

Robitaille, T. P., Whitney, B. A., Indebetouw, R. et al., 2006, ApJS, 167, 2, 256–285
Rocha, W. R. M., Rachid, M. G., Olsthoorn, B. et al., 2022, A&A, 668, A63
Rocha, W. R. M., van Dishoeck, E. F., Ressler, M. E. et al., 2024, A&A, 683, A124
Rodgers, S. D. & Charnley, S. B., 2001, MNRAS, 320, 4, L61–L64
Rubin, M., Altwegg, K., Balsiger, H. et al., 2019a, MNRAS, 489, 1, 594–607
Rubin, M., Bekaert, D. V., Broadley, M. W. et al., 2019b, ACS Earth and Space Chemistry, 3, 9, 1792–1811
Sagan, C., 1972, Nature, 238, 5359, 77–80
Saha, A., Tej, A., Liu, H.-L. et al., 2022, MNRAS, 516, 2, 1983–2005
Sakai, N. & Yamamoto, S., 2013, Chemical Reviews, 113, 12, 8981–9015
Sakai, N., Sakai, T., Hirota, T. & Yamamoto, S., 2008, ApJ, 672, 1, 371–381
Sakai, N., Sakai, T., Hirota, T. et al., 2009, ApJ, 697, 1, 769–786
Sánchez-Monge, Á., Cesaroni, R., Beltrán, M. T. et al., 2013, A&A, 552, L10
Sandell, G., 2000, A&A, 358, 242–256
Santos, J. C., Chuang, K.-J., Lamberts, T. et al., 2022, ApJ, 931, 2, L33
Schilke, P., Groesbeck, T. D., Blake, G. A. et al., 1997, ApJS, 108, 1, 301–337
Schutte, W. A., Gerakines, P. A., Geballe, T. R. et al., 1996, A&A, 309, 633–647
Schutte, W. A., Boogert, A. C. A., Tielens, A. G. G. M. et al., 1999, A&A, 343, 966–976
Scibelli, S., Shirley, Y., Vasyunin, A. & Launhardt, R., 2021, MNRAS, 504, 4, 5754–5767
Segura-Cox, D. M., Looney, L. W., Tobin, J. J. et al., 2018, ApJ, 866, 2, 161
Shalabiea, O. M. & Greenberg, J. M., 1994, A&A, 290, 266–278
Shannon, R. J., Caravan, R. L., Blitz, M. A. & Heard, D. E., 2014, Phys. Chem. Chem. Phys., 16, 3466–3478
Shu, F. H., 1977, ApJ, 214, 488–497
Shu, F. H., Adams, F. C. & Lizano, S., 1987, ARA&A, 25, 23–81
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, 2, 135
Skouteris, D., Balucani, N., Ceccarelli, C. et al., 2019, MNRAS, 482, 3, 3567–3575
Slavicinska, K., Rachid, M. G., Rocha, W. R. M. et al., 2023, A&A, 677, A13
Slavicinska, K., van Dishoeck, E. F., Tychoniec, Ł. et al., 2024, A&A, 688, A29
Slavicinska, K., Boogert, A. C. A., Tychoniec, Ł. et al., 2025a, A&A, 693, A146
Slavicinska, K., Tychoniec, Ł., Navarro, M. G. et al., 2025b, ApJ, 986, 2, L19
Snow, T. P. & McCall, B. J., 2006, ARA&A, 44, 1, 367–414
Snyder, L. E., Buhl, D., Schwartz, P. R. et al., 1974, ApJ, 191, L79
Soifer, B. T., Puetter, R. C., Russell, R. W. et al., 1979, ApJ, 232, L53–L57
Soma, T., Sakai, N., Watanabe, Y. & Yamamoto, S., 2018, ApJ, 854, 2, 116
Spezzano, S., Caselli, P., Bizzocchi, L. et al., 2017, A&A, 606, A82
Sung, K. & Tidwell, T. T., 1998, The Journal of Organic Chemistry, 63, 26, 9690–9697
Sutton, E. C., Blake, G. A., Masson, C. R. & Phillips, T. G., 1985, ApJS, 58, 341–378
Suutarinen, A., 2015, omnifit: v0.2.1, doi: 10.5281/zenodo.35536
Tabone, B., Bettoni, G., van Dishoeck, E. F. et al., 2023, Nature Astronomy, 7, 805–814

Tan, J. C., Beltrán, M. T., Caselli, P. et al., 2014, H. Beuther, R. S. Klessen, C. P. Dullemond & T. Henning, editors, Protostars and Planets VI, 149–172
Tang, X. D., Henkel, C., Wyrowski, F. et al., 2018, A&A, 611, A6
Taquet, V., López-Sepulcre, A., Ceccarelli, C. et al., 2015, ApJ, 804, 2, 81
Taquet, V., Wirström, E. S. & Charnley, S. B., 2016, ApJ, 821, 1, 46
Tercero, F., López-Pérez, J. A., Gallego, J. D. et al., 2021, A&A, 645, A37
Terwisscha van Scheltinga, J., Ligterink, N. F. W., Boogert, A. C. A. et al., 2018, A&A, 611, A35
Terwisscha van Scheltinga, J., Marcandalli, G., McClure, M. K. et al., 2021, A&A, 651, A95
Tielens, A. G. G. M. & Hagen, W., 1982, A&A, 114, 2, 245–260
Towner, A. P. M., Brogan, C. L., Hunter, T. R. & Cyganowski, C. J., 2021, ApJ, 923, 2, 263
Tsang, W., 1988, Journal of Physical and Chemical Reference Data, 17, 2, 887–951
Tsang, W. & Hampson, R. F., 1986, Journal of Physical and Chemical Reference Data, 15, 3, 1087–1279
Tychoniec, Ł., van Dishoeck, E. F., van't Hoff, M. L. R. et al., 2021, A&A, 655, A65
van Broekhuizen, F. A., Keane, J. V. & Schutte, W. A., 2004, A&A, 415, 425–436
van Broekhuizen, F. A., Pontoppidan, K. M., Fraser, H. J. & van Dishoeck, E. F., 2005, A&A, 441, 1, 249–260
van de Hulst, H. C., 1946, Recherches Astronomiques de l'Observatoire d'Utrecht, 11, 2.i–2
van der Marel, N., Booth, A. S., Leemker, M. et al., 2021, A&A, 651, L5
van der Tak, F. F. S., van Dishoeck, E. F., Evans, I., Neal J. et al., 1999, ApJ, 522, 2, 991–1010
van der Tak, F. F. S., van Dishoeck, E. F., Evans, I., Neal J. & Blake, G. A., 2000, ApJ, 537, 1, 283–303
van der Tak, F. F. S., Chavarría, L., Herpin, F. et al., 2013, A&A, 554, A83
van Dishoeck, E. F. & Black, J. H., 1986, ApJS, 62, 109
van Dishoeck, E. F., Tychoniec, Ł., Rocha, W. R. M. et al., 2025, A&A, 699, A361
van Gelder, M. L., Tabone, B., Tychoniec, Ł. et al., 2020, A&A, 639, A87
van Gelder, M. L., Nazari, P., Tabone, B. et al., 2022, A&A, 662, A67
van Gelder, M. L., Francis, L., van Dishoeck, E. F. et al., 2024a, A&A, 692, A197
van Gelder, M. L., Ressler, M. E., van Dishoeck, E. F. et al., 2024b, A&A, 682, A78
van 't Hoff, M. L. R., 2019, Ph.D. thesis, University of Leiden, Netherlands
van't Hoff, M. L. R., Harsono, D., van Gelder, M. L. et al., 2022, ApJ, 924, 1, 5
Varricatt, W. P., Davis, C. J., Ramsay, S. & Todd, S. P., 2010, MNRAS, 404, 2, 661–720
Vastel, C., Ceccarelli, C., Lefloch, B. & Bachiller, R., 2014, ApJ, 795, 1, L2
Vastel, C., Bottinelli, S., Caux, E. et al., 2015, SF2A-2015: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, 313–316
Vazart, F., Ceccarelli, C., Balucani, N. et al., 2020, MNRAS, 499, 4, 5547–5561
Vázquez-Semadeni, E., González-Samaniego, A. & Colín, P., 2017, MNRAS, 467, 2, 1313–1328
Vázquez-Semadeni, E., Palau, A., Ballesteros-Paredes, J. et al., 2019, MNRAS, 490,

