



**HAL**  
open science

# HyGAL: Characterizing the Galactic Interstellar Medium with Observations of Hydrides and Other Small Molecules. I. Survey Description and a First Look Toward W3(OH), W3 IRS5, and NGC 7538 IRS1

A. M. Jacob, D. A. Neufeld, P. Schilke, H. Wiesemeyer, W. -J. Kim, S. Bialy, M. Busch, D. Elia, E. Falgarone, M. Gerin, et al.

## ► To cite this version:

A. M. Jacob, D. A. Neufeld, P. Schilke, H. Wiesemeyer, W. -J. Kim, et al.. HyGAL: Characterizing the Galactic Interstellar Medium with Observations of Hydrides and Other Small Molecules. I. Survey Description and a First Look Toward W3(OH), W3 IRS5, and NGC 7538 IRS1. *The Astrophysical Journal*, 2022, 930, 10.3847/1538-4357/ac5409 . insu-03717118

**HAL Id: insu-03717118**

**<https://insu.hal.science/insu-03717118v1>**

Submitted on 9 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



# HyGAL: Characterizing the Galactic Interstellar Medium with Observations of Hydrides and Other Small Molecules. I. Survey Description and a First Look Toward W3(OH), W3 IRS5, and NGC 7538 IRS1

A. M. Jacob<sup>1,2</sup>, D. A. Neufeld<sup>1</sup>, P. Schilke<sup>3</sup>, H. Wiesemeyer<sup>2</sup>, W.-J. Kim<sup>3</sup>, S. Bialy<sup>4</sup>, M. Busch<sup>1</sup>, D. Elia<sup>5</sup>, E. Falgarone<sup>6</sup>, M. Gerin<sup>7</sup>, B. Godard<sup>8</sup>, R. Higgins<sup>3</sup>, P. Hennebelle<sup>9</sup>, N. Indriolo<sup>10</sup>, D. C. Lis<sup>11</sup>, K. M. Menten<sup>2</sup>, A. Sanchez-Monge<sup>3</sup>, T. Möller<sup>3</sup>, V. Ossenkopf-Okada<sup>3</sup>, M. R. Rugel<sup>2</sup>, D. Seifried<sup>3</sup>, P. Sonnentrucker<sup>12</sup>, S. Walch<sup>3</sup>, M. G. Wolfire<sup>4</sup>, F. Wyrowski<sup>2</sup>, and V. Valdivia<sup>13,14</sup>

<sup>1</sup> William H. Miller III Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA; [ajacob51@jhu.edu](mailto:ajacob51@jhu.edu)

<sup>2</sup> Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany

<sup>3</sup> I. Physikalisches Institut, Universität zu Köln, Zùlpicher Str. 77, D-50937 Köln, Germany

<sup>4</sup> Department of Astronomy, University of Maryland, College Park, MD 20742-2421, USA

<sup>5</sup> INAF—Istituto di Astrofisica e Planetologia Spaziali, via Fosso del Cavaliere 100, I-00133, Roma, Italy

<sup>6</sup> Laboratoire de Physique de l'ENS, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, 24 rue Lhomond, F-75005 Paris, France

<sup>7</sup> LERMA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universités, F-75014 Paris, France

<sup>8</sup> Observatoire de Paris, École normale supérieure, Université PSL, Sorbonne Université, CNRS, LERMA, F-75005, Paris, France

<sup>9</sup> AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, F-91191, Gif-sur-Yvette, France

<sup>10</sup> AURA for the European Space Agency (ESA), Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

<sup>11</sup> Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

<sup>12</sup> European Space Agency (ESA), ESA Office, Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

<sup>13</sup> Laboratoire AIM, Paris-Saclay, CEA/IRFU/SAP-CNRS-Université Paris Diderot, F-91191 Gif-sur-Yvette Cedex, France

<sup>14</sup> Department of Physics, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan

Received 2021 December 17; revised 2022 February 8; accepted 2022 February 9; published 2022 May 12

## Abstract

The HyGAL Stratospheric Observatory for Infrared Astronomy legacy program surveys six hydride molecules—ArH<sup>+</sup>, OH<sup>+</sup>, H<sub>2</sub>O<sup>+</sup>, SH, OH, and CH—and two atomic constituents—C<sup>+</sup> and O—within the diffuse interstellar medium (ISM) by means of absorption-line spectroscopy toward 25 bright Galactic background continuum sources. This detailed spectroscopic study is designed to exploit the unique value of specific hydrides as tracers and probes of different phases of the ISM, as demonstrated by recent studies with the Herschel Space Observatory. The observations performed under the HyGAL program will allow us to address several questions related to the life cycle of molecular material in the ISM and the physical processes that impact the phase transition from atomic to molecular gas, such as: (1) What is the distribution function of the H<sub>2</sub> fraction in the ISM? (2) How does the ionization rate due to low-energy cosmic rays vary within the Galaxy? (3) What is the nature of interstellar turbulence (e.g., typical shear or shock velocities), and what mechanisms lead to its dissipation? In this overview, we discuss the observing strategy, the synergies with ancillary and archival observations of other small molecules, and the data reduction and analysis schemes we adopted; and we present the first results obtained toward three of the survey targets, W3(OH), W3 IRS5, and NGC 7538 IRS1. Robust measurements of the column densities of these hydrides—obtained through widespread observations of absorption lines—help address the questions raised, and there is a very timely synergy between these observations and the development of theoretical models, particularly pertaining to the formation of H<sub>2</sub> within the turbulent ISM. The provision of enhanced HyGAL data products will therefore serve as a legacy for future ISM studies.

