Bridging the gap between physics and chemistry in early stages of star formation

Nazari, P.

Chapter 1

Introduction

The urge to understand our place in the Universe on both spatial and temporal axes, is one of the primary characteristics that sets us apart from the other inhabitants of this little blue dot. This curiosity to find answers to questions such as *what is the origin of the Universe?* and *what is the origin of Life?* has immensely impacted our lives in spiritual, philosophical, and technological manners, from cave people to the modern-day scientists. This evolution has only become possible through the advancement in technology and engineering to build increasingly powerful telescopes and instruments, in addition to the advancement in various scientific fields such as physics, chemistry, biology, and more importantly, combinations of these subjects. In particular, physics and chemistry are appropriately intertwined in the field of molecular astrophysics (or astrochemistry) which is the theme of this thesis.

Starting from the *origin of the Universe* and its connection to astrochemistry, just after Big Bang there were no atoms or molecules because the Universe was so hot that they could not exist (Weinberg 1977). It took about 3 minutes to become cool enough that nucleosynthesis could take place and form fully ionized hydrogen, deuterium, helium and a small amount of lithium (Fig. 1.1 left panel). The Universe kept expanding and cooling until the nucleosynthesis stopped. Later at a redshift (z) of around 1000 the densities and temperatures were low enough for the neutral molecules to form (Dalgarno 2006) and at $z{\sim}20-30$, first stars, which then formed galaxies and among those the Milky Way (Fig. 1.1 middle panel).

Moving to the *origin of Life* and its connection to astrochemistry, there are many chemical reactions that occur in the molecular clouds of our Galaxy (Fig. 1.1 right panel). These clouds are where stars and planets are born which again have their own unique chemical environment. Assuming that life originates on planets, it is impossible to learn about life unless we learn about the physical and chemical laws that govern the process of star and planet formation. This thesis attempts to bridge between chemistry and physics of the nascent stellar systems that later turn

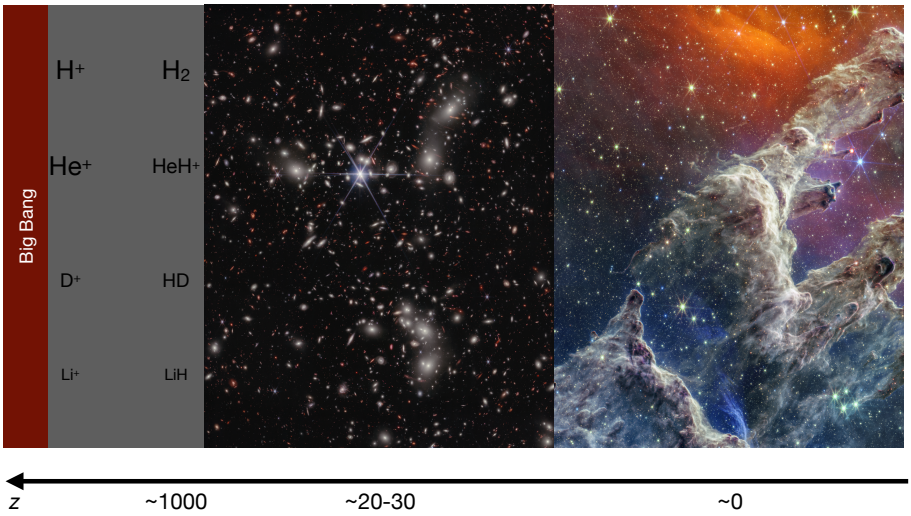

Figure 1.1: A schematic for the origin and formation of the Universe as we know it. The middle panel is the image of massive galaxy cluster Abell 2744 taken by the *JWST*-NIRCam (NASA, ESA, CSA, and STScI, I. Labbe and R. Bezanson). In the background of this image some of the highest redshift galaxies known to date have been discovered. The right most panel is an image of the Pillars of Creation taken by *JWST*-NIRCam and *JWST*-MIRI (NASA, ESA, CSA, and STScI).

into planetary systems like our own Solar System. Here, observations of Atacama Large Millimeter/submillimeter Array (ALMA) and *James Webb* Space Telescope (*JWST*) are combined with radiative transfer models to bring us one step closer to understanding our origins.

1.1 Star formation

1.1.1 Low-mass stars and planetary systems

Formation of low-mass stars begins from large molecular clouds where the densities ($\sim 10^4 - 10^6\,\mathrm{cm}^{-3}$) and temperatures ($\sim 10 - 20\,\mathrm{K}$) are relatively low (see review by Bergin & Tafalla 2007). These clouds are not completely smooth and the over-dense regions can start collapsing when the Jeans mass is reached (Jeans 1928) due to for example an external trigger or turbulence. The Jeans mass is defined as the mass threshold at which the inner gaseous pressure can no longer support gravity. Collapse continues while the temperature stays constant until the core becomes optically thick to radiation. Therefore, the radiation pressure stops the collapse and a First Hydrostatic Core forms (Larson 1969). Given that radiation is trapped inside, the temperature keeps rising until it is high enough for H_2 to dissociate ($\sim 2000\,\mathrm{K}$). Hence, the core cools down and collapse continues again to form a protostar.

This newly born protostar will go through four subsequent phases before turning into a main-sequence star hosting potentially a planetary system (see Fig. 1.2). Technically, there is a difference between the *Stages* and *Classes* (Lada 1987; Shu et al. 1987; André et al. 1993; Robitaille et al. 2006) that the protostellar system goes through. Stages are defined based on parameters such as age, envelope mass, disk mass and protostellar mass while Classes are defined based on the observed flux. However, generally there is a good match between them and hence in this thesis the two words are used interchangeably. It is worth noting that although this is generally true, sometimes there can be a mismatch between the two because factors such as orientation of the object can affect the observational parameters (e.g., Crapsi et al. 2008). The Class system is defined by the slope of the spectral energy distribution of the system at infrared wavelengths (Lada 1987)

$$\alpha_{\rm IR} = \frac{d\log \lambda F_\lambda}{d\log \lambda}, \tag{1.1}$$

where F_λ is the observed flux at wavelength λ between $\sim 2 - 25\,\mu$m.

The infrared slope is not defined for a Class 0 system because these systems are very faint in infrared wavelengths. However, other parameters have been defined to classify these sources, such as bolometric temperature ($T_{\rm bol}$) and ratio of submillimeter luminosity to bolometric luminosity ($L_{\rm submm}/L_{\rm bol}$; André et al. 1993; Chen et al. 1995). A Class 0 system has a bolometric temperature of less than 70 K and a $L_{\rm submm}/L_{\rm bol}$ of $> 0.5\%$. In this stage the envelope mass is much higher than the mass of the central protostar. As the rotating infalling envelope collapses, due to conservation of angular momentum the envelope flattens and a disk forms. However, if angular momentum is not extracted from the disk, accretion cannot continue, while accretion rates are found to be relatively high in this stage ($> 10^{-6}\,{\rm M}_\odot\,{\rm yr}^{-1}$; see review by Hartmann et al. 2016). Therefore, two mechanisms, viscosity or disk winds, are suggested to extract angular momentum from the (inner) disks (Lynden-Bell & Pringle 1974; Blandford & Payne 1982). The disks that have just started forming in this phase are relatively small (Segura-Cox et al. 2018).

A Class I system is defined as one with $\alpha_{\rm IR} > -0.3$ but sometimes a further classification defines systems with $0.3 > \alpha_{\rm IR} > -0.3$ as flat spectrum sources and only those with $\alpha_{\rm IR} > 0.3$ as true Class I (Greene et al. 1994). These systems have their $T_{\rm bol}$ between 70 K and 650 K. The disk is more evolved in this phase and the protostar has accreted more from the envelope such that now $M_\star > M_{\rm env} > M_{\rm disk}$. Stages 0 and I are expected to last about $10^4 - 10^5$ years (Kristensen & Dunham 2018). Most of the work of this thesis is on Class 0 and I (also Stage 0 and Stage I) protostellar systems, where the envelope mass is still significant and the central protostar is embedded within it. In the following chapters these two phases are sometimes referred to as the *protostellar phase* or the *embedded phases of star formation*.

A Class II system is defined as that with $-1.6 < \alpha_{\rm IR} < -0.3$ and $T_{\rm bol}$ between 650 K and 2800 K. At this stage the envelope has already dissipated and $M_\star > M_{\rm disk} > M_{\rm env}$. The disk in this phase is much more evolved and is called a *protoplanetary disk* with highly structured gas and dust profiles (e.g., Casassus