3, 3061–3097
Walsh, C., Loomis, R. A., Öberg, K. I. et al., 2016, ApJ, 823, 1, L10
Watanabe, N. & Kouchi, A., 2002, ApJ, 571, 2, L173–L176
Watson, W. D., 1974, ApJ, 188, 35–42
Wells, M., Pel, J. W., Glasse, A. et al., 2015, PASP, 127, 953, 646
Wenzel, G., Cooke, I. R., Changala, P. B. et al., 2024, Science, 386, 6723, 810–813
Wenzel, G., Speak, T. H., Changala, P. B. et al., 2025, Nature Astronomy, 9, 262–270
Whitney, B. A., Wood, K., Bjorkman, J. E. & Cohen, M., 2003a, ApJ, 598, 2, 1079–1099
Whitney, B. A., Wood, K., Bjorkman, J. E. & Wolff, M. J., 2003b, ApJ, 591, 2, 1049–1063
Whittet, D. C. B., Gerakines, P. A., Hough, J. H. & Shenoy, S. S., 2001, ApJ, 547, 2, 872–884
Widicus Weaver, S. L., Laas, J. C., Zou, L. et al., 2017, ApJS, 232, 1, 3
Williams, G. M., Cyganowski, C. J., Brogan, C. L. et al., 2022, MNRAS, 509, 1, 748–762
Williams, J. P. & Cieza, L. A., 2011, ARA&A, 49, 1, 67–117
Willner, S. P., Gillett, F. C., Herter, T. L. et al., 1982, ApJ, 253, 174–187
Wilson, R. W., Jefferts, K. B. & Penzias, A. A., 1970, ApJ, 161, L43
Wilson, T. L. & Rood, R., 1994, ARA&A, 32, 191–226
Woitke, P., Min, M., Pinte, C. et al., 2016, A&A, 586, A103
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K. et al., 2010, AJ, 140, 6, 1868–1881
Wright, G. S., Wright, D., Goodson, G. B. et al., 2015, PASP, 127, 953, 595
Wright, G. S., Rieke, G. H., Glasse, A. et al., 2023, PASP, 135, 1046, 048003
Wu, Y. W., Sato, M., Reid, M. J. et al., 2014, A&A, 566, A17
Xu, L.-H., Fisher, J., Lees, R. M. et al., 2008, Journal of Molecular Spectroscopy, 251, 1-2, 305–313
Xue, C., Byrne, A. N., Morgan, L. et al., 2025, ApJS, 281, 1, 9
Yamato, Y., Notsu, S., Aikawa, Y. et al., 2024, AJ, 167, 2, 66
Yang, Y.-L., Sakai, N., Zhang, Y. et al., 2021, ApJ, 910, 1, 20
Yang, Y.-L., Green, J. D., Pontoppidan, K. M. et al., 2022, ApJ, 941, 1, L13
Yarnall, Y. Y. & Hudson, R. L., 2022, ApJ, 931, 1, L4
Yocum, K. M., Milam, S. N., Gerakines, P. A. & Widicus Weaver, S. L., 2021, ApJ, 913, 1, 61
Young, C. H. & Evans, N. J., II, 2005, ApJ, 627, 1, 293–309
Zhang, B., Zheng, X. W., Reid, M. J. et al., 2009, ApJ, 693, 1, 419–423
Zingsheim, O., Müller, H. S. P., Lewen, F. et al., 2017, Journal of Molecular Spectroscopy, 342, 125–131
Zinnecker, H. & Yorke, H. W., 2007, ARA&A, 45, 1, 481–563
Zubko, V. G., Mennella, V., Colangeli, L. & Bussoletti, E., 1996, MNRAS, 282, 4, 1321–1329

Nederlandse Samenvatting

Sterren worden geboren samen met moleculen. Lang voordat een jonge ster zichtbaar wordt, is haar geboorte-omgeving al in stilte nieuwe moleculaire soorten aan het maken onder extreme omstandigheden. Dit proefschrift houdt zich bezig met die verborgen chemische geschiedenis: hoe eenvoudige atomen in het interstellaire medium worden omgevormd tot steeds complexere moleculen terwijl sterren ontstaan en evolueren.

Stervorming: een fysisch en chemisch proces

Sterren ontstaan binnen koude, dichte gebieden van de interstellaire ruimte die bekendstaan als **moleculaire wolken**. Deze moleculaire wolken behoren tot de koudste en rustigste plaatsen in het heelal, met temperaturen die zo laag zijn (< 10 K, of equivalent, $< -260°$C) dat beweging vrijwel bevriest. Het zijn echter precies deze kalme, ijzige omstandigheden die het mogelijk maken dat zowel sterren als moleculen zich vormen.

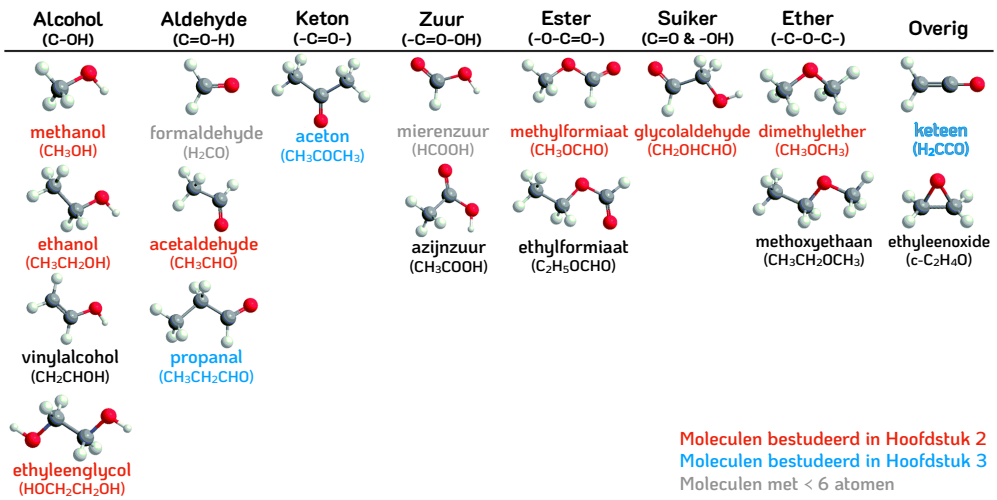
Terwijl de zwaartekracht langzaam materiaal binnen een wolk samenbrengt, wordt het gas dichter en kouder. In dit stadium beginnen atomen en eenvoudige moleculen (bijv. H en CO) zich vast te hechten aan kleine stofkorrels en deze geleidelijk te bedekken met dunne lagen ijs. Hoewel onzichtbaar voor het blote oog, zijn deze ijzige korrels cruciaal voor de voortgaande chemie: zij bieden oppervlakken waarop atomen elkaar kunnen ontmoeten, reageren en uitgroeien tot nieuwe moleculen.

Naarmate de instorting voortduurt, vormt zich een protoster in het centrum van de wolk. Terwijl het omringende materiaal geleidelijk wordt verwarmd door de groeiende ster, beginnen de ijsmantels op stofkorrels op te warmen, zich te herschikken en uiteindelijk te verdampen naar de gasfase. Gebieden dicht bij de protoster worden compacte reservoirs van warm, molecuul-rijk gas. Deze gebieden, bekend als **hete kernen** of **hete corino's**, zijn de plaatsen waar astronomen één van de rijkste chemie in de ruimte waarnemen.

Complexe organische moleculen: rijzende sterren in de astrochemie

Onder de vele moleculen die in stervormingsgebieden worden aangetroffen, zijn astronomen vooral geïnteresseerd in **complexe organische moleculen**, of **COM's**. Dit zijn koolstofhoudende moleculen die uit ten minste zes atomen bestaan—veel eenvoudiger dan DNA of eiwitten, maar volgens de maatstaven van het heelal al als "complex" beschouwd (zie Fig. 1 voor een voorbeeld van zuurstofhoudende COM's, of O-COM's).

Lange tijd gingen wetenschappers ervan uit dat dergelijke moleculen voornamelijk in de gasfase worden opgebouwd uit eenvoudigere soorten nadat zij van stofkorrels zijn verdampt. In de afgelopen twee decennia hebben laboratoriumexperimenten en astrochemische modellen echter aangetoond dat dit beeld onvolledig is. Veel gasfase-



Figuur 1: Overzicht van enkele veelvuldig gedetecteerde zuurstofhoudende COM's (O-COM's), gecategoriseerd op basis van hun functionele groepen. Overgenomen uit Fig. 12 van Maity et al. (2015).

reacties blijken te inefficiënt te zijn om de waargenomen abundanties van COM's te verklaren.