*Unified Astronomy Thesaurus concepts:* [Interstellar abundances \(832\)](#); [Astrochemistry \(75\)](#); [Diffuse interstellar bands \(379\)](#); [Molecular spectroscopy \(2095\)](#); [Interstellar medium \(847\)](#)

## 1. Introduction

Diffuse clouds, being at the crossroad between the low-density, mainly neutral and atomic gas phase, and the dense and fully molecular cloud phase where stars are born, not only trace this crucial transition but play an active role in the life cycle of gas in the interstellar medium (ISM). They provide the raw materials from which dense molecular clouds form. Being continuously replenished by feedback from late stages of star formation and stellar evolution, diffuse clouds form a reservoir of heavy elements that is subsequently responsible for initiating

the growth of chemical complexity in the ISM. Therefore, the study of diffuse clouds is crucial in advancing our interpretation of a variety of physical and chemical processes of broad applicability in astrophysics. These processes include those governing the transition from H- to H<sub>2</sub>-dominated gas (Federman et al. 1979; Sternberg et al. 2014; Bialy & Sternberg 2016), an essential ingredient in any complete theory of star formation and galaxy evolution.

Being poorly shielded against photodissociating interstellar radiation fields, the low-density ( $n_{\text{H}} \sim 10^2\text{--}10^3 \text{ cm}^{-3}$ ) gas comprising diffuse clouds was originally predicted to be devoid of molecules (Eddington 1926). However, observations at ultraviolet, visible, and radio wavelengths have revealed an astounding inventory of molecules (Goss 1968; Whiteoak & Gardner 1970; Lucas & Liszt 1996; Snow &

McCall 2006; McGuire 2022). The surprisingly rich chemistry present in diffuse clouds can be attributed to exothermic reactions driven by cosmic-ray ionization and large fluctuations in the thermal pressure of the media. In more recent years, the molecular richness of diffuse clouds has been further explored by means of absorption spectroscopy at far-infrared (FIR) and submillimeter wavelengths against the background continuum emission of bright star-forming regions in both the Milky Way and external galaxies (see Gerin et al. 2016, for a complete overview). These observations—largely performed with the Herschel Space Observatory (HSO; Pilbratt et al. 2010), the Atacama Pathfinder Experiment (APEX; Güsten et al. 2006), and the Stratospheric Observatory for Infrared Astronomy (SOFIA; Young et al. 2012)—have not only extended absorption-line studies over Galactic scales, but have also renewed interest in the study of hydrides, defined here as molecules containing a single heavy element atom with one or more hydrogen atoms.

Historically, hydrides have the distinction of being the first gas-phase molecules to be observed in the ISM, through the successful detection and identification of CH and CH<sup>+</sup> alongside CN at optical wavelengths near 4300 Å (Dunham 1937; Swings & Rosenfeld 1937; McKellar 1940). Thereafter, the ground-state transitions of OH near 18 cm became the first molecular signature to be detected in the ISM at radio wavelengths (Weinreb et al. 1963). Moreover, being among the first species to form in gas with both atomic and molecular hydrogen through relatively simple chemical pathways, hydrides play an important role in driving the chemistry of the ISM. Furthermore, because the FIR rotational transitions of most hydride species possess high rates of spontaneous radiative decay and in turn high critical densities, hydride molecules are typically in the ground rotational state under diffuse cloud conditions ( $n_{\text{H}} \sim 10\text{--}500 \text{ cm}^{-3}$ ); their measured absorption-line strengths therefore yield robust and accurate estimates of the molecular column densities.

Observations performed over the past decade have allowed the ground-state rotational transitions of several interstellar hydrides to be observed for the first time at high spectral resolution and provided a wealth of new information and unique clues to our understanding of the physics of the diffuse ISM. These include sensitive tracers of atomic gas and the cosmic-ray ionization rate like ArH<sup>+</sup> (Schilke et al. 2014; Bialy et al. 2019; Jacob et al. 2020b), OH<sup>+</sup>, H<sub>2</sub>O<sup>+</sup>, and H<sub>3</sub>O<sup>+</sup> (Gerin et al. 2010a; Gupta et al. 2010; Neufeld et al. 2010a; Ossenkopf et al. 2010; Wyrowski et al. 2010; Indriolo et al. 2012, 2015); molecular CO-dark H<sub>2</sub> gas tracers such as HF (Neufeld et al. 2010b; Sonnentrucker et al. 2015), CH (Gerin et al. 2010b; Wiesemeyer et al. 2018; Jacob et al. 2019), and OH (Wiesemeyer et al. 2016); and tracers of the dissipation of turbulence or the turbulent transport at the interface of different phases such as CH<sup>+</sup> and SH<sup>+</sup> (Falgaroni et al. 2010; Menten et al. 2011; Godard et al. 2014). While the careful analysis of hydride abundances obtained from these observations have demonstrated the potential of absorption-line studies of interstellar hydrides to probe the physical and chemical conditions in diffuse clouds, previous observations were limited to a fairly small set of sight lines almost entirely within the inner Galaxy.

With the end of the Herschel mission, SOFIA is the only observatory that can provide access that is almost unhindered

by atmospheric absorption in the terahertz frequency range, where the rotational transitions of most interstellar hydrides lie. Moreover, the German REceiver for Astronomy at Terahertz Frequencies (GREAT)<sup>15</sup> instrument on SOFIA enables high-resolution spectroscopy for the first time at frequencies that were inaccessible to the Heterodyne Instrument for the Far-infrared (HIFI; de Graauw et al. 2010) on Herschel, permitting absorption-line studies of the OH, SH, CH, and O transitions targeted in the HyGAL program.

Taking advantage of this unique capability, the HyGAL legacy survey aims to characterize the diffuse Galactic ISM, significantly expanding the set of observations available to probe the diffuse ISM throughout the Galaxy with observations of six key hydrides (ArH<sup>+</sup>, p-H<sub>2</sub>O<sup>+</sup>, OH<sup>+</sup>, SH, OH, and CH) along with two atomic constituents of the diffuse ISM (C<sup>+</sup> and O). The acquisition and analysis of these data are also very timely, as theoretical simulations of the turbulent multiphase ISM are starting to include chemistry, thereby allowing observations to be compared to specific model predictions for the first time. This paper presents an overview of the HyGAL program and the analysis of the first data collected from SOFIA Cycle 8 observations toward three of the targeted sources, W3(OH), W3 IRS5, and NGC 7538 IRS1. Section 2 details the main scientific goals of the HyGAL program, with the observations, use of archival and complementary data, while the data reduction strategies are described in Section 3. We present and discuss the results in Sections 4 and 5, and summarize our findings in Section 6.