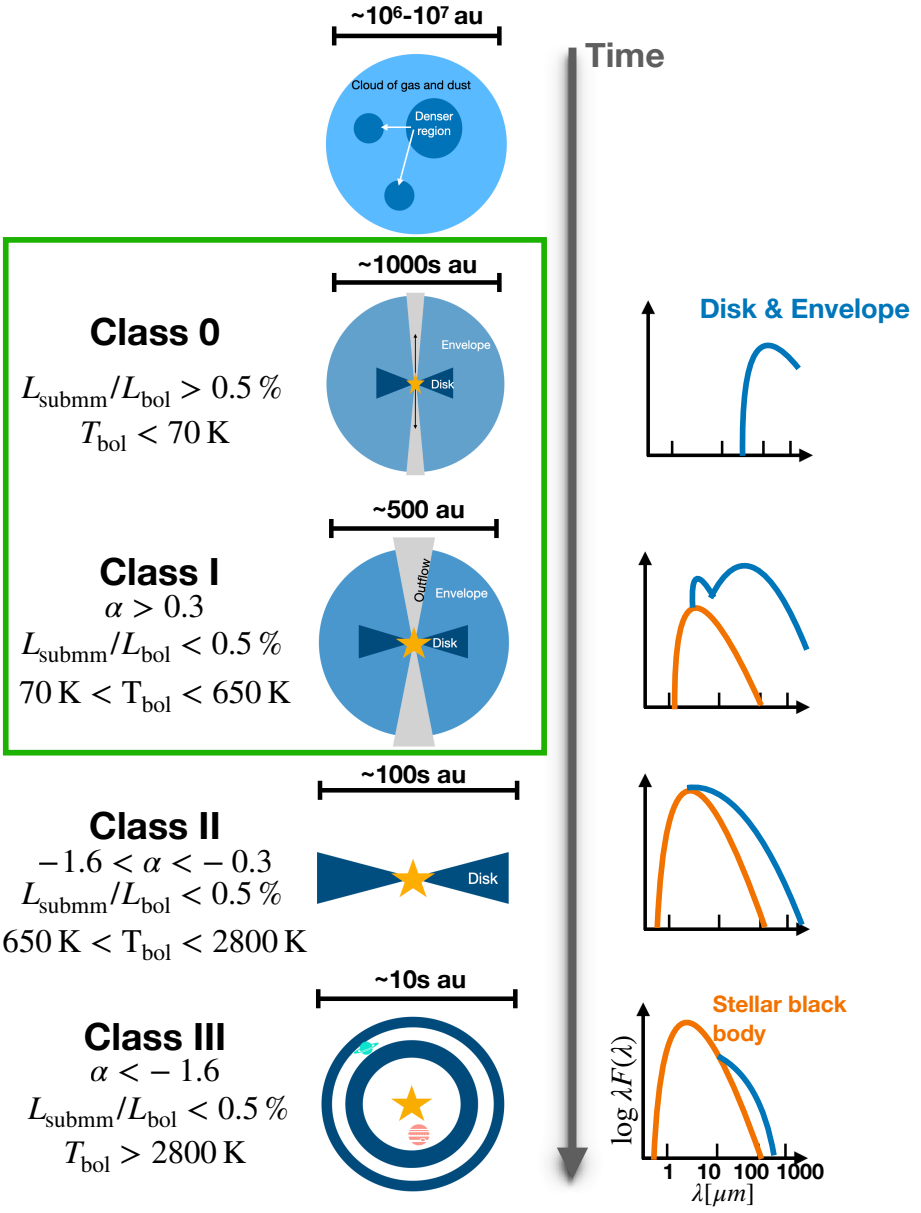

Figure 1.2: A schematic view of various phases of star formation on the left. A rough spectral energy distribution diagram for each Class is also presented on the right. The focus of this thesis is the early phases of star formation (Class 0 and I) which are indicated by a green box.

et al. 2013; Qi et al. 2013; van der Marel et al. 2013; Fedele et al. 2017; Andrews et al. 2018; Rosotti et al. 2021). The ages of protoplanetary disks are on the order of 10^5 to a few 10^6 years and the jets discovered in these objects are atomic and ionized (Podio et al. 2011).

Finally, Class III disks (also known as debris disks) are much fainter almost gas-poor disks (see review by Hughes et al. 2018). In this phase all the material in the disks is already turned into planets or dissipated by photoevaporation and therefore, the small amount of dust left from collisions of planetesimals as a *debris disk* will produce the infrared excess with $\alpha_{\rm IR} < -1.6$. The bolometric temperature is also relatively high ($T_{\rm bol} > 2800\,{\rm K}$). In this stage the central pre-main sequence star keeps contracting until it starts hydrogen fusion and reaches the main sequence. This phase lasts about 10^6 to a few 10^7 years and once (almost) all the remnant disk is gone the final planetary system lives for $> 10^7$ years.

1.1.2 High-mass stars

Formation of high-mass stars is not as well-understood as their low-mass counterparts. This is because these objects are less commonly observed which in turn can be attributed to their fast evolution ($\sim 10^4 - 10^5\,{\rm yr}$) and their larger distances (Beuther et al. 2007; Mottram et al. 2011; Lumsden et al. 2013). Their formation is still under debate with various theories proposed. However, the main two theories for their formation are monolithic collapse (or core accretion) and competitive accretion (Bonnell & Bate 2006; Myers et al. 2013; Tan et al. 2014; Motte et al. 2018). The former theory proposes a similar formation mechanism to the low-mass counterparts. The latter requires formation of these massive objects in clusters of low-mass stars, where the one in the center of the gravitational potential is expected to become the most massive.

For the purposes of this thesis, it is important to highlight two characteristics of these proposed theories. First, disks are expected to form in both theories, although the stability of such disks is under debate (Ahmadi et al. 2019; Johnston et al. 2020). Second, both theories predict a *hot core* phase, which is similar to the Class 0 phase of low-mass star formation. In other words, it is expected that at some point in their lifetime a highly embedded and active protostar warms up the inner envelope (and the disk). An important difference in this phase between low- and high-mass protostars is the start of formation of an inner hypercompact HII region where the gas is completely ionized due to significant amounts of ultraviolet (UV) photons emitted by the central protostar (e.g., Keto 2003). These regions are defined to have extents of less than 0.05 pc (Kurtz 2005).

1.2 Chemical evolution

Understanding how chemistry evolves as a protostar forms provides important clues on constraining the physical conditions at each stage. To learn about the chemical evolution from clouds to later stages, it is important to start with the initial elemental abundances in the interstellar medium (ISM). The gas in molec-

ular clouds is mainly composed of hydrogen (∼90%) and helium (∼10%) with much lower abundances of other elements. From those the next most abundant ones are oxygen, carbon and nitrogen with abundances of $\sim 4 \times 10^{-4}$, $\sim 3 \times 10^{-4}$ and $\sim 7.5 \times 10^{-5}$ (Wilson & Rood 1994). The dust in these clouds only comprises around 1% of the total mass and is made of elements such as silicon and iron which have abundances of $\lesssim 3 \times 10^{-5}$. However, these dust grains and lower-abundance elements have an important role in advancing the chemistry and formation of major species such as H_2, H_2O and NH_3 (Tielens 2013; van Dishoeck et al. 2013). The following sections dive deeper into the chemical reactions that start taking place in the molecular clouds and continue to the protostellar phase.

1.2.1 Prestellar phase

The prestellar phase (see top panel of Fig. 1.2) starts from diffuse molecular clouds where the densities are at $\sim 10^2\,\mathrm{cm}^{-3}$. The denser regions in these clouds can have densities of $\sim 10^3\,\mathrm{cm}^{-3}$ and temperatures of $\gtrsim 20\,\mathrm{K}$ with $A_\mathrm{v} \sim 1 - 2\,\mathrm{mag}$, this is called the translucent phase (e.g., van Dishoeck et al. 2014). At this stage atoms such as oxygen, carbon and nitrogen stick to the dust grains and later can be hydrogenated through diffusion of hydrogen atoms over the surface. This process can produce molecules such as H_2O, CH_4 and NH_3 on the surface of dust grains as ices in this translucent phase (e.g., Caselli & Ceccarelli 2012; Tielens 2013; Cuppen et al. 2017).

Later as these denser regions collapse, the densities increase to $\gtrsim 10^4\,\mathrm{cm}^{-3}$ and the temperatures drop to below $\sim 15\,\mathrm{K}$ ($A_\mathrm{v} \gtrsim 9\,\mathrm{mag}$). At these cold temperatures CO that is formed in the gas freezes out onto the dust grains, producing a second layer of ice on top of the previously made layer of H_2O mixed with CH_4 and NH_3 (see review by Boogert et al. 2015). It is important to note that these dense and cold conditions are also present far from the central protostar (∼1000 s au) in the outer envelope. Therefore, similar chemical conditions are expected at those far distances even after the birth of the protostar (see Fig. 1.3; Herbst & van Dishoeck 2009). After the CO freeze-out other more complex molecules such as HCO and H_2CO (formaldehyde) can form in these ices through hydrogenation of CO. An important and one of the simplest *complex organic molecules* (COMs) that is thought to form through this mechanism is methanol (CH_3OH; Watanabe & Kouchi 2002; Fuchs et al. 2009; Santos et al. 2022). A COM is defined as a carbon containing molecule with 6 or more atoms (Herbst & van Dishoeck 2009). At this dense and cold phase chemical complexity does not stop at methanol. Laboratory experiments show that molecules as complex as those with > 8 atoms can already form in this phase through hydrogenation or reaction of precursors of methanol such as CH_2OH and HCO. Examples of such species are glycolaldehyde ($HC(O)CH_2OH$), ethylene glycol ($H_2C(OH)CH_2OH$) and glycerol ($HOCH_2CH(OH)CH_2OH$; Fedoseev et al. 2015a; Fedoseev et al. 2017). Other important large species that could form in this phase are three-carbon species such as propanal (CH_3CH_2CHO) and 1-propanol ($CH_3CH_2CH_2OH$; Qasim et al. 2019a).

An important recent discovery in laboratory experiments is the possibility of forming some of the COMs such as acetaldehyde (CH_3CHO) and ethanol (C_2H_5OH)

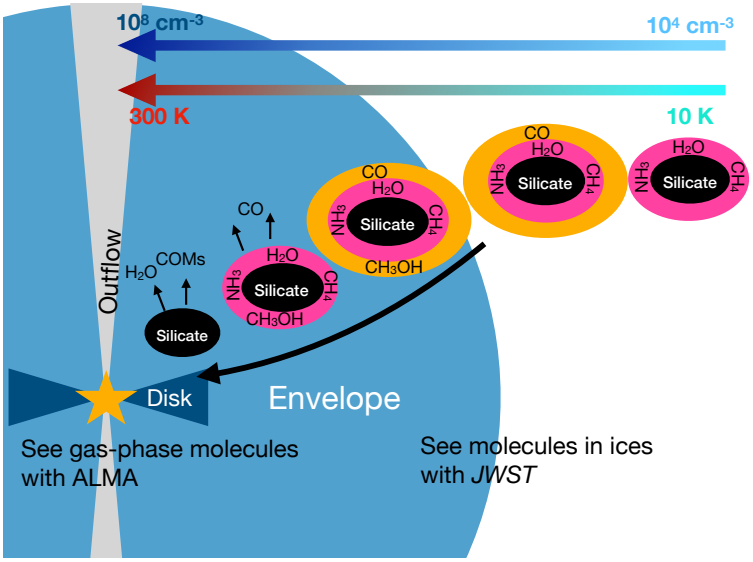

Figure 1.3: A schematic view of evolution of chemistry from the cold prestellar phase toward the protostar (figure after Herbst & van Dishoeck 2009).

starting from acetylene (C_2H_2) rather than species related to CO, already prior to the CO freeze-out stage (Chuang et al. 2020, 2021). More complex species such as n- and i-propanol ($CH_3CH_2CH_2OH$ and $H_3CCHOHCH_3$) and n- and i-propenol ($H_3CCHCHOH$ and $H_3CCOHCH_2$) are also found to form under conditions that may be representative of prior to the CO freeze-out in the laboratory (Qasim et al. 2019b). However, C_2H_2 is not yet detected in the early phases of star formation.