Waarnemingen hebben bovendien laten zien dat gasfase-COM's al aanwezig zijn in extreem koude gebieden voordat protosterren worden geboren, waar vrijwel niets zou moeten kunnen bewegen. Deze ontdekking dwong tot een heroverweging hoe de chemie werkt in de vroege stadia van stervorming. Als gevolg hiervan verschoof meer aandacht naar chemie in de vaste fase op de oppervlakken van stofkorrels, waar atomen en kleine moleculen zich stap voor stap kunnen combineren, zelfs bij zeer lage temperaturen. Latere theoretische studies zijn nog een stap verder gegaan en hebben reactiepaden voorgesteld die geen diffusie van moleculen vereisen, maar in plaats daarvan lokaal verlopen wanneer reactanten naast elkaar worden gevormd.

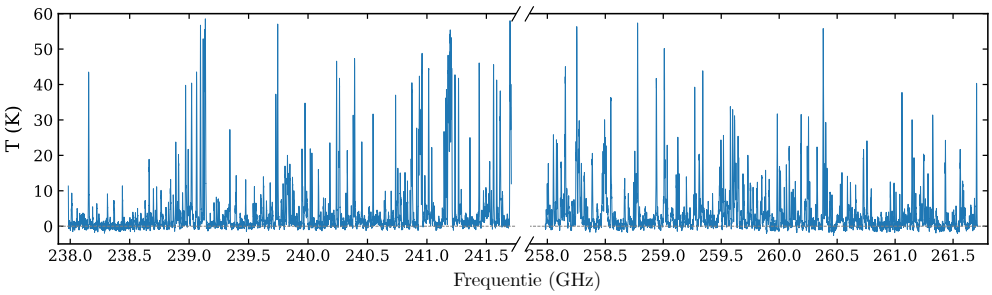
Tegenwoordig wordt de vorming van COM's begrepen als een combinatie van processen die plaatsvinden in koude ijzen, tijdens geleidelijke opwarming, en later in de gasfase. De vraag is nu niet langer of de oppervlaktechemie op stofkorrels van belang is, maar hoe verschillende chemische mechanismen in de loop van de tijd samenwerken.

Het waarnemen van COM's in het gas: de route tot nu toe

Moleculen kunnen zich op verschillende manieren bewegen. Naast de beweging als geheel kunnen zij ook roteren en trillen, waarbij straling met specifieke energieën wordt uitgezonden. Wanneer deze straling wordt gemeten over verschillende golflengten (die representatief zijn voor energieën), vormt zij wat astronomen een **spectrum** noemen. Door een waargenomen spectrum te analyseren, kunnen astronomen vaststellen welke moleculen aanwezig zijn en hun fysische eigenschappen meten, zoals abundantie en temperatuur.

Decennialang hebben astronomen interstellaire moleculen voornamelijk in de gasfase waargenomen, aangezien hun rotatiebewegingen signalen produceren op millimeter-

en submillimeter- (ook wel "radio"-) golflengten die relatief eenvoudig met radiotelescopen kunnen worden waargenomen. Gasfasewaarnemingen zijn vooral krachtig in hete-kerngebieden dichtbij protosterren, waar de ijsmantels van stofkorrels zijn verdampt. Inventarisaties van deze gebieden onthullen een buitengewone rijkdom aan moleculen, vaak met tientallen lijnen die samengepakt zijn binnen kleine frequentiebereiken (zoals te zien is in Fig. 2). In het bijzonder is de Atacama Large Millimeter/submillimeter Array (ALMA) revolutionair voor dit vakgebied, doordat deze zowel de gevoeligheid biedt die nodig is om zeldzame moleculen te detecteren als de ruimtelijke resolutie die nodig is om compacte chemische gebieden te isoleren.



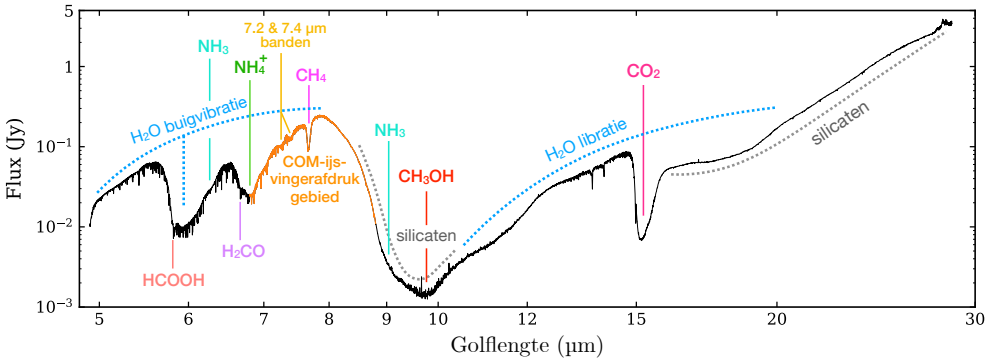
Figuur 2: ALMA-spectrum van G19.88-0.53, één van de massieve hete kernen die zijn bestudeerd in Hoofdstukken 2 & 3. De meeste weergegeven lijnen worden uitgezonden door gasfase-COM's.

Waarnemingen van gasfase-COM's tonen zowel overeenkomsten als verschillen in hun abundanties tussen protostellaire bronnen. De abundantie van sommige COM's (bijv. CH_3OCH_3 en CH_3OCHO) blijft consistent over bronnen met verschillende massa's, terwijl die van andere moleculen (bijv. CH_3CHO en C_2H_5OH) sterker varieert. Dergelijke patronen wijzen erop dat de moleculen onder vergelijkbare omstandigheden zijn gevormd, waarschijnlijk in het allervroegste begin van de stervorming, maar suggereren ook aanvullende reacties nadat moleculen in de gasfase zijn gesublimeerd. Gasfasewaarnemingen kunnen ons echter niet vertellen wat er vóór verdamping is gebeurd—een beperking die kritisch wordt wanneer men de volledige vormings- en evolutie geschiedenis van COM's probeert te begrijpen.

Het waarnemen van COM's in ijs: de ontbrekende helft van het beeld

Hoewel gasfasewaarnemingen zeer krachtig zijn, onthullen zij slechts een deel van de chemische inventaris. In de vroegste stadia van stervorming zijn de meeste moleculen opgesloten in ijsmantels op stofkorrels. Deze ijzen kunnen, in tegenstelling tot hun gasfase tegenhangers, niet roteren en radiogolven uitzenden. In plaats daarvan kunnen ze infrarood (IR) licht via verschillende trillingen, zoals buigen en strekken.

Helaas zijn infraroodwaarnemingen veel moeilijker en vanaf de grond beperkt door de absorptie van de aardatmosfeer. Eerdere infraroodtelescopen detecteerden voornamelijk eenvoudige ijzen (bijv. H_2O, CO en CO_2), waarbij methanol (CH_3OH) eruit sprong als het enige COM-ijs dat overtuigend in de ruimte werd gedetecteerd. Gelukkig heeft de lancering van de *James Webb* Space Telescope (JWST) in 2021 deze



Figuur 3: JWST-spectrum van B1-c, een bekende lage-massa protoster die is bestudeerd in Hoofdstuk 4. De absorptie banden van belangrijke interstellaire ijzen en het vingerafdrukgebied van COM-ijzen (6.8–8.8 μm) zijn in verschillende kleuren aangegeven.

situatie ingrijpend veranderd. Dankzij zijn uitzonderlijke gevoeligheid en spectrale resolutie kan JWST zwakke absorpties detecteren die worden geproduceerd door grotere moleculen (bijv. COM's groter dan methanol) in de vaste fase (zie Fig. 3 voor een JWST-spectrum van een protoster).