## 2. HyGAL: Key Science Goals

The processes by which molecular clouds—the sites of star formation, traced by CO—form out of the atomic medium and the timescales on which molecular cloud formation occurs are crucial missing links in theories of star formation (Mac Low & Klessen 2004; McKee & Ostriker 2007; Kennicutt & Evans 2012; Padoan et al. 2014; Krumholz & Federrath 2019; Rosen et al. 2020). Therefore, investigations of the diffuse atomic and molecular media are critical for our understanding of the formation and evolution of molecular clouds. While this topic is highly complex, the observations planned under the HyGAL legacy project will address several related questions about the ISM and star formation, including:

1. What is the distribution of the molecular fraction in various phases (diffuse atomic gas, diffuse molecular gas, and dense molecular gas) of the cold and warm neutral media of the ISM?
2. How does the ionization rate due to low-energy cosmic rays, an important source of heating in the ISM and responsible for regulating star formation, vary within the Galaxy?
3. What is the nature of interstellar turbulence (e.g., the typical shear or shock velocities) and what mechanisms lead to its dissipation?

The following subsections detail the potential use of interstellar hydrides in addressing these key questions.

<sup>15</sup> GREAT is a development by the MPI für Radioastronomie and the KOSMA/Universität zu Köln, in cooperation with the DLR Institut für Optische Sensorsysteme.

### 2.1. Probing the Transition from Diffuse Atomic to Molecular Gas

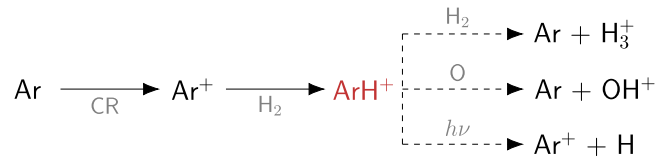
The molecular fraction,  $f(\text{H}_2)$ , is a fundamental quantity that may be used to describe the transition between atomic and molecular gas within the cold neutral medium. It may be expressed either as a ratio of volume densities,  $f^n(\text{H}_2)$ , or column densities,  $f^N(\text{H}_2)$ :

$$f^n(\text{H}_2) = \frac{2n(\text{H}_2)}{n(\text{H}) + 2n(\text{H}_2)}, \quad (1)$$

$$f^N(\text{H}_2) = \frac{2N(\text{H}_2)}{N(\text{H}) + 2N(\text{H}_2)}, \quad (2)$$

where  $n(X)$  and  $N(X)$  refer to the volume density (in units of  $\text{cm}^{-3}$ ) and column density (in units of  $\text{cm}^{-2}$ ) for  $X = \text{H}, \text{H}_2$ , respectively.

Despite being the most abundant molecular species in the ISM, molecular hydrogen has been very difficult to observe directly, except through its ground electronic transitions near 1100 Å at far-ultraviolet wavelengths in absorption (Savage et al. 1977; Rachford et al. 2002, 2009; Pan et al. 2005; Burgh et al. 2007; Shull et al. 2021)—observations of which are restricted to bright nearby hot stars and toward bright QSOs—and through its emission in infrared (IR) rovibrational transitions (Burton et al. 1992; Habart et al. 2011; Santangelo et al. 2014). Rovibrational emission from  $\text{H}_2$  is typically excited in shock-heated or UV-irradiated regions, so UV absorption lines observed toward nearby hot stars provide the only direct probe of the cold  $\text{H}_2$  in the ISM. Because it is homonuclear and symmetric,  $\text{H}_2$  does not possess a permanent dipole moment and radiates only through weak quadrupole transitions of its excited rotational and rovibrational states at wavelengths below  $\sim 28 \mu\text{m}$  ( $h\nu/k_B \sim 500$  K). Furthermore, with high-lying and widely spaced energy levels, the rotational excitation of  $\text{H}_2$  becomes important only for gas temperatures  $T \geq 80$  K, making infrared  $\text{H}_2$  emissions from the cold ISM unobservable in practice. The challenges involved in observing molecular hydrogen are overcome by using chemically associated species and dust as surrogates for  $\text{H}_2$ . While the integrated intensity of low-lying rotational transitions of CO scaled by an empirical conversion factor  $X(\text{CO})$  has been the most commonly used proxy for measuring the total molecular reservoir, there is mounting evidence for the presence of excess molecular material not traced by CO, dubbed as CO-dark  $\text{H}_2$  gas (Grenier et al. 2005), which may constitute up to 50% of the total mass of the molecular gas along any given sight line (Pineda et al. 2013). This excess molecular gas points to the inability of CO to act as a proxy for  $\text{H}_2$ , in regions of low dust column densities (or visual extinction) where CO readily undergoes photodissociation. In recent years, through observations of their radio and FIR rotational transitions, hydrides such as CH (Gerin et al. 2010a; Wiesemeyer et al. 2018; Jacob et al. 2019), OH (Allen et al. 2015; Engelke & Allen 2018; Rugel et al. 2018; Busch et al. 2019, 2021), and HF (Neufeld et al. 2010b; Sonnentrucker et al. 2015) have been established as important tracers of CO-dark  $\text{H}_2$  gas in addition to the [C II] 158  $\mu\text{m}$  line (Langer et al. 2014) and the  $J = 1 - 0$  rotational transition of  $\text{HCO}^+$  (Lucas & Liszt 1996; Gerin et al. 2010b). We note that the CH– $\text{H}_2$  relationship was first established by Federman (1982) and later by Sheffer et al. (2008) and then Weselak (2019) using optical observations of the  $\text{A}^2\Delta - \text{X}^2\Pi$  system of CH at 4300 Å and the (2–0), (3–0), and (4–0)



**Figure 1.** Schematic representation of the argon chemistry in diffuse clouds. Solid arrows display chemical pathways leading to the formation of  $\text{ArH}^+$  initiated by cosmic rays (CR), while the dashed arrows indicate competing reactions that reduce the abundance of this cation.