Formation of nitrogen-bearing COMs, however, is less understood. Perhaps one reason for that is the ∼5 times lower elemental abundance of nitrogen compared with oxygen in the ISM. Therefore, it has been more difficult to observe nitrogen-bearing species. Nevertheless, apart from NH_3 which forms during the translucent phase, OCN^- is another important simple molecule that is thought to form in the dense and cold phase from NH_3 and isocyanic acid (HNCO) through acid-base reactions in the ice (Schutte et al. 1999; Novozamsky et al. 2001; Schutte & Khanna 2003). An important group of nitrogen-bearing molecules are those with a CN bond, such as OCN^-, but also hydrogen cyanide (HCN), methyl cyanide (CH_3CN) and ethyl cyanide (C_2H_5CN). Recent laboratory work has found that starting from pure CH_3CN ice in translucent or dense core phases many other simple and complex species such as HCN or HNC, C_2H_3CN and C_2H_5CN can form through UV photolysis. Other complex nitrogen bearing species that can form during the prestellar phase are glycine (NH_2CH_2COOH) and methylamine (CH_3NH_2; Ioppolo et al. 2021). This work suggests that during the translucent phase, NH_2CH_2 is formed along with methane and ammonia which can contribute to formation of methylamine and glycine.

1.2.2 Protostellar phase and later

It is important to note that formation of many COMs was originally thought to take place in the later warm-up and the protostellar phase, where the higher dust temperatures increase the level of diffusion through ice (Garrod & Herbst 2006; Garrod 2013). Thermal processing during the warm-up phase has proven effective in forming COMs such as aminomethanol (NH_2CH_2OH; Theulé et al. 2013). Moreover, UV-induced chemistry of methanol-rich ices is found to be effective in producing O-bearing COMs such as ethanol (C_2H_5OH) and dimethyl ether (CH_3OCH_3; Öberg et al. 2009b). However, more recent laboratory and modeling studies seem to favor the earlier phases both the translucent as well as the dark and dense cloud phases (see Sect, 1.2.1). For example the most recent version of MAGICKAL models find that the bulk of many oxygen- and nitrogen-bearing COMs form in the dense and cold phase (Garrod et al. 2022). Nevertheless, they also find that a significant amount of certain molecules can form in the later warm-up phase closer to the protostar in ices (Fig. 1.3). Two examples of such molecules are formamide (NH_2CHO) and isocyanic acid, which seem to form in the warm-up stage of Garrod et al. (2022) after initial formation in the cold phase.

Closer to the protostar as the temperatures gets higher, the various ice layers start desorbing (Fig. 1.3) and result in the observed rich spectra of many of the already discussed molecules in the gas phase around protostars. In particular, at first molecules that are less bound to the ices (i.e., lower binding energies) such as CO, N_2, and CH_4 start sublimating and molecules with higher binding energies such as water sublimate closer to the protostar at $T \gtrsim 100\,\mathrm{K}$. In this high-temperature gas, potentially more complex chemistry can take place, however, the importance of these high-temperature gas-phase routes is debated (see reviews by Herbst & van Dishoeck 2009; van Dishoeck 2014; Ceccarelli et al. 2022).

In particular, there are a few relevant molecules for this thesis whose formation in the gas or on ices is argued to this date. One of these molecules is NH_2CHO. Both gas and ice formation pathways are suggested for this molecule (Raunier et al. 2004; Jones et al. 2011; Barone et al. 2015; Codella et al. 2017; Skouteris et al. 2017; Dulieu et al. 2019; Chuang et al. 2022; Douglas et al. 2022) and some studies argue for its formation being linked to HNCO (Haupa et al. 2019). Two other examples are acetaldehyde (CH_3CHO; Vazart et al. 2020; Chuang et al. 2021; Fedoseev et al. 2022) and CH_3CN (Huntress & Mitchell 1979; Willacy et al. 1993; Mackay 1999; Garrod et al. 2008; Walsh et al. 2014), where both ice and gas formation mechanisms have been suggested.

After the protostellar phase, when the envelope is completely dissipated (see Sect. 1.1.1), the observed gas-phase spectra of the protoplanetary disks become 'deserted' compared with what is observed in the protostellar phase (Fig. 1.4). It is still debated whether the reason for this is simply lower disk temperatures resulting in all the previously observed COMs to freeze out onto the dust grains (*inheritance scenario*) or the previously observed COMs are destroyed (*reset scenario*) and chemistry occurs from simple species in situ (see review by Öberg et al. 2023). Nevertheless, in recent years a few COMs have been discovered in protoplanetary disks (Öberg et al. 2015; Walsh et al. 2016; Booth et al. 2021; Brunken et al. 2022).

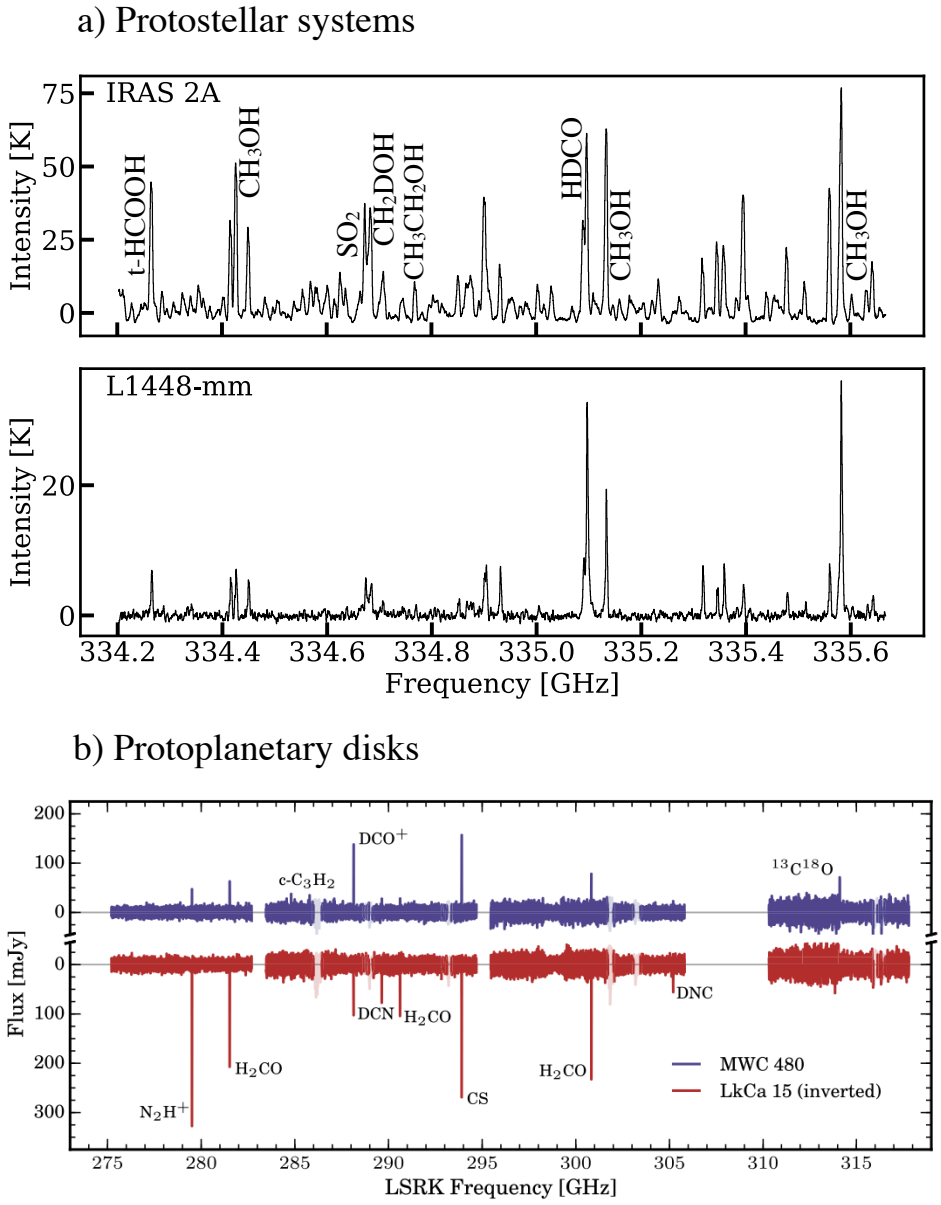

Figure 1.4: *a)* ALMA gas-phase spectra of the Class 0 protostellar systems of IRAS 2A and L1448-mm at ∼1 mm (project ID: 2021.1.01578.S; PI: B. Tabone). The spectrum is relatively line rich for both sources with more lines toward IRAS 2A. *b)* ALMA gas-phase spectra of two protoplanetary disks from Loomis et al. (2020). Note the large range in frequency and low line density for these two sources (less than one line per GHz) compared with the protostellar systems in panel *a* (more than ∼25 lines per GHz).