JWST-waarnemingen openen een nieuw venster op de chemie in de ruimte, maar brengen ook veel uitdagingen met zich mee. De vingerafdrukken van ijzen, die in infraroodspectra aanwezig zijn in de vorm van "absorptiebanden", hebben doorgaans geen geïsoleerde en voorspelbare vormen zoals de "emissielijnen" die door gasvormige moleculen worden geproduceerd. Aangezien de profielen van de banden van moleculair ijs onregelmatig variëren met temperatuur en mengverhoudingen (d.w.z. welke soorten gemengd zijn en in welke verhoudingen), is het meten van hun spectra in laboratoria essentieel voor het analyseren van de waarnemingsgegevens. Als een opkomend vakgebied binnen de astrochemie is de studie van COM-ijzen afhankelijk van een nauwe samenwerking tussen waarnemingen, laboratoriumexperimenten en de voortdurende verfijning van analysemethoden.

Dit proefschrift: het volgen van COM's van ijs naar gas

Het doel van dit proefschrift betreft observationele studies van COM's in zowel gas als ijs met behulp van twee van de krachtigste telescopen ter wereld—ALMA en JWST. Het eerste deel van dit proefschrift (Hoofdstukken 2–3) breidt traditionele studies van gasfase-COM's in individuele bronnen uit naar systematische analyses van grote steekproeven. In het tweede deel (Hoofdstukken 4–5) worden de gasfasewaarnemingen vervolgens verbonden met de nieuw beschikbare waarnemingen van COM-ijzen door JWST. De combinatie van ALMA- en JWST-waarnemingen stelt ons in staat een meer complete ijs-naar-gas-evolutie van COM's te volgen tijdens de vroege stadia van stervorming.

Het eerste deel van dit proefschrift richt zich op ALMA-waarnemingen van gasfase-COM's in de hete kern van een dozijn massieve protosterren. Specifiek bestudeert **Hoofdstuk 2** zes zuurstofhoudende COM's (d.w.z. de in rood gemarkeerde moleculen, met uitzondering van methanol, in Fig. 1) en vergelijkt hun abundantie tussen 14

massieve bronnen en vijf lage massa bronnen. Door deze zes moleculen te normaliseren ten opzichte van methanol (het eenvoudigste en meest abundante COM), toont dit hoofdstuk een contra-intuïtief resultaat: de abundantieverhoudingen tussen deze O-COM's en methanol lijken niet te variëren met de stellaire lichtkracht of massa. Hoewel de abundantieverhoudingen voor verschillende moleculen een uiteenlopende mate van spreiding vertonen, blijven de algemene trends constant over vele ordes van grootte in stellaire lichtkracht.

Hoofdstuk 3 breidt de systematische studie van gasfase-COM's uit Hoofdstuk 2 uit naar aceton (CH_3COCH_3) en propanal (C_2H_5CHO), twee grotere O-COM's met drie koolstofatomen die tot nu toe alleen in individuele bronnen zijn bestudeerd. Naast het meten van hun fysische eigenschappen onderzoekt dit hoofdstuk verschillende mogelijke vormingspaden door hun gasfase-abundanties te vergelijken met die van andere mogelijk relevante moleculen, zoals acetaldehyde (CH_3CHO) en keteen (H_2CCO). Dit hoofdstuk vat ook de gasfase-abundantieverhoudingen samen tussen methanol en alle acht O-COM's die zijn bestudeerd in Hoofdstukken 2 en 3, en onthult interessante resultaten waaruit blijkt dat aldehyden (moleculen met een –CHO-groep) in het algemeen minder abundant zijn in de gasfase, terwijl moleculen met CH_3O-radicalen tot de meest abundante O-COM's in waarnemingen behoren. Deze resultaten tonen het belang aan van een systematische analyse en vragen bovendien om meer aandacht voor aceton, een abundant O-COM dat door de astrochemische gemeenschap vaak over het hoofd is gezien.

Het tweede deel van dit proefschrift verlegt de focus van gas naar ijs. **Hoofdstuk 4** laat de mogelijkheid zien om vaste-fase-COM's die groter zijn dan methanol met JWST te detecteren en demonstreert de volledige procedure om de abundantie van COM-ijs te kwantificeren vanuit het zogenoemde "COM-ijs-vingerafdrukgebied" (d.w.z. 6.8–8.8 µm) in JWST-spectra. Dit hoofdstuk presenteert tevens de eerste directe gas-tot-ijs-vergelijking van COM's in dezelfde protostellaire systemen. Door de abundantie van vier veelvuldig gedetecteerde O-COM's in beide fasen te meten, laat dit hoofdstuk zien dat sommige moleculen intact lijken te blijven tijdens de overgang van ijs naar gas, terwijl andere dat niet doen. De resultaten geven aan dat zowel over-erving vanuit ijzen als daaropvolgende reacties in de gasfase een belangrijke rol kunnen spelen bij het vormgeven van de COM-chemie.

Hoofdstuk 5 breidt de studie van het COM-ijs-vingerafdrukgebied uit Hoofdstuk 4 uit naar een grotere steekproef van protostellaire bronnen en ontwikkelt een systematische methode om zwakke ijskenmerken uit infraroodspectra te isoleren. Door de eigenschappen van absorptiebanden (bijv. hun piekposities en breedtes) te meten en deze te vergelijken met laboratoriumgegevens, kunnen de meest plausibele ijssoorten of -mengsels ge-identificeerd worden die verantwoordelijk zijn voor deze banden. Dit hoofdstuk biedt de eerste systematische analyse van het COM-ijs-vingerafdrukgebied en versterkt het bewijs dat chemie in de vaste fase wijdverbreid is tijdens de vroege stervorming.

Toekomstperspectief: een nieuw tijdperk van de astrochemie

We zijn nu een nieuw tijdperk van de astrochemie binnengetreden. Terwijl ALMA blijft zorgen voor spectra en kaarten van moleculair gas van hoge kwaliteit, heeft JWST de deur geopend naar systematische studies van grotere moleculen in ijs. In plaats van ijschemie indirect af te leiden uit gasfasewaarnemingen, kunnen we nu beide reservoirs

rechtstreeks waarnemen en ze kwantitatief met elkaar vergelijken. In de komende jaren zullen grootschalige waarnemingsprogramma's deze studies uitbreiden naar veel meer bronnen met verschillende massa's. Tegelijkertijd zullen laboratoriumexperimenten en chemische modellen worden ingezet om diepgaandere analyses van waarnemingsgegevens mogelijk te maken. Samen brengen deze ontwikkelingen ons dichter bij het beantwoorden van bredere vragen over de chemische oorsprong van planetaire systemen, en uiteindelijk de processen die tot leven leiden.

English summary

Stars are born together with molecules. Long before a young star becomes visible, its natal environment is already quietly assembling new molecular species under extreme conditions. This thesis is concerned with that hidden chemical history: how simple atoms in the interstellar medium are transformed into increasingly complex molecules as stars form and evolve.

Star formation: a physical and chemical process

Stars form within cold, dense regions of interstellar space known as **molecular clouds**. These molecular clouds are some of the quietest places in the Universe, with temperatures so low ($\sim$ 10 K, or equivalently, $< -260°$C) that motion nearly freezes. However, it is exactly these calm, frigid conditions that allow both stars and molecules to form.

As gravity slowly pulls material together inside a cloud, gas becomes denser and colder. At this stage, atoms and simple molecules (e.g., H and CO) begin to stick to tiny dust grains and gradually coat them with thin layers of ice. Although invisible to the naked eye, these icy grains are crucial to the ongoing chemistry: they provide surfaces where atoms can meet, react, and grow into new molecules.

As the collapse continues, a protostar forms at the center of the cloud. While the surrounding material is gradually heated by the growing star, icy mantles on dust grains begin to warm, rearrange, and eventually evaporate into the gas phase. Regions close to the protostar become compact reservoirs of warm, molecule-rich gas. These regions, known as **hot cores** or **hot corinos**, are where astronomers observe some of the richest chemistry in space.

Complex organic molecules: a rising star in astrochemistry

Among the many molecules found in star-forming regions, astronomers are especially interested in **complex organic molecules**, or **COMs**. These are carbon-containing molecules made of at least six atoms—far simpler than DNA or proteins, but already considered "complex" by the standards of the Universe (see Fig. 1 for a sample of oxygen-bearing COMs, or O-COMs).