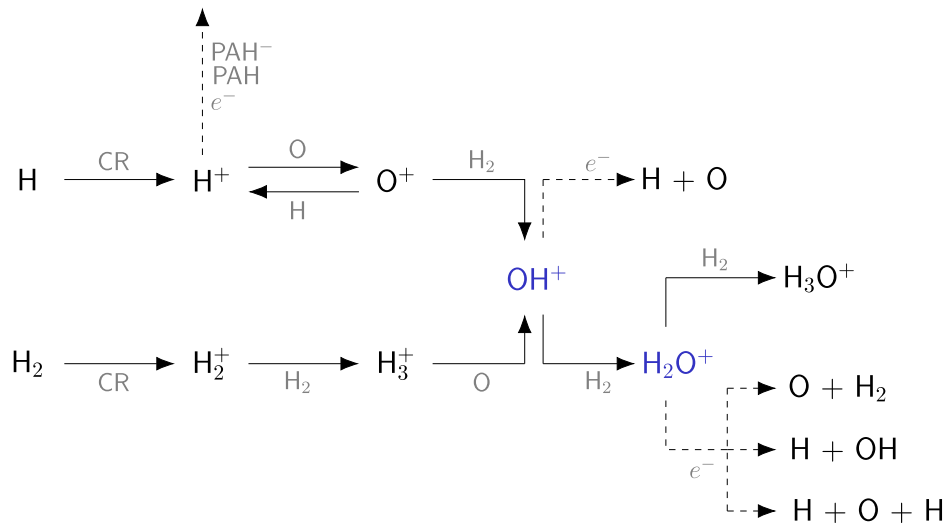
bands of the Lyman B–X transitions of  $\text{H}_2$  toward stars located in the local diffuse ISM ( $10^{19} < N(\text{H}_2) < 10^{21} \text{ cm}^{-2}$ ). The validity of this relationship between CH and  $\text{H}_2$  over Galactic scales is poorly known, having been determined only using a sample of nearby stars that are bright at visible wavelengths, and might be dependent on variations in the elemental abundance of carbon across the Galaxy. Nevertheless, a linear CH– $\text{H}_2$  relationship with a constant slope has been widely adopted in previous studies, an approach we follow here.

The neutral hydrides that are observed in the HyGAL project—CH, OH, and SH—are known to show peak abundances in gas with a relatively large fraction of hydrogen in molecular form ( $f^n(\text{H}_2) > 0.1$ ), tracing diffuse molecular and translucent clouds (as per the classification of the neutral medium by Snow & McCall 2006). In contrast, the hydride ions targeted in this study— $\text{ArH}^+$ ,  $\text{p-H}_2\text{O}^+$ , and  $\text{OH}^+$ —are rapidly destroyed by reactions with  $\text{H}_2$  and are therefore most abundant in material with molecular fractions of a few percent (see Figures 1 and 2 and 1 of Bialy et al. 2019). The  $n(\text{OH}^+)/n(\text{H}_2\text{O}^+)$  ratio is a particularly valuable probe of the  $\text{H}_2$  fraction (Neufeld et al. 2010a) because the reaction between  $\text{H}_2$  and  $\text{OH}^+$ , which forms  $\text{H}_2\text{O}^+$ , competes with the dissociative recombination of  $\text{OH}^+$  (Figure 2); thus, the  $n(\text{OH}^+)/n(\text{H}_2\text{O}^+)$  ratio is larger in regions of small  $f^n(\text{H}_2)$  and smaller in regions of large  $f^n(\text{H}_2)$ .

An analysis of  $\text{OH}^+$  and  $\text{H}_2\text{O}^+$  absorption features observed using Herschel along sight lines toward bright Galactic continuum sources (Indriolo et al. 2015) showed these molecules to trace  $\text{H}_2$  fractions of a few percent ( $f^N(\text{H}_2) = 10^{-2} - 10^{-1}$ ). However, having been determined using ratios of column densities rather than volume density (i.e.,  $N(\text{OH}^+)/N(\text{H}_2\text{O}^+)$  rather than  $n(\text{OH}^+)/n(\text{H}_2\text{O}^+)$ ), the derived molecular fractions only represent averages across the entire region where the  $\text{OH}^+$  and  $\text{H}_2\text{O}^+$  absorption features are produced. Therefore, for any given sight line, these  $f^N(\text{H}_2)$  values may differ substantially from those derived using measurements of  $n(\text{H}_2)$  and  $n(\text{H})$ . The argonium ion  $\text{ArH}^+$ , on the other hand, traces material of even lower molecular fraction ( $f^N(\text{H}_2) = 10^{-4} - 10^{-2}$ ) and high cosmic-ray ionization rates, which paradoxically makes it a better tracer of almost purely atomic hydrogen gas than H I itself, as discussed in Schilke et al. (2014) and Neufeld & Wolfire (2016). Taken together, the observations of both the ionized and neutral hydride species carried out in the HyGAL project trace different cloud layers across any given sight line and reveal the transition from diffuse atomic ( $n_{\text{H}} = 10 - 100 \text{ cm}^{-3}$ ;  $f^n(\text{H}_2) < 0.1$ ) to diffuse molecular ( $n_{\text{H}} = 100 - 500 \text{ cm}^{-3}$ ;  $f^n(\text{H}_2) > 0.1$ ) and translucent clouds ( $n_{\text{H}} = 500 - 5000 \text{ cm}^{-3}$ ;  $f^n(\text{H}_2) \sim 0.8$ ; Snow & McCall 2006).