Although the origin of some of these detected COMs is still debated, IRS 48 is one of the unique sources with an icy dust trap irradiated by the central protostar resulting in the COMs sublimating from the ices (Brunken et al. 2022).

1.3 Techniques

1.3.1 Observations with ALMA and JWST

Two important telescopes used in this thesis are ALMA and *JWST*. Here ALMA is mainly used for gas-phase emission lines of simple and complex molecules in the inner regions of the protostellar systems. On the other hand, *JWST* is used for ice observations of those molecules in absorption. Figures 1.4 and 1.5 present examples of ALMA and *JWST* spectra and how various species are identified through narrow emission or broad absorption features. It is worth noting that both telescopes are used for other types of observations such as infrared gas-phase observations with *JWST* or detailed dust continuum analysis with ALMA which are not the focus of this thesis. Two particularly important characteristics of ALMA and *JWST* for the work of this thesis are their high sensitivity and their long wavelength range coverage. The former is needed to observe the optically thin isotopologues of the most abundant species with bright emission lines such as methanol and methyl cyanide with ALMA and to observe any COM other than methanol with *JWST*. The latter is necessary to study protostellar systems where dust extinction is an issue for gas-phase molecular observations in the inner regions.

The ground based telescope, ALMA, is an interferometer that is composed of many dishes (or antennas). The angular resolution of a telescope is proportional to λ/D, where λ is the wavelength and D is the diameter of the telescope which for an interferometer can be substituted by the baseline (i.e., the distance between antennas). The sensitivity on the other hand, depends on the collecting area of the telescope and quality of the detectors. The noise scales as $1/(t_{\rm int}\Delta\nu)$, where $\Delta\nu$ expresses the frequency range over which the sensitivity is calculated and $t_{\rm int}$ is the integration time of the observations. For the type of ALMA observations studied here, a sensitivity of $\sim$5 mJy km s^{-1} in a $\sim0.5''$ beam is needed. ALMA observes the visibilities in the (u,v) plane where u and v are defined as the length of the projected baseline between antennas on the sky in units of wavelength. These are then Fourier transformed to obtain the intensities in the image plane (see Jackson 2008 for more detail on basics of interferometric observations). The 54 12-m antennas of ALMA can be distributed in various configurations to give different angular resolutions ($\sim0.01'' - 9''$) and the 12 7-m dishes are used in a compact configuration for large scale observations. These configurations make observations at different frequency ranges (or at different Bands) from 35 GHz to 950 GHz. Moreover, ALMA can have spectral resolution of as high as around 0.03 MHz which corresponds to $\sim$0.26 km s^{-1} at the lowest ALMA frequencies and $\sim$0.01 km s^{-1} at the highest ALMA frequencies. Therefore, the gas-phase emission lines from the protostellar systems (normally FWHM of $\sim1-3$ km s^{-1}) can be fully resolved. This thesis uses ALMA Bands 3, 6, and 7 (wavelengths ranging from $\sim$3 mm to $\sim$0.8 mm) at moderate to high angular resolutions (i.e., $\sim0.1'' - 0.5''$).

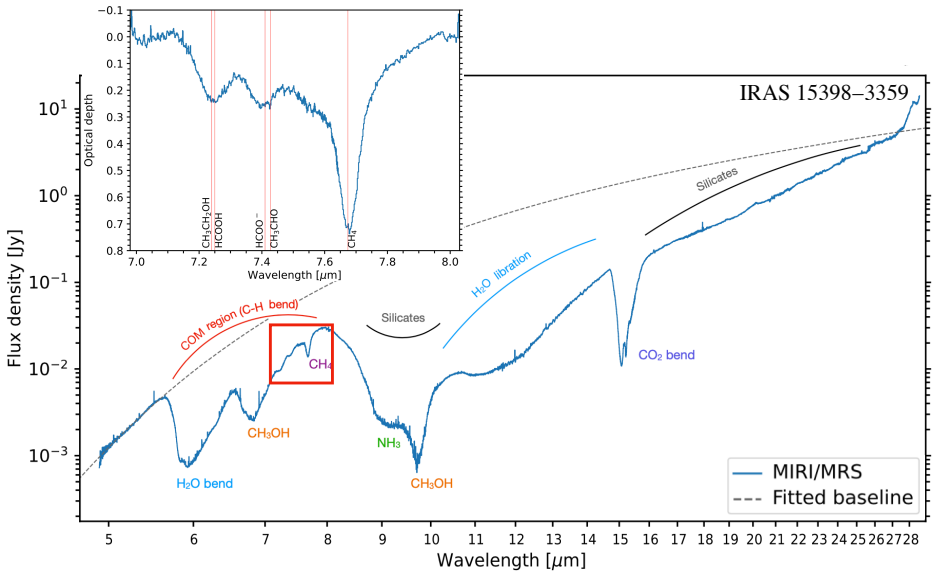

Figure 1.5: Ice spectrum of the Class 0 protostellar system IRAS 15398-3359 with *JWST*-MIRI showing many ice absorption features. *Inset:* Zoom-in of the 7-8 μm regime (indicated by the red box) where ice features of CH_3CH_2OH and CH_3CHO are identified (figure adapted from Yang et al. 2022).

On the other hand, *JWST* has a single primary mirror with a size of $\sim$6.5 m that gathers the light and reflects it onto a secondary mirror which focuses the light into the Integrated Science Instrument Module. This telescope is arguably the most powerful space science telescope to date. It has four instruments: Near-Infrared Camera (NIRCam), Near-Infrared Spectrograph (NIRSpec), Near-Infrared Imager and Slitless Spectrograph/Fine Guidance Sensor (NIRISS/FGS), and Mid-Infrared Instrument (MIRI). The first three instruments observe a wavelength range of $\sim 0.6 - 5\,\mu$m, while MIRI covers the wavelength range of $\sim 5 - 28\,\mu$m. Various spectroscopic modes and (angular and spectral) resolutions are available for each instrument. For spectroscopic observations of single protostellar systems the Integral Field Units (IFU) onboard MIRI and NIRSpec are the most useful. These two instruments can spatially resolve an object down to $\sim 0.2''$ and have a maximum spectral resolution of $R\sim$3500 (Rieke et al. 2015; Jakobsen et al. 2022). For the purposes of analyzing the gas-phase emission lines and subtracting them from the spectra to reveal the ice features in the $\sim 3 - 28\,\mu$m range, spectroscopy with medium or high spectral resolution ($R \gtrsim 1000$) is needed. This corresponds to medium-resolution spectroscopic mode on MIRI (i.e., MIRI-MRS) and NIRSpec-IFU disperser/filter combination of G395H/F290LP. Because the ice features are relatively broad they get spectrally resolved, while the gas-phase lines are narrow and cannot be fully resolved.

1.3.2 Radiative transfer and column density measurement

One of the most important concepts in astrophysics is understanding how light propagates through a medium before reaching us. This is governed by the radiative transfer equation which in the case of no scattering is given by

$$\frac{dI_\nu}{ds} = -\kappa_\nu(s)I_\nu + \epsilon_\nu(s), \tag{1.2}$$

where I_ν is the intensity at frequency ν along the path s, κ_ν is the absorption coefficient and ϵ_ν is the emission coefficient. This can be interpreted as follows: the intensity that reaches us from an object is affected by absorption and emission along its path. There are codes that use Monte Carlo techniques to reproduce the dust temperature and emission from molecules. One of these Monte Carlo radiative transfer codes is RADMC-3D (Dullemond et al. 2012) which has been used in *Chapters 7-9* of this thesis. RADMC-3D calculates both the dust temperature and line emission from the gas by assuming that the dust and the gas are thermally coupled. The dust temperature is calculated by solving Eq. (1.2) for all frequencies. In simple words RADMC-3D takes a fixed number of photons as input, calculates their energies based on the total luminosity and emit them one by one. As a photon hits a cell it can be absorbed or scattered, if it is absorbed, it is immediately re-emitted at another direction and another wavelength (Bjorkman & Wood 2001). Given that absorption of a photon will increase the energy of each cell, it also increases the temperature. The initial photon will go through the grid in this manner until it escapes the edge of the grid. The final temperature is what remains after all photons have gone through the grid. The line radiative transfer implies computing images or spectra of a molecule at a specific set of frequencies. This is more complex to explain in simple words and thus only some basics are touched upon.

There are two ways to excite a gas-phase molecule's various vibrational and rotational states: thermally (i.e, collisionally) or radiatively (see van der Tak et al. 2007 for more detail). In both cases the molecule can emit a photon afterwards to de-excite but in high-density environments collisions dominate excitation and de-excitation. This condition is called local thermodynamic equilibrium (LTE) where the emission from molecules can be explained by Boltzmann distribution at a single excitation temperature. In low-density environments on the other hand, radiation dominates excitation and de-excitation.

Rotational states are only available for gas-phase molecules as they can freely tumble in space. However, the vibrational motions are available for both gas-phase and solid-state molecules. Solid-state molecules that are sufficiently refractory to have their vibrational modes thermally excited without sublimating can emit photons with an intensity high enough to be detectable. An example of this phenomenon is the crystalline water ice emission feature at $\sim 45 - 63\,\mu$m (Malfait et al. 1999; Molinari et al. 1999). However, generally ice features are only observed in absorption because the temperatures needed to excite them thermally is larger than their sublimation temperature. Therefore, they will be in the gas before they can be observed in emission as part of the ice matrix. Gas-phase molecules can

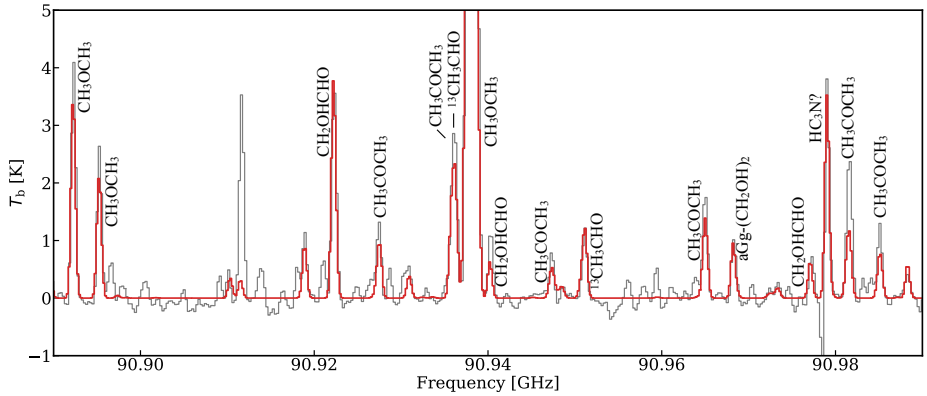

Figure 1.6: Spectrum of IRAS 16293-2422 B taken by Band 3 (∼3 mm) of ALMA (black). Red shows a CASSIS model as the sum of models for individual molecules.

also be observed in absorption and that usually occurs when a colder cloud of gas is in front of a hotter one that is emitting.