For a long time, scientists assumed that such molecules are mainly built upon simpler species in the gas phase after evaporated from dust grains. Over the past two decades, however, laboratory experiments and astrochemical models have revealed that this picture is incomplete. Many gas-phase reactions turn out to be too inefficient to explain the observed abundances of COMs. Observations have also shown that gas-phase COMs already exist in extremely cold regions before protostars are born, where almost nothing should be able to move. This discovery forced a rethink of how chemistry works in the early stages of star formation. As a result, more attention turned to solid-phase chemistry on grain surfaces, where atoms and small molecules can combine step by step, even at very low temperatures. Subsequent theoretical

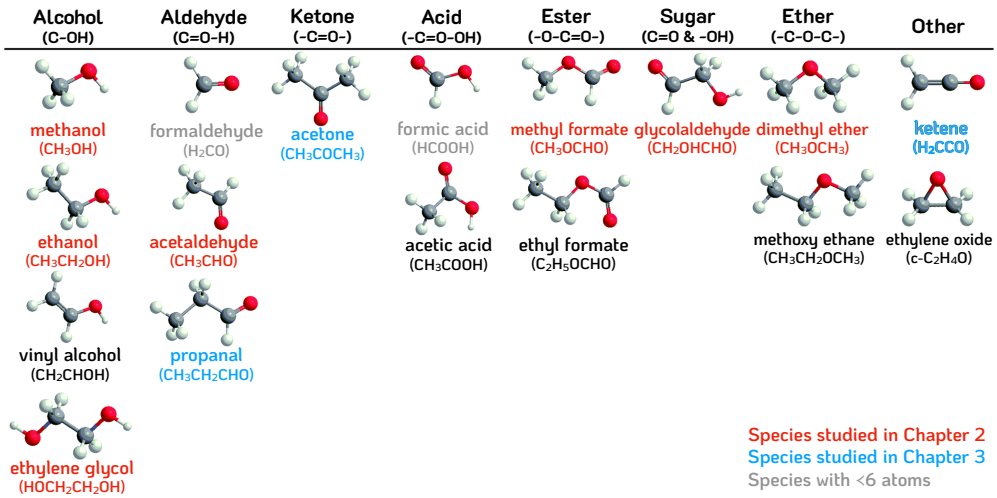


Figure 1: Summary of some commonly detected oxygen-bearing COMs (O-COMs), categorized by their functional groups. Adopted from Fig. 12 of Maity et al. (2015).

studies have gone a step further, proposing reaction pathways that do not require molecules to diffuse but instead proceed locally when reactants are created next to one another.

Today, the formation of COMs is understood as a combination of processes taking place in cold ices, during gradual warm-up, and later in the gas. The question now is no longer whether grain-surface chemistry matters, but how different chemical mechanisms work together over time.

Observing COMs in the gas: the path so far

Molecules can move in different ways. Besides moving as a whole, they can also rotate and vibrate, producing radiation of specific energies. When this radiation is measured across different wavelengths (which represent energies), it forms what astronomers call a **spectrum**. By analyzing an observed spectrum, astronomers can identify which molecules are present and measure their physical properties such as abundance and temperature.

For decades, astronomers have primarily observed interstellar molecules in the gas phase, since their rotational motions produce signals at millimeter and submillimeter (also known as "radio") wavelengths that are relatively easy to observe with radio telescopes. Gas-phase observations are especially powerful in hot core regions close to protostars, where icy mantles have evaporated from dust grains. Surveys of these regions reveal an extraordinary richness of molecules, often with dozens of lines packed into small frequency ranges (as shown in Fig. 2). In particular, the Atacama Large Millimeter/submillimeter Array (ALMA) has been revolutionary in the field, offering both the sensitivity needed to detect rare molecules and the spatial resolution needed to isolate compact chemical regions.

Observations of gas-phase COMs show both similarities and differences in their abundances among protostellar sources. The abundances of some COMs (e.g., CH_3OCH_3

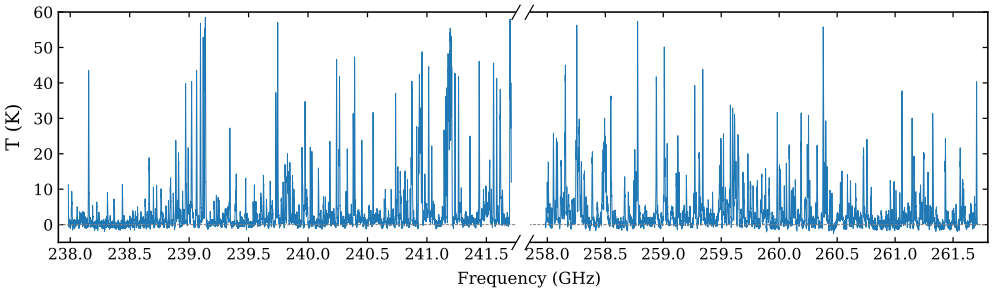


Figure 2: ALMA spectrum of G19.88-0.53, one of the high-mass hot cores studied in Chapters 2 & 3. Most of the displayed lines are emitted by gas-phase COMs.

and CH_3OCHO) remain consistent across sources with different masses, while those of others (e.g., CH_3CHO and C_2H_5OH) vary more significantly. Such patterns hint at similar formation environments that are more likely realized in the very beginning of star formation, but also suggest additional processing after molecules sublimate into the gas phase. However, gas-phase observations alone cannot tell us what happened before evaporation—a limitation that becomes critical when trying to understand the full formation and evolutionary history of COMs.

Observing COMs in ices: the missing half of the picture

While gas-phase observations are powerful, they reveal only part of the chemical inventory. In the earliest stages of star formation, most molecules are locked up in icy mantles on dust grains. These ices are not able to rotate and emit radio waves like their gaseous counterparts, but instead they can absorb infrared (IR) light through various vibrations such as bending and stretching.

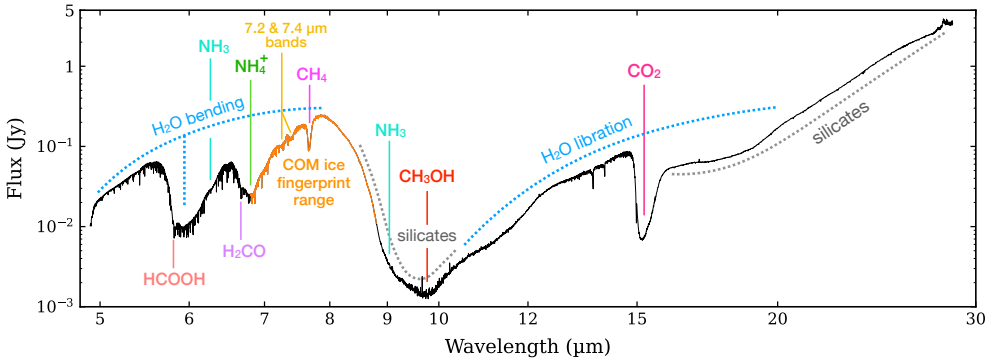


Figure 3: JWST spectrum of B1-c, a famous low-mass protostar studied in Chapter 4. The absorption features of major interstellar ices and the fingerprint range of COM ices (6.8–8.8 μm) are labeled in different colors.

Unfortunately, infrared observations are challenging and limited from the ground due to the absorption of Earth's atmosphere. Earlier infrared telescopes mainly de-

tected simple ices (e.g., H_2O, CO, and CO_2), with methanol (CH_3OH) stood out as the only COM ice that was firmly detected in space. Fortunately, the launch of the *James Webb* Space Telescope (JWST) in 2021 changed this situation dramatically. With its exceptional sensitivity and spectral resolution, JWST can detect weak absorption features produced by larger molecules (e.g., COMs larger than methanol) in the solid phase (see Fig. 3 for a JWST spectrum of a protostar).

JWST observations open a new window on chemistry in space, but also bring a lot of challenges. The fingerprints of ices, which are present in the form of "absorption bands" in infrared spectra, usually do not have isolated and predictable shapes as the "emission lines" produced by gaseous molecules. Since the band shapes of molecular ices vary irregularly with temperature and mixing conditions (i.e., which species are mixed and in what proportions), measuring their spectra in laboratories is essential for analyzing the observational data. As an emerging field in astrochemistry, the study of COM ices relies on close interplay between observations, laboratory experiments, and continued refinement of analysis strategies.

This thesis: tracing COMs from ice to gas

This thesis aims at advancing the observational studies of COMs in both gas and ice using two of the most powerful telescopes in the world—ALMA and JWST. The first part of this thesis (Chapters 2–3) extends traditional case studies of gas-phase COMs in individual sources to systematic analyses of large samples. The gas-phase observations are then connected to the newly available observations of COM ices by JWST in the second part (Chapters 4–5). The combination of ALMA and JWST observations enables us to trace a more complete ice-to-gas evolution of COMs during early stages of star formation.