### 2.2. Probing the Cosmic-ray Ionization Rate in the Diffuse ISM

In addition to providing valuable probes of the molecular fraction in the Galactic ISM, the observed abundances of the hydride ions  $\text{OH}^+$ ,  $\text{H}_2\text{O}^+$ , and  $\text{ArH}^+$  allow us to estimate the



**Figure 2.** Schematic representation of the oxygen chemistry in diffuse clouds. Solid arrows display chemical pathways leading to the formation of oxygen hydride cations— $\text{OH}^+$  and  $\text{H}_2\text{O}^+$ —while the dashed arrows indicate competing reactions that reduce the abundance of these cations.

cosmic-ray ionization rates. Since the first ionization potentials of oxygen (13.62 eV) and argon (15.76 eV) are higher than that of hydrogen (13.59 eV), both atoms remain primarily in their neutral states in the cold neutral medium as they are well-shielded from extreme-ultraviolet radiation capable of ionizing them. Therefore, the formation of oxygen- and argon-bearing hydride ions such as  $\text{OH}^+$ ,  $\text{H}_2\text{O}^+$ , and  $\text{ArH}^+$  is driven by cosmic-ray ionization (Gerin et al. 2010a; Neufeld et al. 2010a; Neufeld & Wolfire 2016). The formation of  $\text{ArH}^+$  ensues via the cosmic-ray ionization of Ar to produce  $\text{Ar}^+$ , which quickly reacts with  $\text{H}_2$  to form  $\text{ArH}^+$  (see Figure 1), while oxygen chemistry in diffuse clouds, initiated by the cosmic-ray ionization of H or  $\text{H}_2$ , follows a sequence of ion-neutral reactions resulting in the formation of  $\text{OH}^+$  and  $\text{H}_2\text{O}^+$ , as shown in Figure 2.

Carrying out a detailed analysis of data available from Herschel on  $\text{OH}^+$ ,  $\text{H}_2\text{O}^+$ , and  $\text{ArH}^+$ , Neufeld & Wolfire (2017) obtained an estimate of the cosmic-ray ionization rate within diffuse atomic clouds with an average value  $\zeta_p(\text{H}) = (2.2 \pm 0.3) \times 10^{-16} \text{ s}^{-1}$ . A similar value is obtained by Jacob et al. (2020b) using  $\text{ArH}^+$  data collected using the APEX 12 m telescope. While this value is in excellent agreement with entirely independent estimates of  $\zeta_p(\text{H})$  obtained for diffuse molecular clouds (with  $f(\text{H}_2) > 0.1$ ) using observations of  $\text{H}_3^+$  (McCall et al. 2003; Indriolo et al. 2007; Indriolo & McCall 2012), it is 1 order of magnitude larger than the typical cosmic-ray ionization rates inferred for dense molecular clouds (Caselli et al. 1998; Bialy et al. 2022) and expectations based on direct measurements of cosmic rays obtained with the Voyager I spacecraft (Cummings et al. 2016). This suggests that cosmic rays are excluded from dense molecular clouds (e.g., Padovani et al. 2009, 2020) and that the cosmic-ray flux in the vicinity of the solar system may be atypical. Furthermore, results obtained from existing  $\text{H}_3^+$  data by Neufeld & Wolfire (2017) suggest at a marginal level of significance that the cosmic-ray ionization rates in diffuse molecular clouds decrease with cloud extinction for  $A_V \geq 0.5$ .

Observations of the key cosmic-ray ionization rate tracers,  $\text{OH}^+$ ,  $\text{H}_2\text{O}^+$ , and  $\text{ArH}^+$ , to be performed in the HyGAL project—in combination with ancillary observations (discussed further in Section 3.3) of tracers of higher total gas column and

molecular fraction such as  $\text{HCO}^+$  and  $\text{CO}$ —will not only more than double the sample of Galactic sight lines toward which the cosmic-ray ionization rate has been determined, but will also provide better statistics to understand the nature of the cosmic-ray ionization rates in different cloud environments.

### 2.3. Probing the Chemistry of Warm Phases of the ISM

While the chemistry of  $\text{OH}^+$ ,  $\text{H}_2\text{O}^+$ , and  $\text{ArH}^+$  discussed in Sections 2.1, and 2.2 proceed via exothermic reactions that are rapid at temperatures typical of the diffuse ISM ( $\leq 100 \text{ K}$ ; Snow & McCall 2006), that of other hydrides and in particular those containing sulfur (like  $\text{SH}$ ,  $\text{SH}^+$ ,  $\text{H}_2\text{S}$ ) and  $\text{CH}^+$ , are distinctive in that their formation only takes place by means of endothermic reactions. The latter may be driven by elevated gas temperatures or ion-neutral drift in shocks associated with turbulent dissipation regions; or via turbulent transport at the interface between the cold and warm neutral media (Myers et al. 2015; Valdivia et al. 2017; Moseley et al. 2021).

Observationally,  $\text{SH}$  was first detected using the GREAT instrument by Neufeld et al. (2012) in absorption toward the star-forming region W49N and subsequently detected toward a handful of other star-forming regions (Neufeld et al. 2015). The measured abundances of  $\text{SH}$  in the diffuse ISM are found to be many orders of magnitude larger than those predicted for quiescent, cold clouds by standard models of photodissociation regions (PDRs), suggesting that some fraction of the volume is occupied by a warmer phase heated by shocks or other processes that lead to the dissipation of interstellar turbulence (Godard et al. 2009, 2014; Neufeld et al. 2015). Similarly, Walch et al. (2015) have found using the SILCC<sup>16</sup> simulations that the shock heating of a large fraction of gas that neighbors molecular clouds lies above the thermal equilibrium curve thereby pushing this material into the thermally unstable regime. While such modeling efforts facilitate the interpretation of the chemistry of sulfur-bearing hydrides, the abundance estimates are based on a limited number of detections in diffuse clouds.