The emission from gas-phase molecules and thus the column density of those molecules in the gas can be found by starting from the initial radiative transfer equation (i.e., Eq. 1.2). Assuming no absorption and optically thin lines Eq. (1.2) can be re-written as

$$dI_\nu = \frac{h\nu}{4\pi} A_{\rm u,l} n_{\rm u}^{\rm thin} \phi(\nu) ds, \tag{1.3}$$

where h is the Planck's constant, $A_{\rm u,l}$ is the Einstein-A coefficient (radiative decay rate for a transition from an upper lever, u to a lower level l), $n_{\rm u}^{\rm thin}$ is the number density of molecules in energy level u, and $\phi(\nu)$ is the line profile which is normalized to 1 when integrated over all frequencies. In Eq. (1.3) the expression before ds is the emission coefficient (ϵ) in Eq. (1.2) which has been assumed as emissivity due to spontaneous radiative decay. Next, Eq. (1.3) can be integrated over s to turn $n_\nu^{\rm thin}$ into $N_\nu^{\rm thin}$ (column density of molecules in level u). During this integration the background emission is also assumed negligible. Then, the remaining equation can be integrated over the frequencies of the emitted line. If the frequency scale is changed to velocity scale ($\frac{d\nu}{\nu} = \frac{dV}{c}$) and the intensity is written in terms of the flux density and the solid angle subtended by the object (i.e., F_ν/Ω)

$$\int F_\nu dV = \frac{N_{\rm u}^{\rm thin} A_{\rm u,l} \Omega h c}{4\pi}. \tag{1.4}$$

The term $N_{\rm u}^{\rm thin}$ needs to be rewritten in terms of $N_{\rm tot}$. These two terms can be linked by using the Boltzmann population distribution and assuming LTE conditions

$$\frac{N_{\rm u}}{g_{\rm u}} = \frac{N_{\rm tot}}{Q(T_{\rm ex})} e^{-E_{\rm u}/k_{\rm B}T_{\rm ex}}, \tag{1.5}$$

where $g_{\rm u}$ is the upper state degeneracy, $E_{\rm u}$ is the upper state energy level, $k_{\rm B}$ is the Boltzmann constant, $T_{\rm ex}$ is the excitation temperature, and $Q(T_{\rm ex})$ is the partition function. Therefore, combining equations (1.4) and (1.5) the integrated line intensity can be written as a function of the total column density and excitation temperature. Hence, knowing the column density and excitation temperature the integrated line intensity can be calculated. By fitting the emission lines of a molecule assuming a Gaussian function for each line, one can extract the column density and excitation temperature. More complexity can be added to this approach by accounting for the line optical depth and background emission (for details see e.g., Möller et al. 2017; Loomis et al. 2018). Various tools have been developed for the purpose of fitting emission lines and extracting the column density and excitation temperature. One of these tools is CASSIS which has been used throughout this thesis (Vastel et al. 2015). An example of such fitting is presented in Fig. 1.6 where various molecules are first fitted separately and the final model in red is the sum of the individual models. A molecule is often reported as detected if it has at least three lines at a 3σ level with no over-prediction of emission lines in the spectrum. However, a larger number of detected lines will increase the accuracy of the determined column density and excitation temperature.

The column density of molecules in ices can be measured again starting from the basic Eq. (1.2). However, in the case of ices the emission coefficient can be set to zero. Integrating over that equation gives

$$I_\nu = I_\nu(0) e^{-\tau}, \tag{1.6}$$

which can be rewritten as

$$\tau = -\ln\left(I_\nu / I_\nu(0)\right). \tag{1.7}$$

The ice spectra are often converted into optical depth through Eq. (1.7). Then the optical depth is integrated over the absorption band and is related to the total column density through the band strength A by $N = \int \tau d\tilde{\nu}/A$, where $\tilde{\nu}$ is the wavenumber.

1.4 Protostellar environment

Protostellar systems consist of three main components: envelope, disk, and outflow. These three are shown in the sketch of Fig. 1.3. The gas-phase spectra of protostellar systems are generally line-rich due to the elevated temperatures (van 't Hoff et al. 2020b). Various molecules are known to trace a different component because of the different physical conditions present in each region which will govern the chemical reactions that can take place (Tychoniec et al. 2021). The following sections describe these three components and the observations toward them in more detail.

1.4.1 Envelope

The envelope around the protostar is composed of gas and ice-coated dust. The density and temperature in the envelope based on observations and radiative transfer models (Ulrich 1976; Adams & Shu 1985; Kristensen et al. 2012) generally follow a power-law relation with radius that decreases further from the central protostar.

1.4.1.1 Gas phase

In the gas the colder outer envelope is normally traced by $H^{13}CO^+$, N_2H^+, NH_3 and some of their deuterated isotopologues. On the other hand, some of the tracers of the warm inner envelope are H_2CS, OCS, and HCOOH. It is still not completely clear whether gas-phase COMs trace the warm inner envelope, the disk surface or the base of the outflow but they certainly trace the warm and hot inner regions after sublimating from the ices or direct formation in the gas ($\gtrsim 100\,\mathrm{K}$, see Fig. 1.3). In fact, similar to the simpler species, they are bound to their ice matrix at varying levels. Thus, those with lower binding energies sublimate from the ice at lower temperatures further from the central protostar and those with higher binding energies sublimate closer to the protostar, resulting in an expected onion-like structure (see Fig. 1.7). However, a complication is that the binding energies of COMs depend on the ice matrix that contains them (Collings et al. 2004; Minissale et al. 2022). For example, if they are mixed with water, they could have a similar sublimation temperature to water. Yet, this is not always the case, for example formamide (NH_2CHO) sublimates at higher temperatures than water even when mixed with water in laboratory experiments (Dawley et al. 2014; Urso et al. 2017; Chaabouni et al. 2018; Slavicinska et al. 2023). Therefore, it is challenging to determine their sublimation temperature in space without knowing their ice matrix.

From an observational perspective, COMs around low- and high-mass protostars have been observed in the gas phase for decades (e.g., Blake et al. 1987; van Dishoeck et al. 1995a). At first by single-dish ground based telescopes (e.g., Gibb et al. 2000; Cazaux et al. 2003) and later by space observatories such as *Herschel* (e.g., Crockett et al. 2014; Neill et al. 2014). However, those observatories lacked the sensitivity to detect the weaker, optically thin lines and because of their poor angular resolution the inner warm regions of protostellar systems were highly diluted in the large beams of those telescopes. In recent years, COMs are mostly observed by interferometers such as the Submillimeter Array (SMA), Northern Extended Millimeter Array (NOEMA), and ALMA with much higher angular resolution and sensitivity. Although dozens of oxygen-, nitrogen- and sulfur-bearing COMs have been detected toward protostars in the gas-phase (McGuire 2022), directly mapping their emitting regions around low-mass protostars has proven difficult. This is because these species generally have high sublimation temperatures and thus physical resolutions of $\sim 10 - 20\,\mathrm{au}$ are needed to map them. This is in contrast to simpler species such as $C^{18}O$ and N_2H^+, with sublimation temperatures of $\sim 20 - 30\,\mathrm{K}$, where it has been possible to map their *snowlines* with interferometric observations of physical resolutions of around 100s au (e.g., Jørgensen 2004; Anderl et al. 2016).

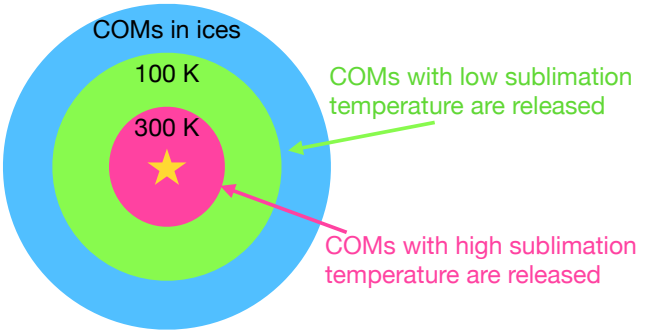

Figure 1.7: A schematic of the expected onion-like structure around a protostar as COMs with different sublimation temperatures are released into the gas from grains.

Nevertheless, segregation among various COMs around a few low- and high-mass protostars has been suggested in the literature for years based on the measured excitation temperatures of those species (Bisschop et al. 2007; Jørgensen et al. 2018; Manigand et al. 2020; Bianchi et al. 2022). The chemical separation in the hot regions around protostars has been directly mapped but mainly around massive objects (Wyrowski et al. 1999; Bisschop et al. 2008; Allen et al. 2017; Csengeri et al. 2019; Law et al. 2021). In the low-mass regime, HH 212 is one of the few sources where the emitting regions of a few COMs have been mapped with spatial resolution of around 20 au and found to be different (Lee et al. 2022). More high angular resolution observations of hot cores are needed to constrain the emitting regions of various COMs. This is particularly crucial because it directly affects our conclusions of gas-phase column density ratios and their comparison with ice observations (see *Chapters 2, 6 and 9*) given the difficulty to estimate the sublimation temperatures of COMs in space (see the start of this section).