The first part of this thesis focuses on ALMA observations of gas-phase COMs in the hot core of a dozen of high-mass protostars. Specifically, **Chapter 2** studies six oxygen-bearing COMs (i.e., those highlighted in red other than methanol in Fig. 1) and compares their abundances across 14 high-mass sources and five low-mass sources. By normalizing these six molecules to methanol (the simplest and most abundant COM), this chapter shows a counterintuitive result: the abundance ratios between these O-COMs and methanol seem not vary with stellar luminosity or mass. Although the abundance ratios show varying degrees of scatter for different molecules, the overall trends remain constant across stellar luminosity.

Chapter 3 extends the systematic study of gas-phase COMs in Chapter 2 to acetone (CH_3COCH_3) and propanal (C_2H_5CHO), two larger O-COMs with three carbon atoms that have only been studied in individual sources. In addition to measuring their physical properties, this chapter examines several possible formation pathways by checking their gas-phase abundances with respect to those of other potentially relevant molecules such as acetaldehyde (CH_3CHO) and ketene (H_2CCO). This chapter also summarizes the gas-phase abundance ratios between methanol and all the eight O-COMs studied in Chapters 2 and 3, revealing interesting results that aldehydes (those with –CHO group) are overall depleted in the gas phase, while molecules with CH_3O radicals are among the most abundant O-COMs in observations. These results show the importance of systematic analysis, and also call for more attention to acetone, which is an abundant O-COM but has often been overlooked by the community.

The second part of this thesis shifts the focus from gas to ice. **Chapter 4** shows the possibility of detecting solid-phase COM larger than methanol with JWST and demonstrates the full procedure of how to quantify the COM ice abundances from the "COM ice fingerprint range" (i.e., 6.8–8.8 µm) in JWST spectra. This chapter also presents the first direct gas-to-ice comparison of COMs in the same protostellar systems. By measuring the abundances of four commonly detected O-COMs in both phases, this chapter shows that some molecules appear to keep intact during the transition from ice to gas, while others do not. The results indicate that both inheritance from ices and subsequent gas-phase reprocessing can play significant roles in shaping the COM chemistry.

Chapter 5 expands the study of the COM ice fingerprint range in Chapter 4 to a larger sample of protostellar sources, and comes up with a systematic method for isolating weak ice features from infrared spectra. By measuring the properties of absorption bands (e.g., their peak positions and widths) and comparing them with laboratory data, we can identify the most plausible ice species or mixtures responsible for these bands. This chapter provides the first systematic analysis of the COM ice fingerprint range, strengthening the case that solid-phase chemistry is widespread during early star formation.

Looking ahead: a new era of astrochemistry

We have now entered a new era of astrochemistry. While ALMA continues to provide high-quality spectra and maps of molecular gas, JWST has opened the door to systematic studies of bigger molecules in ice. Instead of inferring ice chemistry indirectly from gas-phase observations, we can now observe both reservoirs directly and compare them quantitatively. In the coming years, large observational programs will expand these studies to many more sources with different masses. At the same time, laboratory experiments and chemical models will be leveraged to enable deeper analyses of observational data. Together, these advances bring us closer to addressing broader questions about the chemical origins of planetary systems, and ultimately, the pathways that lead to life.

Publications

Refereed publications as first author

1. *CoCCoA: Complex Chemistry in hot Cores with ALMA. The chemical evolution of acetone from ice to gas*
   Chen, Y., Garrod, R. T., Rachid, M., van Dishoeck, E. F., Brogan, C., Loomis, R., Lipnicky, A., McGuire, B. A., 2025, Astronomy & Astrophysics, 696, A198. (Chapter 4)

2. *JOYS+: The link between the ice and gas of complex organic molecules. Comparing JWST and ALMA data of two low-mass protostars*
   Chen, Y., Rocha, W. R. M., van Dishoeck, E. F., van Gelder, M. L., Nazari, P., Slavicinska, K., Francis, L., Tabone, B., Ressler, M. E., Klaassen, P. D., Beuther, H., Boogert, A. C. A., Gieser, C., Kavanagh, P. J., Perotti, G., Le Gouellec, V. J. M., Majumdar, L., Güdel, M., and Henning, T., 2024, Astronomy & Astrophysics, 690, A205. (Chapter 3)

3. *CoCCoA: Complex Chemistry in hot Cores with ALMA. Selected oxygen-bearing species*
   Chen, Y., van Gelder, M. L., Nazari, P., Brogan, C. L., van Dishoeck, E. F., Linnartz, H., Jørgensen, J. K., Hunter, T. R., Wilkins, O. H., Blake, G. A., Caselli, P., Chuang, K.-J., Codella, C., Cooke, I., Drozdovskaya, M. N., Garrod, R. T., Ioppolo, S., Jin, M., Kulterer, B. M., Ligterink, N. F. W., Lipnicky, A., Loomis, R., Rachid, M. G., Spezzano, S., and McGuire, B. A., 2023, Astronomy & Astrophysics, 678, A137. (Chapter 2)

Submitted publication as first author

1. *JOYS+: An overview of the ice features in the COM fingerprint range (6.8–8.8 μm)*
   Chen, Y., van Dishoeck, E. F., Francis, L., Slavicinska, K., Tychoniec, Ł., van Gelder, M. L., Boogert, A. C. A., Schravendijk, C. A. F., Beuther, H., Caratti o Garatti, A., Eenhorst, E. K. H., Gieser, C., Henning, Th., Güdel, M., Le Gouellec, V. J. M., McClure, M. K., Narang, M., Nazari, P., Reyes, S. D., Tolman, M., and Yang, Y.-L., submitted to Astronomy & Astrophysics. (Chapter 5)

Refereed publications as co-author

1. *Expanding the ice inventory of NGC1333 IRAS 2A with INDRA using JWST observations: Tracing organic refractories and beyond*

Rayalacheruvu, P., Majumdar, L., Rocha, W. R. M., Ressler, M. E., Giri, P. R., Maitrey, S., Willacy, K., Lis, D. C., **Chen, Y.**, Klaassen, P. D., 2025, The Astrophysical Journal Supplement Series, 281, 51.

2. *Sulfur oxides tracing streamers and shocks at low-mass protostellar disk–envelope interfaces*
Liu, X.-C., van Dishoeck, E. F., Hogerheijde, M. R., van Gelder, M. L., **Chen, Y.**, Liu, T., van't Hoff, M., Drozdovskaya, M. N., Artur de la Villarmois, E., Mai, X.-F., Tychoniec, Ł., 2025, Astronomy & Astrophysics, 701, A141

3. *HDO ice detected toward an isolated low-mass protostar with JWST*
Slavicinska, K., Tychoniec, Ł., Navarro, M. G., van Dishoeck, E. F., Tobin, J. J., van Gelder, M. L., **Chen, Y.**, Boogert, A. C. A., Drechsler, W. B., Beuther, H., Caratti o Garatti, A., Megeath, S. T., Klaassen, P. D., Looney, L. W., Kavanagh, P. J., Brunken, N. G. C., Sheehan, P., Fischer, W. J., 2025, The Astrophysical Journal Letters, 986, L19.

4. *JWST Observations of Young protoStars (JOYS): Overview of program and early results*
van Dishoeck, E. F., Tychoniec, Ł., Rocha, W. R. M., Slavicinska, K., Francis, L., van Gelder, M. L., Ray, T. P., Beuther, H., Caratti o Garatti, A., Brunken, N. G. C., **Chen, Y.**, Devaraj, R., Geers, V. C., Gieser, C., Greene, T. P., Justtanont, K., Le Gouellec, V. J. M., Kavanagh, P. J., Klaassen, P. D., Janssen, A. G. M., Navarro, M. G., Nazari, P., Notsu, S., Perotti, G., Ressler, M. E., Reyes, S. D., Sellek, A. D., Tabone, B., Tap, C., Theijssen, N. C. M. A., Colina, L., Güdel, M., Henning, Th., Lagage, P.-O., Östlin, G., Vandenbussche, B., Wright, G. S., 2025, Astronomy & Astrophysics, 699, A361.