Observations of  $\text{SH}$  performed in the HyGAL program will provide high-quality spectra toward individual targets probing

<sup>16</sup> SILCC- Simulating the Life Cycle of molecular Clouds.

**Table 1**  
Spectroscopic Properties of the Studied Species and Transitions

Species	Transition		Frequency (GHz)	$A_{u,l}$ ( $s^{-1}$ )	$E_u$ (K)	GREAT Channel	HPBW (")	$\eta_{MB}^a$
	$J' - J''$	$F' - F''$						
ArH <sup>+</sup>	1 – 0	...	617.5252(2)	0.0045	29.63	4G1	44.7	0.61
p-H <sub>2</sub> O <sup>+</sup>	3/2 – 1/2	...	604.6841(8)	0.0013	29.20	4G1	44.7	0.61
$N_{K_a K_c} = 1_{0,0} - 1_{0,1}$	3/2 – 3/2	...	607.2258(2)	0.0062	29.20			
OH <sup>+</sup>	2 – 1	5/2 – 3/2	971.8038(15) <sup>b</sup>	0.0182	46.64	4G2	28.8	0.50
( $N = 1-0$ )		3/2 – 1/2	971.8053(15)	0.0152				
		3/2 – 3/2	971.91920(10)	0.0030				
SH	5/2 – 3/2	2 <sup>+</sup> – 2 <sup>-</sup>	1382.9040(1)	0.0005	66.40	4G3	19.3	0.60
$^2\Pi_{3/2}, N = 2$		3 <sup>+</sup> – 2 <sup>-</sup>	1382.9086(1)	0.0047				
		2 <sup>+</sup> – 1 <sup>-</sup>	1382.9152(1)	0.0042				
		2 <sup>-</sup> – 2 <sup>+</sup>	1383.2350(1)	0.0005				
		3 <sup>-</sup> – 2 <sup>+</sup>	1383.2397(1)	0.0047				
		2 <sup>-</sup> – 1 <sup>+</sup>	1383.2462(1) <sup>b</sup>	0.0042				
OH	5/2 – 3/2	2 <sup>-</sup> – 2 <sup>+</sup>	2514.2987(9)	0.0137	120.75	4G4	11.2	0.52
$^2\Pi_{3/2}, N = 2 - 1$		3 <sup>-</sup> – 2 <sup>+</sup>	2514.3167(9) <sup>b</sup>	0.1368				
		2 <sup>-</sup> – 1 <sup>+</sup>	2514.3532(9)	0.1231				
CH	3/2 – 1/2	1 <sup>-</sup> – 1 <sup>+</sup>	2006.74886(6)	0.0111	96.31	LFA	13.5	0.66
$^2\Pi_{3/2}, N = 2 - 1$		1 <sup>-</sup> – 0 <sup>+</sup>	2006.76258(6)	0.0223				
		2 <sup>-</sup> – 1 <sup>+</sup>	2006.79906(6) <sup>b</sup>	0.0335				
C II $^2P_{3/2} - ^2P_{1/2}$	2 – 1	...	1900.5369(13)	$2.32 \times 10^{-6}$	91.21	LFA	14.1	0.66
O I $^3P_1 - ^3P_2$	2 – 1	...	4744.7775(1)	$8.91 \times 10^{-5}$	227.76	HFA	6.3	0.65

**Notes.** The spectroscopic data are taken from the Cologne Database for Molecular Spectroscopy (CDMS; Müller et al. 2005). The H<sub>2</sub>O<sup>+</sup> frequencies were refined considering astronomical observations (see Appendix A of Müller et al. 2016) for which the upper level energies are given with respect to the ground state of p-H<sub>2</sub>O<sup>+</sup> ( $N_{K_a K_c} = 1_{0,1}$ ) and the CH frequencies are taken from Truppe et al. (2014). For the rest frequencies, the numbers in parentheses give the uncertainty in the last listed digit.

<sup>a</sup> The main-beam efficiencies for each channel (and pixel where applicable) are determined for each flight series, listed here are the typical values.

<sup>b</sup> Indicates the hyperfine-structure transition that was used to set the velocity scale in the analysis.

different sight lines. In combination with ancillary data on other small molecules containing sulfur, such as CS and H<sub>2</sub>S, they will provide key information about the overall sulfur chemistry in the diffuse and translucent clouds. Interpreted in the context of chemical models of shocks and turbulent dissipation regions, the observed abundances of sulfur-bearing molecules further promises to provide important constraints on the nature of interstellar turbulence.

### 3. Observations

#### 3.1. Source Selection

The goals of the HyGAL project discussed in Section 2 are met by performing absorption-line spectroscopy of foreground interstellar gas along the sight lines to bright background continuum sources. The most useful sight lines intersect as many spiral arms as possible, including those in the outer Galaxy (i.e., sources located outside the solar circle at Galactocentric distances,  $R_{GAL} > 8.15$  kpc), to probe the structure of the ISM. Our main selection criterion, however, is driven by sensitivity. Since the signal-to-noise ratio obtained for an absorption feature is proportional to the continuum flux for a given column density, the most favorable sources are those possessing the largest continuum fluxes within the far-IR wavelength range (119–494  $\mu$ m) encompassing the targeted line transitions (see Table 1). Suitable background continuum sources were identified using the source catalog of the Hi-GAL Herschel key-guaranteed time project (Molinari et al. 2010), which provides continuum fluxes at 70, 160, 250, 350, and 500  $\mu$ m. The set of sources targeted in this survey consists of 25 submillimeter continuum sources in the Galactic plane with

160  $\mu$ m (1.9 THz) continuum fluxes in excess of 2000 Jy for sources in the inner Galaxy and 1000 Jy for those in the outer Galaxy, as estimated from the Hi-GAL source catalog (Elia et al. 2021). The positions of the background sources were checked against known millimeter-wave sources, H II regions and masers using the SIMBAD<sup>17</sup> astronomical database (Wenger et al. 2000). In addition, the HyGAL source list also includes a few sources for which partial spectroscopic data are available from Herschel/HIFI, most often collected as a part of the PRISMAS<sup>18</sup> guaranteed time key program, but where the complete set of species targeted in HyGAL has not been observed. Therefore, the completion of this legacy program will roughly double the number of sight lines along which the hydrides targeted in this survey have been observed; equally importantly, it will extend observations toward regions in the outer Galaxy that were previously unexplored. The HyGAL sources are situated over a wide range of Galactic longitude, lying at estimated distances from the Sun ranging from 1.3 to 13 kpc.