Originally chemical complexity was analyzed in detail for single sources. The two 'poster child' low- and high-mass systems with extensive COM studies are IRAS 16293-2422 (hereafter IRAS16293) and Sagittarius B2 (hereafter SgrB2), where Jørgensen et al. (2020) review their chemistry. IRAS16293 is the low-mass protostar toward which methanol was first detected (van Dishoeck et al. 1995a) and many other studies targeted its COMs (e.g., Cazaux et al. 2003; Bisschop et al. 2008; Ceccarelli et al. 2010; Jørgensen et al. 2012; Kahane et al. 2013). Recently, the Protostellar Interferometric Line Survey (PILS; Jørgensen et al. 2016, 2018) studied COMs of this protostellar system in the entire Band 7 frequency range of ALMA ($\sim$329.147 − 362.896 GHz) and detected many COMs and their minor isotopologues in a low mass system for the first time (e.g., Coutens et al. 2016; Lykke et al. 2017; Calcutt et al. 2018a; Manigand et al. 2019, 2021; Coutens et al. 2022).

The Galactic Center massive star forming region, SgrB2, has also been subject of various molecular line surveys. A few COMs were detected for the first time in the ISM in this system already in the pre-ALMA era, for example methylamine (CH_3NH_2; Kaifu et al. 1974), acetic acid (CH_3COOH; Mehringer et al. 1997), and glycolaldehyde (CH_2OHCHO; Hollis et al. 2000). Later this system was subject of a line survey with the IRAM 30-meter in which more COMs were detected for the

first time in the ISM (Belloche et al. 2008, 2009, 2013). In recent years this region was observed by the Exploring Molecular Complexity with ALMA (EMoCA) survey and the Re-exploring Molecular Complexity with ALMA (ReMoCA) survey at higher angular resolution and sensitivity at $\sim 84.1 - 114.4\,\mathrm{GHz}$ (Belloche et al. 2016, 2019; Müller et al. 2016; Bonfand et al. 2017). These observations resolved the region into various hot cores and, because of their better angular resolution and sensitivity, provided more robust column densities for COMs.

Although detailed study and line surveys of individual sources are extremely valuable, in order to understand *how universal chemistry is* the COMs need to be analyzed in large samples of systems preferably scattered across the galaxy in various star forming regions. Such large samples with single dish telescopes already show interesting trends such as constant ratio of CH_3OCHO/CH_3OCH_3 across a sample with large mass and luminosity range (Coletta et al. 2020). However, single dish telescopes did not have the power to detect optically thin minor isotopologues of abundant species and suffer from beam dilution. Large samples with ALMA and NOEMA only became available in the past few years. In particular, the Perseus ALMA Chemistry Survey (PEACHES; Yang et al. (2021)) considered COMs in a sample of $\sim$50 protostellar systems. However, this survey did not go very deep and only detected the less abundant COMs in a handful of sources of their sample (< 10). Therefore, large samples of less abundant COMs are still missing for low-mass protostellar systems and thus a statistical analysis of COM chemistry. For high-mass protostars, *Chapter 3* of this thesis detects less abundant N-bearing COMs in a large sample of high-mass protostellar systems ($\gtrsim 30$) using the ALMA Evolutionary study of High Mass Protocluster Formation in the Galaxy (ALMAGAL) survey. Moreoevr, recently Chen et al. (2023) used the survey of Complex Chemistry in hot Cores with ALMA (CoCCoA) to provide column density ratios of six O-bearing COMs around 14 protostars and in general found constant ratios across a large mass and luminosity range which could point to a universal chemistry and bulk formation of these species in similar physical environments, likely the prestellar ices.

The recent ALMA surveys have also started to paint a clear picture on detection statistics of COMs around low- and high-mass protostellar systems. Yang et al. (2021) for example found methanol in only 56% of their Perseus sample, CH_3OCHO and CH_3CN in 32% and 40% of their sample. A similar or lower detection rate is also found for Orion (Bouvier et al. 2022; Hsu et al. 2022). Figure 1.4a shows examples of spectra for two low-mass Class 0 systems with varying levels of line richness. Moreover, van Gelder et al. (2022b) consider warm methanol mass in 184 low- and high-mass protostellar systems and find a $\sim$4 orders of magnitude spread in those values. The origin of this spread is an integral part of this thesis and has been suggested to be either due to optically thick dust or small scale source structures such as disks (see Section 1.4.2 for more discussion).

1.4.1.2 Ices

There is a wealth of information on simple species in ices of protostellar envelopes obtained with previous space telescopes such as *Infrared Space Observatory* (ISO)

and *Spitzer* as well as ground-based telescopes such as the ESO-Very Large Telescope. Species such as CO, CO_2, NH_3, H_2O and OCN^- have been commonly observed in these systems (Gerakines et al. 1999; Boogert et al. 2000; Keane et al. 2001; van Broekhuizen et al. 2004; Öberg et al. 2008; Pontoppidan et al. 2003, 2008; Bottinelli et al. 2010; Boogert et al. 2015). However, when it comes to COMs the only one firmly detected with the previous telescopes in ices is methanol (e.g., Dartois et al. 1999; Öberg et al. 2011). This is because of the low abundance of COMs and low sensitivity of telescopes prior to *JWST*. However, there were hints of some COM features in the protostellar ices from spectra of pre-*JWST* era. Some examples are suggestions of NH_2CHO, C_2H_5OH, and CH_3CHO in ices by Schutte et al. (1999) and Öberg et al. (2011).

JWST with its unprecedented sensitivity ($\sim$2 orders of magnitude higher than *Spitzer*) and higher spectral resolution in the critical 3-10 μm range has the ability to firmly detect many of the previously suggested COMs in ices. Yang et al. (2022) already identified features that could be attributed to C_2H_5OH and CH_3CHO in the protostellar envelope of RAS 15398-3359. Moreover, other ongoing works (Rocha et al. submitted; Chen et al. in preparation) are analyzing abundances of O-bearing COMs using the ENIIGMA fitting tool (Rocha et al. 2021) in the guaranteed time observation program JOYS+ (van Dishoeck et al. 2023). *Chapter 6* of this thesis presents the first tentative detection of a nitrogen-bearing COM (i.e., CH_3CN) in ices around two protostars. It is important to note that these advances have only become possible because of the new infrared laboratory spectra of O- and N-bearing COM features (e.g., Rachid et al. 2020, 2022; Terwisscha van Scheltinga et al. 2021; Slavicinska et al. 2023), summarized in the Leiden Ice Database for Astrochemistry (LIDA; Rocha et al. 2022).

1.4.2 Disk

The existence of disks around Class 0 protostars has been subject of debate for years. They were often assumed to be too small to have a considerable effect on the chemical composition and physical conditions of the system. This meant that many models were assumed to be one-dimensional. However, single-source studies started to detect disks in a handful of Class 0 and I sources with the SMA and Combined Array for Millimeter-wave Astronomy (Lee 2011; Tobin et al. 2012; Yen et al. 2013). At the same time with ALMA more Class 0 and I sources started to show disks (e.g., Murillo et al. 2013; Tobin et al. 2016a; Yen et al. 2017; Harsono et al. 2018; Maret et al. 2020). In particular, recent surveys find that disks of $\sim 10 - 100$ au seem to be more common than exception (Tobin et al. 2020). One of these is the VLA/ALMA Nascent Disk and Multiplicity (VANDAM) survey which observed all embedded protostars in Perseus and 328 systems in Orion at an angular resolution between $0.1''$ and $0.2''$. A few of the many outcomes of this survey are that disks in Class 0 objects are common, dust disk sizes show a decreasing trend from Class 0 to Class I sources, and dust substructures can form early in the disk lifetime (Segura-Cox et al. 2018; Tobin et al. 2020; Sheehan et al. 2020). Another recent survey with ALMA, Early Planet Formation in Embedded Disks (eDisk), searched for disks and substructures within them as signs of early

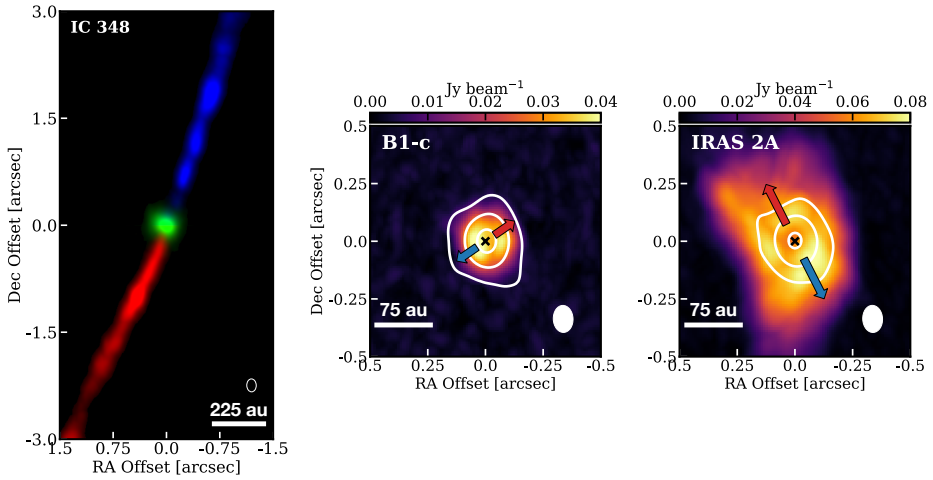

Figure 1.8: *Left panel:* High angular resolution observations of SiO (8-7; red and blue) peak intensity map and the continuum image (green) in IC348, *Middle and right panels:* CH_3OH ($2_{2,1,0} - 3_{1,2,0}$) peak intensity maps in B1-c and IRAS2A. The crosses show the peak of the continuum emission (data from ALMA program with PI: B. Tabone and project ID: 2021.1.01578.S).

planet formation in 12 Class 0 and 7 Class I disks with an angular resolution of $\sim 0.04''$ (around twice better than VANDAM). They also found that disks are common around embedded protostars, however, they did not find substructures to be common, only some disks show clear patterns.