5. *JWST Observations of Young protoStars (JOYS): Overview of gaseous molecular emission and absorption in low-mass protostars*
van Gelder, M. L., Francis, L., van Dishoeck, E. F., Tychoniec, Ł., Ray, T. P., Beuther, H., Caratti o Garatti, A., **Chen, Y.**, Devaraj, R., Gieser, C., Justtanont, K., Kavanagh, P. J., Nazari, P., Reyes, S. D., Rocha, W. R. M., Slavicinska, K., Güdel, M., Henning, T., Lagage, P.-O., and Wright, G., 2024, Astronomy & Astrophysics, 692, A197.

6. *JWST detections of amorphous and crystalline HDO ice toward massive protostars*
Slavicinska, K., van Dishoeck, E. F., Tychoniec, Ł., Nazari, P., Rubinstein, A. E., Gutermuth, R., Tyagi, H., **Chen, Y.**, Brunken, N. G. C., Rocha, W. R. M., Manoj, P., Narang, M., Thomas Megeath, S., Yang, Y.-L., Looney, L. W., Tobin, J. J., Beuther, H., Bourke, T. L., Linnartz, H., Federman, S., Watson, D. M., and Linz, H., 2024, Astronomy & Astrophysics, 688, A29.

7. *JWST Observations of Young protoStars (JOYS): Linked accretion and ejection in a Class I protobinary system.*
Tychoniec, Ł., van Gelder, M. L., van Dishoeck, E. F., Francis, L., Rocha, W. R. M., Caratti o Garatti, A., Beuther, H., Gieser, C., Justtanont, K., Linnartz, H., Le Gouellec, V. J. M., Perotti, G., Devaraj, R., Tabone, B., Ray, T. P., Brunken,

N. G. C., **Chen, Y.**, Kavanagh, P. J., Klaassen, P. D., Slavicinska, K., Güdel, M., and Östlin, G., 2024, Astronomy & Astrophysics, 687, A36.

8. *JWST Observations of Young protoStars (JOYS+): Detecting icy complex organic molecules and ions. I. CH_4, SO_2, $HCOO^-$, OCN^-, H_2CO, HCOOH, CH_3CH_2OH, CH_3CHO, CH_3OCHO, and CH_3COOH.*
Rocha, W. R. M., van Dishoeck, E. F., Ressler, M. E., van Gelder, M. L., Slavicinska, K., Brunken, N. G. C., Linnartz, H., Ray, T. P., Beuther, H., Caratti o Garatti, A., Geers, V., Kavanagh, P. J., Klaassen, P. D., Justtanont, K., **Chen, Y.**, Francis, L., Gieser, C., Perotti, G., Tychoniec, Ł., Barsony, M., Majumdar, L., Le Gouellec, V. J. M., Chu, L. E. U., Lew, B. W. P., Henning, T., and Wright, G., 2024, Astronomy & Astrophysics, 683, A124.

9. *JOYS+: Mid-infrared detection of gas-phase SO_2 emission in a low-mass protostar. The case of NGC 1333 IRAS 2A: Hot core or accretion shock?*
van Gelder, M. L., Ressler, M. E., van Dishoeck, E. F., Nazari, P., Tabone, B., Black, J. H., Tychoniec, Ł., Francis, L., Barsony, M., Beuther, H., Caratti o Garatti, A., **Chen, Y.**, Gieser, C., Le Gouellec, V. J. M., Kavanagh, P. J., Klaassen, P. D., Lew, B. W. P., Linnartz, H., Majumdar, L., Perotti, G., and Rocha, W. R. M., 2024, Astronomy & Astrophysics, 682, A78.

10. *Which molecule traces what: Chemical diagnostics of protostellar sources.*
Tychoniec, Ł., van Dishoeck, E. F., van't Hoff, M. L. R., van Gelder, M. L., Tabone, B., **Chen, Y.**, Harsono, D., Hull, C. L. H., Hogerheijde, M. R., Murillo, N. M., and Tobin, J. J., 2021, Astronomy & Astrophysics, 655, A65.

Submitted publications as co-author

1. *JOYS: Launching and destruction of dust in protostellar jets. The case of BHR71-IRS1 with JWST/MIRI*
Tychoniec, Ł., Francis, L., Navarro, M. G., Vorster, J. M., van Dishoeck, E. F., Caratti o Garatti, A., Assani, K., Le Gouellec, V. J. M., Tabone, B., Klaassen, P. D., Janssen, A. G. M., Justtanont, K., Harsono, D., Nazari, P., Reyes, S. D., Slavicinska, K., Gieser, C., Bourke, T., Yang, Y.-L., Nisini, B., Giannini, T., Beuther, H., Devaraj, R., Ray, T. P., Brunken, N. G. C., **Chen, Y.**, van Gelder, M. L., submitted to Astronomy & Astrophysics.

2. *JOYS: Linking the molecular ice and gas-phase composition towards the high-mass hot core IRAS 18089-1732*
Gieser, C., Rocha, W. R. M., **Chen, Y.**, Slavicinska, K., van Dishoeck, E. F., Nazari, P., Francis, L., Beuther, H., Reyes, S. D., Caratti o Garatti, A., Klaassen, P. D., Vorster, J. M., Navarro, M. G., submitted to Astronomy & Astrophysics.

3. *Measuring accretion rates onto high-mass protostars with JWST*
Reyes, S. D., Beuther, H., van Dishoeck, E. F., Gieser, C., Caratti o Garatti, A., Tychoniec, Ł., Kavanagh, P. J., Klaassen, P. D., Justtanont, K., Francis, L., Le Gouellec, V. J. M., Devaraj, R., Ray, T. P., **Chen, Y.**, Rocha, W. R. M., van Gelder, M. L., submitted to Astronomy & Astrophysics.

4. *JOYS+: Analyses of OCN^-, N_2O, NO, and complex cyanides in ices. Thermal processing results in modest enhancement of OCN^- ice*
Nazari, P., Brunken, N. G. C., **Chen, Y.**, Slavicinska, K., van Dishoeck, E. F., Rocha, W. R. M., Boogert, A. C. A., Navarro, M. G., Le Gouellec, V. J. M., Francis, L., Tychoniec, Ł., Caratti o Garatti, A., Gieser, C., Greene, T. P., Kavanagh, P. J., submitted to Astronomy & Astrophysics.

Curriculum Vitae

Twenty-eight years ago, I was born on the 4th of May, the Youth Day in China, which was established to commemorate the patriotic May Fourth Movement in 1919. Although both of my parents were born and bred in the north, they migrated to the tropical city of Guangzhou in the early 1990s, which is where I grew up.

The first fifteen years of my life had nothing to do with astronomy, and the reason was simple: you have very little chance to pay attention to the night sky in a metropolis* with a rainy climate, towering buildings, and severe light pollution. After entering secondary school at 12, I became a boarding student and couldn't hang out with my playmates after school anymore. I gradually lost contact with them and focused more on my studies instead. Luckily, I inherited my parents' talent for taking exams and was admitted to one of the best high schools in the province—Guangdong Experimental High School, the place where all the stories began.

I still clearly remember the moment when my attention was captured by a photo in the admission brochure. It was a photo of the award ceremony of Chinese National Astronomy Olympiad (CNAO) in 2012. Everyone in the photo was smiling so brightly that I couldn't help imagining myself as one of them. As a result, I joined the astronomy club (SSAA†) in the first month of school. It turned out that SSAA is such an amazing student club that it became the best part of my school life. There were two academic lessons per week taught by second-year students, where I was able to learn basic astronomy from scratch. When the lessons switched from introducing the seasonal night sky to learning theories and equations, the class began to thin out. However, I was one of the few who stayed on, not only because I found astronomy interesting, but also because I was deeply touched by the enthusiasm and devotion of those senior students. I had been studying hard to meet the expectations of people around, but never thought about chasing a dream of my own. For a 15-year-old teen, nothing could be more exciting than exploring the scenery off the beaten track.

Driven by strong curiosity, I devoted most of my free time to learning astronomy. After only three months of study, I placed among the top 20 in the high school division of a provincial astronomy competition. Several months later, I became a finalist of CNAO and won a silver medal right before my 16th birthday—it took me only eight months to progress from a beginner to a medalist. Even now, I can hardly imagine how enthusiastic I was at that time. At the end of my first year in high school, I took over the astronomy club as the vice president, and attended a nationwide astronomy summer school in Beijing. From then on, I set my sights on majoring in astronomy. However, only a handful of universities in China offered astronomy programs at that time, and admission to most of them was very competitive. In the final semester, I gradually began to struggle under the prolonged high intensity of studies, and finally fell ill on the eve of the college entrance examination (also known as *Gaokao*). However,

*Guangzhou is the third largest city in China with a population of over 20 million.