Table 2 provides an overview of the properties of the HyGAL source sample, and Figure 3 shows their distribution in the Galactic disk. Furthermore, a brief description of each source and its line-of-sight properties is given in Appendix A.

#### 3.2. SOFIA Observations

The science goals discussed in Section 2 are achieved by exploiting the complete capabilities of the high-resolution

<sup>17</sup> See <http://simbad.u-strasbg.fr/simbad/>.

<sup>18</sup> PRobing InterStellar Molecules with Absorption line Studies (PRISMAS; Gerin et al. 2010a).

**Table 2**  
HyGAL Source Parameters

#	Source Designation	R.A. (hh:mm:ss)	Decl. (dd:mm:ss)	Gal. Long. (deg)	Gal. Lat. (deg)	$v_{\text{LSR}}$ (km s <sup>-1</sup> )	$d$ [Ref] (kpc)	$R_{\text{GAL}}$ (kpc)
a	HGAL284.015–00.86	10:20:16.1	–58:03:55.0	284.016	–0.857	9.0	5.7 [1]	9.0
b	HGAL285.26–00.05	10:31:29.5	–58:02:19.5	285.263	–0.051	3.4	4.3 [2]	8.2
c	G291.579–00.431	11:15:05.7	–61:09:40.8	291.579	–0.431	13.6	8.0 [3]	9.3
d	IRAS 12326-6245	12:35:35.9	–63:02:29.0	301.138	–0.225	–39.3	4.6 [4]	7.2
e	G327.3–00.60	15:53:05.0	–54:35:24.0	327.304	–0.551	–46.9	3.1 [5]	6.2
f	G328.307+0.423	15:54:07.2	–53:11:40.0	328.309	+0.429	–93.6	5.8 [6]	4.6
g	IRAS 16060–5146	16:09:52.4	–51:54:58.5	330.953	–0.182	–91.2	5.3 [7]	4.5
h	IRAS 16164–5046	16:20:11.9	–50:53:17.0	332.827	–0.551	–57.3	3.6 [8]	5.4
i	IRAS 16352–4721	16:38:50.6	–47:28:04.0	337.404	–0.403	–41.4	12.3 [4]	5.1
j	IRAS 16547–4247	16:58:17.2	–42:52:08.9	343.126	–0.063	–30.6	2.7 [8]	5.8
k	NGC 6334 I	17:20:53.4	–35:47:01.5	351.417	+0.645	–7.4	1.3 [9]	7.0
l	G357.558–00.321	17:40:57.2	–31:10:59.3	357.557	–0.321	5.3	9.0–11.8 [10]	1.0–3.6
m	HGAL0.55–0.85	17:50:14.5	–28:54:30.7	0.546	–0.851	16.7	7.7–9.2 [11]	0.4–1.0
n	G09.62+0.19	18:06:14.9	–20:31:37.0	9.620	+0.194	4.3	5.2 [12]	3.3
o	G10.47 + 0.03	18:08:38.4	–19:51:52.0	10.472	+0.026	67.6	8.6 [13]	1.6
p	G19.61–0.23	18:27:38.0	–11:56:39.5	19.608	–0.234	40.8	12.6 [14]	4.7
q	G29.96–0.02	18:46:03.7	–02:39:21.2	29.954	–0.016	97.2	6.7 [6]	4.5
r	G31.41+0.31	18:47:34.1	–01:12:49.0	31.411	+0.307	98.2	4.9 [15]	5.0
s	W43 MM1	18:47:47.0	–01:54:28.0	30.817	–0.057	97.8	5.5 [15]	5.0
t	G32.80+0.19	18:50:30.6	–00:02:00.0	32.796	+0.191	14.6	13.0 [16]	7.4
u	G45.07+0.13	19:13:22.0	+10:50:54.0	45.071	+0.133	59.2	4.3 [17]	6.2
v	DR21	20:39:01.6	+42:19:37.9	81.681	0.537	–4.0	1.5 [18]	7.4
w	<b>NGC 7538 IRS1</b>	<b>23:13:45.3</b>	<b>+61:28:11.7</b>	<b>111.542</b>	<b>0.777</b>	<b>–59.0</b>	<b>2.6 [19]</b>	<b>9.8</b>
x	<b>W3 IRS5</b>	<b>02:25:40.5</b>	<b>+62:05:51.0</b>	<b>133.715</b>	<b>1.215</b>	<b>–39.0</b>	<b>2.3 [20]</b>	<b>9.9</b>
y	<b>W3(OH)</b>	<b>02:27:04.1</b>	<b>+61:52:22.1</b>	<b>133.948</b>	<b>1.064</b>	<b>–48.0</b>	<b>2.0 [20, 21]</b>	<b>9.6</b>

**Notes.** Those entries highlighted in bold text refer to the sources discussed in this work the remaining sources have either only recently been observed or are yet to be observed.