Nevertheless, disks are expected to form early in the protostellar stage as a result of angular momentum conservation with the density concentrated around the midplane. The density structure of a disk in vertical hydrostatic equilibrium is (Shakura & Sunyaev 1973; Pringle 1981)

$$\rho = \rho_{\rm mid}(R) e^{\frac{-z^2}{2H(R)^2}} \tag{1.8}$$

where R and z are the radius and height in cylindrical coordinates. $H(R)$ is the disk scale height which is a function of radius and $\rho_{\rm mid}$ is the mid-plane density which is normally assumed as a power-law relation in radius decreasing at larger distances from the protostar. The disk midplane is usually a few orders of magnitude denser than the envelope surrounding it and hence it is also much colder that the inner envelope ($\sim 20 - 30$ K versus > 100 K) as the protostellar radiation cannot not reach deep within the disk. These characteristics also change the chemical reactions expected to take place in the disk mid plane and at the disk surface (e.g., Bosman et al. 2018; Calahan et al. 2022). An important effect of the disk on the physical conditions is the decrease in temperature behind where the disk is located due to *disk shadowing*. A consequence of disk shadowing was suggested to be a decrease in the amount of warm sublimated water in the inner regions of protostellar systems (e.g., Persson et al. 2012, 2016) to solve the 'missing

water problem' (e.g., Visser et al. 2013; van Dishoeck et al. 2021). *Chapters 7 and 8* of this thesis particularly focus on whether disks and optically thick dust can decrease the methanol emission using radiative transfer models as a solution to the large scatter observed in warm methanol mass across many systems (van Gelder et al. 2022b).

Disks in embedded systems are best traced by minor isotopologues of CO although other tracers can also be used for their detection and characterization such as SO and H_2CO (Tychoniec et al. 2021). It is not clear whether COMs trace the disk. For example, Fig. 1.8 presents high angular resolution observations of methanol in two Class 0 systems and without further analysis it is not possible to speculate on the origin of this emission. However, it is interesting to note the lower intensity of methanol emission on the continuum peak in B1-c and IRAS 2A which could be tracing the disk (see *Chapter 10* for more discussion).

Finally, an important recent conclusion of disk surveys has been that there is simply not enough mass in the later Class II disks for the observed planet masses. Thus, at least the cores of giant planets should start forming early, already in the Class 0 or I phase (Manara et al. 2018; Tychoniec et al. 2018). However, there is still a debate on whether most of planet formation can take place in the later stages, as originally thought, through the mass provided by for example streamers (e.g., Akiyama et al. 2019; Huang et al. 2021; Gupta et al. 2023). If planets form earlier than previously expected, more planet formation models are needed to assess how and if planets can form in these turbulent early phases (Cridland et al. 2022).

1.4.3 Jet and the outflow

Protostellar systems often have a high-velocity collimated jet and a lower-velocity wider outflow. In terms of diagnostics, jets are often traced by the SiO emission (left panel of Fig. 1.8), while the wide-angle outflow is traditionally traced by SO and CO (Rodriguez et al. 1980; Lada 1985; Schmid-Burgk et al. 1990; Guilloteau et al. 1992; Podio et al. 2021; Tychoniec et al. 2021). The origins of both high- and low-velocity components are under debate. Starting with the highly collimated component, there are two leading models that predict them based on magneto-centrifugal forces. These two models are the X-wind model (Shu et al. 2000) and the magneto-hydrodynamical (MHD) disk-wind model (Konigl & Pudritz 2000). Both predict extraction of angular momentum from the (inner) disk although at different magnitude and radii. The X-wind model predicts extraction of angular momentum on the order of $\sim$10 au km s^{-1} at radii of $\sim$0.05 au while the disk-wind model predicts extraction of angular momentum as large as $\sim 100 - 200$ au km s^{-1} from a range of radii in the disk, $\sim 1 - 40$ au (see review by Lee 2020).

The origin of the low-velocity component is also predicted by the two leading models described above. In the X-wind model the wide-angle wind is launched at the same radius as the jet ($\sim$0.05 au), while in the disk-wind model the low-velocity outflow is launched as shells from a set of radii in the disk (Ferreira et al. 2006). Another reason for the observed wide-angle outflows could be laterally produced bow-shocks by the central jet (e.g., Tafalla et al. 2017). Which of the above

mechanisms is indeed the underlying reason is still under debate (see reviews by Arce et al. 2007; Pascucci et al. 2023). The diagnostic of the disk-wind as opposed to lateral bow-shocks is that for disk winds an onion-like structure is expected where all shells in this structure launched at various radii are expected to rotate. Evidence for rotating shells in the outflow emission has only been observed in a handful of sources as it requires high angular resolution observations ($\sim$10 – 30 au; Bjerkeli et al. 2016; Hirota et al. 2017; Tabone et al. 2017; Louvet et al. 2018; de Valon et al. 2020). *Chapter 10* of this thesis presents observations of rotating shells in the outflow of L1448-mm and discovers evidence for an MHD disk wind.

Two important properties of a disk wind are its launching radius (r_0) and the gas magnetic lever arm parameter (λ_ϕ). The latter quantifies the specific angular momentum extracted by the magnetic field and given to the gas which can be used to estimate the mass ejection rate of the disk wind. The r_0 and λ_ϕ can be measured using Anderson's and Ferreira's relations as (Anderson et al. 2003; Ferreira et al. 2006)

$$2rv_\phi\Omega_0 = v_{\rm pol}^2 + 3\Omega_0^2 r_0^2, \tag{1.9}$$

and

$$rv_\phi = \Omega_0 r_0^2 \lambda_\phi, \tag{1.10}$$

respectively. In Equations (1.9) and (1.10), $v_{\rm pol}$ is the poloidal velocity, $\Omega_0 = \sqrt{\frac{GM_\star}{r_0^3}}$ is the Keplerian angular velocity at r_0, and v_ϕ is the azimuthal velocity. It is important to note that Eq. (1.9) neglects the gravity well and is only valid far enough from the central protostar. Moreover, it also assumes that $v_{\rm pol} >> v_\phi$.

Finally, an important role of the jet and the outflow is enabling accretion onto the protostar. The jet only enables accretion in the very inner regions ($\lesssim$ 1 au) by extracting angular momentum, while the disk wind is expected to extract angular momentum and enable accretion throughout the disk. Therefore, disk winds can be a solution to the 'angular momentum problem'. The standard model of disk evolution uses viscosity as the main mechanism of extraction of angular momentum (Lynden-Bell & Pringle 1974), while Tabone et al. (2022a,b) recently showed that disk winds could not only explain disk accretion and evolution but also the fast disk dispersal.

1.5 This thesis

1.5.1 A few questions under debate ...

This section summarizes a few main questions in the field of star and planet formation and their link to astrochemistry. It is important to note that this list is not exhaustive and is chosen to match the main concepts discussed in this thesis. From the chemistry side, a question that has been extensively debated is, *Do COMs form in the gas or in ices?* Sections 1.2.1 and 1.2.2 attempt to capture some of this debate. Another question related to the chemistry is, *Inheritance or*

reset? In other words, whether the molecules that are observed in later stages of star formation (e.g., Class II and Class III disks) are inherited from the earlier protostellar phase or whether they are destroyed during the evolution of the system and formed almost from scratch in the Class II disks.

An important question discussed in this thesis is *why some protostellar systems show COMs and some do not?* Although the reason for this could be chemical and thus absence of COMs in some systems, this thesis argues for variations in physical conditions and source structure across protostellar systems as the primary reason. Moreover, the *reason for strong or weak correlation among column densities of various COMs* is still under debate.

There are also many questions regarding the *formation and evolution of disks and later dispersal of the envelope and the disk.* As described briefly in Sect. 1.4.3, disk evolution and accretion can take place if angular momentum is extracted from the disk. The two current leading theories to explain disk evolution are viscously evolving disks (Lynden-Bell & Pringle 1974) or MHD wind-driven disks (Blandford & Payne 1982). Although a huge progress has already been made in this front, more observational evidence and modeling studies are needed to further understand disk evolution. For example, data from upcoming ALMA large programs will provide valuable measurements of disk masses and disk sizes that can be directly compared with the predictions of models to help decide which theory of disk accretion is more dominant.

1.5.2 Overview of each chapter

This thesis is divided into two parts, the first part focusing on the chemical complexity in protostellar systems sets the stage for the second part that connects the observed molecular emission to the physical conditions and structure of those systems. The data used here are either from ALMA or *JWST*. From the ALMA side this thesis provides one of the largest protostellar samples in which COMs are detected (*Chapters 2, 3, 5*), one of the highest angular resolution data sets of protostellar systems (*Chapter 10*), and one of the deepest ALMA Band 3 ($\sim$3 mm) data sets of a low-mass protostellar system (the famous IRAS 16293-2422; *Chapter 4*). The *JWST* data used here provide the first study on complex cyanides in ices with a tentative detection (*Chapter 6*). Finally, this thesis also provides radiative transfer models of protostellar systems with and without a disk in *Chapters 7-9* to consider the effect of source structure on COM emission and column density ratios in low- and high-mass protostellar systems. Overview and the main conclusions of each chapter are provided below:

Part I Origin of chemical complexity in protostellar systems

Chapter 2 High-sensitivity ALMA observations of B1-c and S68N Class 0 protostellar systems mainly at $\sim$1 mm were analyzed. This chapter focuses on observations of nitrogen-bearing COMs and presents detection of NH_2CHO (formamide), C_2H_5CN (ethyl cyanide), HNCO (isocyanic acid) and its minor isotopologues and deuterated isotopologues of CH_3CN (methyl cyanide) in the two systems. The column densities and excitation temperatures of these

molecules were determined using the CASSIS spectral analysis tool. The column density ratios with respect to CH_3OH and HNCO were compared with a handful of available systems in the literature and the need for larger source samples with COM detections was highlighted for robust conclusions on the formation pathways of COMs.