†"SS" is the abbreviation for my high school's name (in Chinese), and "AA" is the abbreviation for "astronomy association".

I never thought about giving up, and somehow my prayers were answered. I delivered my best performance in Gaokao and gained admission to my dream school—Peking University, of which the anniversary happens to be on the same day as my birthday.

Heavy is the head that wears the crown. Although it was a great honor to be a student of PKU, it was not easy to survive there. Since the Department of Astronomy is part of the School of Physics, astronomy students usually take classes together with physics students, many of whom are medalists of National and even International Physics Olympiads. For someone like me with neither a strong aptitude for physics[‡] nor any experience in physics competitions, it was a nightmare to keep up with fellow students. Fortunately, as soon as I started my campus life, I made up my mind to pursue graduate study abroad to engage with the forefront of astronomical research, and my determination to achieve this goal kept me from falling into despair. In my second year, I joined Prof. Gregory Herczeg's group and received basic research training. To prepare for my PhD applications, I also pursued a summer internship with Dr. Lumila Carone at the Max Planck Institute for Astronomy (MPIA) in 2018. In the subsequent fall, I traveled to the University of Michigan, Ann Arbor to conduct my bachelor thesis project on chemical modeling of protoplanetary disks with Prof. Edwin Bergin. These experiences not only equipped me with essential research skills, but also taught me how to navigate life abroad.

The four years I spent as an undergraduate were demanding yet fulfilling, but the ending was not as favorable as that of my high school—I was not able to get admitted to any of the US PhD programs, and instead, I enrolled in a master's program at Leiden University in the fall of 2019. This transition turned out to be a better choice for my career development than what I had planned. At Leiden Observatory, I had the opportunity to do my first master's research project in the group of Prof. Ewine van Dishoeck, who is referred to as *the Queen of Astrochemistry* in the community. Thanks to this experience, I rejoined her group as a PhD student in late 2021. According to the initial Education & Supervision Plan, my PhD research was supposed to study the formation of complex organic molecules in both the gas and the solid phases using ALMA and JWST—yes, my PhD progressed exactly as planned (though probably not as fast as expected). After years of delay, JWST was successfully launched two months after I started my PhD, and I was therefore able to work on first-hand JWST data just after finishing my first ALMA project. In the last year of my PhD, I was very lucky to be granted observing time of both ALMA and JWST as the Principle Investigator (PI), which, as far as I'm concerned, is one of the highest "honors" an observer can receive. I would not have reached this point if I had not chosen to come to Leiden, where I met such a professional and supportive group. Currently, I'm starting a postdoc position in the Center for Astrochemical Studies (CAS) at Max Planck Institute for Extraterrestrial Physics (MPE). I will keep exploring the icy and gaseous worlds of interstellar molecules using the most powerful telescopes in human hands.

Time flies. I still remember it was a clear night in late October of 2012 that I recognized Orion for the first time. Thirteen years later, I'm finally approaching the completion of my doctoral degree. It's been a long journey, beginning with dreams and hopes, but soon filled with challenges, uncertainties, and moments of self-doubt. I considered quitting more than once, yet somehow I arrived here, where I can contentedly say to my fifteen-year-old self: I have seen the scenery you once longed for.

[‡]Both of my parents majored in Chinese language and literature (yes, nothing to do with STEM).

Acknowledgements

In Chinese, my name Yuan(缘) refers to fortunate encounters shaped by both destiny and coincidence. Throughout this journey, I have been very lucky to cross paths with many great people, whose guidance, support, and encouragement have brought me to where I am today. I would like to express my sincere gratitude to everyone who has illuminated my path, including those who may not be individually acknowledged here.

First and foremost, I would like to thank my supervisors, Ewine and Harold, for offering me the opportunity to pursue my PhD, and InterCat, the Center for Interstellar Catalysis at Aarhus University, for providing me the funding over the four years. Ewine, thank you for always taking time out of your busy schedule to meet with me, read my manuscripts, and teach me how to give good presentations. You never lost faith in me, even when my progress was slower than expected. I know from your mentorship that you genuinely hope we can find our own place in academia, rather than simply meeting the requirements to graduate. It was a privilege to be one of your last PhD students, and I hope I have not let you down. Harold, I regret that I did not take more initiative to interact with you before your passing. Thank you for running the Laboratory for Astrophysics, which has been continuously fueling our observational research by providing important lab spectra.

During my six years at Leiden, I also received generous help from many senior group members. Łukasz and Martijn, thank you for guiding me through my first master's project step by step. Martijn, I'm very happy to have you as my co-promoter at my defense! I owe so much to your guidance and dedication. Thank you for staying in the group as a postdoc so that our time overlapped for longer, although I felt sorry that you were always copied on the emails Ewine sent me. Pooneh, I learned a lot from you, not only about doing research, but also about how to grow into an independent researcher. Will, I couldn't imagine how I would have moved forward without your pioneering studies. Thank you for always being patient and helpful whenever I had questions. Łukasz, it was amazing that our paths crossed again, and it is a privilege to have you in the group as an expert in both ALMA and JWST. Logan, thank you for sharing your scripts and humor, both of which are necessary to my life.

The companionship with my fellow students has also meant a lot to me. Katie, you are one of the best people I have met, both as a friend and as a researcher. I will miss our chats about everything and nothing. Nashanty, I love your hearty laugh, even when you were teasing me at the NOVA Fall School. Julia, your diligence and efficiency are truly admirable. I was lucky to have you as the co-mentor of our master's student Kelly. Chen and Yanling, I fondly remember the dinners we shared, the PhD events we attended, and the concert and volleyball matches we watched together.

Beyond those I worked with closely, I would also like to thank the rest of the Leiden-Garching group: Alice, Andrew, Sierra, Nicolás, Franciele, Zak, Xunchuan, Lisa, Marina, Andrés, Margot, Ardjan, Lukas, Teresa, Giulio, Pranjal, Marissa, Milou, for the time we shared at L-G telecons, group retreats, and conferences over the world.

Thank you for broadening my knowledge of disks, jets, simulations and experiments. Femke, Lukas, Tobias, and Mulan, welcome to the group—the torch is now in your hands. Special thanks to other faculty members, Serena, Michiel, Melissa, Nienke, Ko-Ju, Thanja, for keeping astrochemistry thriving at Leiden. Alessia and Josh, thank you for your hard work as TAs for SSE. I learned a lot from working with you. Beth and Karin, you are both fast learners; good luck with your future duties!

Outside Leiden, I also received warm hospitality and valuable guidance. Liv, Karin, Andrew, and all the InterCat folks, thank you for organizing the seminars and hosting us warmly at the yearly retreat. Greg, thank you for supporting me all the way until I got my PhD. I was so lucky to have you as my first academic mentor. Ted, it was my pleasure to see you again after graduation. Doing the bachelor's thesis project with you was the start of my astrochemistry journey. Brett, thank you for your wonderful ALMA data, which became the backbone of this thesis, and for taking care of my JWST program as the US Admin Co-I. Rob, thank you for all the simulation results and your detailed replies to my questions, which made my observational studies much more meaningful. Of course, thank you Silvia, for recognizing my work and offering me the postdoc position in your group. I would also like to express my gratitude to the professors, researchers, and students I met in China. Thank you for your warm welcome and financial support. I really enjoyed our academic discussions and social activities, and I hope to be able to return your kindness in the future.

As someone who spent a lot of time cooking and sleeping, my life outside academia was not very colorful. Even so, I would like to thank my Chinese colleagues for dining out together, which has been an important part of my spare time. Yapeng, Xuechen, Shunsheng, Zhenlin, thank you for your guidance in both life and research. Also thanks to many of my neighbors and friends, with whom I spent enjoyable holidays.

Before closing, I would like to thank my parents, for always respecting and supporting my decisions, even though I chose a less conventional path. Thank you for being my harbor whenever I need to return. As the only child, I wish I could spend more time with you. Finally, thank you—myself—for staying grounded and keeping moving forward, no matter how slow the pace. It was not easy to pursue a PhD abroad, alone, and during the pandemic, but you made it. May the road ahead be kind to you.