Chapter 3 This is the largest and potentially one of the most significant chapters of this thesis where several nitrogen-bearing COMs in addition to methanol (CH_3OH) and the *minor isotopologues* of the most abundant species were detected and analyzed for the first time in a statistically significant sample of sources from the ALMAGAL program. Previously, surveys on COMs with ALMA or other telescopes could only detect these relatively less abundant COMs, such as C_2H_5CN or NH_2CHO, in only $\sim$10 sources, while here those molecules are detected in $\gtrsim$30 sources. A temperature segregation between various molecules was observed that correlates with their sublimation temperatures suggesting an onion-like structure around the central protostar. The most important conclusion of this chapter is the finding of remarkably constant column density ratios among most N-bearing COMs across sources with a large (orders of magnitude) luminosity and mass range, which points to a common formation environment of these species, likely the ices in the prestellar phase. The small scatter in column density ratios was attributed to small variations in prestellar lifetimes or initial conditions of the cold clouds.

Chapter 4 The deepest ALMA Band 3 ($\sim$3 mm) observations of the low-mass protostellar system IRAS 16293-2422 B were analyzed to specifically search for large ($>$ 8 atoms) three-carbon COMs. The frequency range of the observations was chosen to cover i- and n-propanol and glycerol. Long-wavelength observations are thought to be better for finding larger molecules because of less line confusion, lower millimeter dust opacity and, favorable partition function. This chapter highlights the richness of the spectra and the difficulty to observe large COMs even at 3 mm for warm protostellar systems. More than 70 simple and complex molecules are searched for but the only three-carbon COM detected is acetone (CH_3COCH_3). Upper limits are provided for $\sim$40 molecules including glycerol and (i- and n-)propanol which are consistent with previous laboratory and observational works. However, tentative detections of the cyclic COM c-C_2H_4O (ethylene oxide) and the rarely observed prebiotic COM, $HOCH_2CN$ (glycolonitrile) are reported. This chapter also finds that the effect of dust optical depth on column densities and column density ratios is minimal at a location off-source for IRAS 16293-2422 B.

Chapter 5 This chapter analyzes the hot ($\gtrsim$300 K) to warm ($\sim$100 K) gas-phase column density ratios of CH_3CN, C_2H_3CN (vinyl cyanide), CH_3OH and HNCO for $\sim$40 protostellar systems and finds that molecules without oxygen systematically show a factor of $\sim$5 enhancement in their hot component. More oxygen- and nitrogen-bearing species are needed to confirm this result. Nevertheless, the enhancement was interpreted as a signature for destruction

of refractory organics in the hot inner regions of the protostellar systems and the top down formation of molecules without oxygen in the gas phase, hence implying higher carbon to oxygen and nitrogen to oxygen ratios in these inner regions. This chapter particularly highlights the potential of less abundant COMs in constraining the carbon, oxygen and nitrogen budget in protostellar systems in the future.

Chapter 6 This chapter uses the *JWST* data of unprecedented sensitivity for five protostellar systems in the Investigating Protostellar Accretion (IPA) program to search for complex cyanides in ices. The tentative detection of CH_3CN and C_2H_5CN in ices were reported for the first time around two sources. The ratios of nitrogen-bearing species considered in this chapter (OCN^-, CH_3CN, and C_2H_5CN) agree between ice and gas, while the ratios with respect to methanol in ices are around a factor of 5 larger than the gas-phase ratios measured in *Chapter 3*. This was attributed to physical effects, such as the difference in sublimation temperatures of complex cyanides and methanol (see *Chapter 9*). A tentative enhancement of OCN^-, CH_3CN, and C_2H_5CN in warm ices was reported, although they likely start forming in the cold dense cloud. This chapter marks an important step toward understanding COM chemistry but a larger sample of objects studied with MIRI and NIRSpec is required to examine the trends found here statistically.

Part II Effect of physical structure on molecular emission

Chapter 7 Roughly half of protostellar systems show COM emission at millimeter wavelengths and the other half at a similar evolutionary stage do not. The reason for this could be low abundance of COMs in some systems, or COMs could exist in all systems but are not seen in the gas. This chapter uses radiative transfer models with a constant gas-phase abundance of methanol (the simplest and most abundant COM) inside of its snowline to consider whether presence of disk and optically thick dust can reduce methanol emission even if methanol is abundant in low-mass protostellar systems. The conclusion is that models without a disk can only explain observations with strong methanol emission and over-produce the rest even when optically thick dust is included. Models with disks can explain bulk of the observations by decreasing the temperature (disk shadowing) and thus the amount of warm methanol. However, they can only explain systems with low methanol emission and high bolometric luminosity when high-millimeter opacity dust is included and thus continuum over-subtraction becomes effective. In other words, both a disk and optically thick dust in those disks are needed to explain the observations, and absence of millimeter methanol emission does not imply absence of methanol molecules.

Chapter 8 Here the same question as *Chapter 7* is considered but for high-mass protostellar systems. A similar model to that of *Chapter 7* was used but trimmed for high-mass systems. One of the major differences was inclusion of viscous heating in the disk mid-plane. The conclusion of this chapter is that disks and optically thick dust in high-mass protostellar systems are not

as effective as their low-mass counterparts in lowering the methanol emission to explain the large scatter found from ALMA surveys. The disk needs to be large ($\sim$2000 − 2500 au) with high-millimeter opacity dust grains to be effective in lowering the methanol emission. The models with disk and high-millimeter opacity dust can mostly explain the observations at lower luminosities ($\sim10^3$ $L_\odot$) but fail to explain the observations at higher luminosities ($\gtrsim10^4$ $L_\odot$). An alternative solution could be presence of hypercompact or ultracompact HII regions which were not considered in the original models.

Chapter 9 In *Chapter 3*, NH_2CHO/CH_3OH showed one of the largest scatters across many protostellar systems. Such a large scatter (or weak correlation) is usually interpreted as gas-phase formation of NH_2CHO. However, if disks can change the level of molecular emission as found in *Chapters 7 and 8*, they might also change the column density ratios in various systems. This chapter combines the radiative transfer models of *Chapters 7 and 8* with CASSIS modeling, just as done in observations, to investigate the effect of source structure on column density ratios of NH_2CHO/CH_3OH and CH_3CN/CH_3OH. There are two important conclusions from this work: 1. The observational scatter in NH_2CHO/CH_3OH can be partially explained by the difference in source structure and sublimation regions of methanol and formamide across protostellar systems, 2. The difference in sublimation regions of two molecules can produce a 'correction factor' as large as $\sim$10 that can decrease the measured gas-phase ratios compared with their true (i.e., ice) column density ratios.

Chapter 10 Here high angular resolution ALMA observations ($\sim0.1''$, i.e., $\sim$30 au) of the Class 0 protostar L1448-mm are used to resolve the outflow of L1448-mm and search for evidence of a disk wind. Various components of the system (disk, envelope and outflow) were resolved. Rotation signatures were observed in the outflowing emission of CH_3OH and $H^{13}CN$ with a kinematical structure consistent with theoretical predictions of MHD disk winds. This is the first detection of an MHD disk wind candidate in CH_3OH and $H^{13}CN$. The estimated mass loss rate is around four times the mass accretion rate and suggests that the wind can carry enough angular momentum to drive disk accretion.

1.5.3 Summary of the main findings

A combined summary of the major discoveries of this thesis are

- Gas-phase column density ratios of nitrogen-bearing COMs are remarkably constant (with a certain amount of scatter) across many protostellar systems spanning a large mass and luminosity range. This points to their bulk formation in similar physical environments, likely in the prestellar phase as ices followed by their sublimation close to the protostar (*Chapters 2 and 3*). However, effects of subsequent hot gas-phase chemistry for nitrogen-bearing COMs without an oxygen (e.g., CH_3CN) should not be ignored (*Chapter 5*).

- Chemical complexity does not only exist in the gas but can also exist in ISM ices which can be detected and quantified by *JWST* observations (*Chapter 6*), further confirming the conclusion made in the previous point. Detection of higher levels of chemical complexity in the gas (e.g., three-carbon species) is difficult for warm systems even at 3 mm observations (*Chapter 4*).

- Source structure combined with optically thick dust can affect molecular line emission (*Chapters 7 and 8*). This also leads to source structure changing the scatter in column density ratios (or correlation between column densities) if the molecules have different-enough sublimation temperatures. This difference in sublimation temperatures can also move the mean of the gas-phase column density ratios, affecting the interpretations of gas and ice comparison (*Chapter 9*).

- Hot to warm ratios of molecules without oxygen are enhanced by a factor of $\sim$5 compared to those with oxygen. This could be interpreted as destruction of refractory organics in the inner regions and increase of C/O and N/O. Less abundant COMs can help quantify the elemental ratios in protostellar systems as a means to connect the rich chemistry in these systems and that of the planets forming there (*Chapter 5*).

- Observations of disk winds are rare but high-angular resolution ALMA observations can be used to detect and quantify them and prove that disk winds can drive disk accretion. Moreover, CH_3OH and $H^{13}CN$ are suggested as two new disk-wind tracers (*Chapter 10*).

1.5.4 Forward look

An interesting future prospect of chemical studies in the early protostellar phase could be directly linking the chemistry observed in these systems (both simple and complex species) to the planetary compositions that likely start forming in this phase. In other words, comparing the elemental hydrogen, carbon, oxygen, nitrogen and sulfur budgets (or their ratios) in planetary atmospheres and the protostellar disks using chemical models, ALMA and *JWST* observations will provide valuable information on when and how planets form. This link will ultimately help us understand the origin of our Solar System and the life in the Universe.