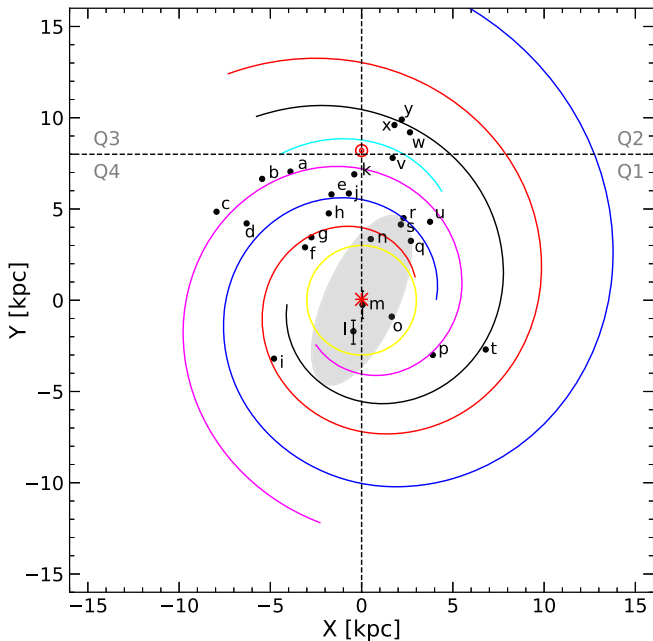
**References.** For the heliocentric distances: [1] Urquhart et al. (2014b); [2] Caswell & Vaile (1995); [3] Lee et al. (2012); [4] Green & McClure-Griffiths (2011), [5] Wienen et al. (2015); [6] Urquhart et al. (2018); [7] Moisés et al. (2011); [8] Giannetti et al. (2014); [9] Immer et al. (2013); [10] Frail et al. (1996) (Lacking distance estimates, for G357.558–00.321 we adopt a distance equivalent to the nearby supernova remnant G357.7–0.1 (MSH 17–39)); [11] Walsh et al. (1998); [12] Sanna et al. (2009); [13] Sanna et al. (2014); [14] Urquhart et al. (2014a); [15] Zhang et al. (2014) (For G31.41+0.31, following Winkel et al. (2017) we adopt a distance which is the average between two sources with similar LSR velocities, G31.28+0.06 and G31.58+0.07 for which accurate distances were measured by Zhang et al. (2014) and consistent with Reid et al. (2017).); [16] Kolpak et al. (2003); [17] Stead & Hoare (2010); [18] Rygl et al. (2012); [19] Moscadelli et al. (2009); [20] Navarete et al. (2019); [21] Hachisuka et al. (2006). These distances were also cross-checked with those recently calculated by Mège et al. (2021).

heterodyne instrument, GREAT on board SOFIA, which provides six frequency bands of which up to five can be used simultaneously. The observations were carried out over several flights as a part of the observatory’s Cycle 8, and the ongoing Cycle 9 campaigns, under the open time project 08\_0038. The receiver configuration consists of the 4GREAT module (Durán et al. 2021) whose different frequency channels are tuned to observe five key hydrides (ArH<sup>+</sup>, H<sub>2</sub>O<sup>+</sup>, OH<sup>+</sup>, SH, and OH), along with the High-frequency Array (HFA) of upGREAT (Risacher et al. 2018), tuned to the 63  $\mu\text{m}$  fine-structure line of atomic oxygen. Additional configurations employ the Low-frequency Array (LFA) to observe CH or the 158  $\mu\text{m}$  [C II] line. The spectroscopic parameters of the different transitions studied and receiver configurations used are presented in Table 1.

The spectra are taken in the double-beam switch mode, chopping at a frequency of 2.5 Hz with a chop throw between 180'' and 300'' and a chop angle between  $-10^\circ$  and  $70^\circ$  (East of North) to account for both atmospheric and instrumental fluctuations that may arise. The receiver is connected to a modified version of the MPIfR Fast Fourier Transform Spectrometer described by Klein et al. (2012). This backend provides a 4 GHz bandwidth per frequency band, per spatial position (equivalent to velocity widths of  $\sim 1945, 1235, 868, 478, 615, \text{ and } 252 \text{ km s}^{-1}$  for the 4G1, 4G2, 4G3, 4G4, LFA,

and HFA receivers, respectively) over 16,384 channels. This implies a native velocity resolution of 0.118, 0.075, 0.053, 0.029, 0.037, and 0.015 km s<sup>-1</sup> for the 4G1, 4G2, 4G3, 4G4, LFA, and HFA receivers, respectively. The high resolution thereby achieved makes it possible to separate the spiral arms from the envelopes of the hot cores.

The 4G1 channel is tuned to observe the 607.225 GHz  $N_{K_a, K_c} = 1_{1,0} - 1_{0,1}$  line of p-H<sub>2</sub>O<sup>+</sup> in the lower sideband (LSB) and the 617.525 GHz  $J = 1 - 0$  transition of ArH<sup>+</sup> in the upper sideband (USB). Previous observations of the 607.225 GHz p-H<sub>2</sub>O<sup>+</sup> line absorption spectra (Schilke et al. 2010; Jacob et al. 2020b) clearly show blended emission features at the systemic velocity of the star-forming regions studied: these arise from the H<sup>13</sup>CO<sup>+</sup> ( $J = 7 - 6$ ) and CH<sub>3</sub>OH  $J_k = 12_2 - 11_1$ E lines at 607.1747 (Lattanzi et al. 2007) and 607.2159 GHz (Belov et al. 1995), respectively. The line of sight also contains the D<sub>2</sub>O ( $J_{K_a, K_c} = 1_{1,1} - 0_{0,0}$ ) transition at 607.349 GHz (Matsushima et al. 2001), although this transition has so far only been observed toward the solar type protostar IRAS16293-2422 (Butner et al. 2007; Vastel et al. 2010) and is unlikely to result in significant contamination. Therefore, in order to decompose contributions from both sidebands and disentangle any additional contamination present in the bandpass, two additional setups are used with offsets in the LO frequency corresponding to Doppler velocities



**Figure 3.** Distribution of the sight lines observed under the HyGAL program (black circles), as viewed from the north Galactic pole. The underlying trace of the Milky Way spiral arm pattern follows the log-periodic model presented in Reid et al. (2019); the Galactic center (red asterisk) is at (0, 0) kpc and the Sun (red Sun symbol) is at (0, 8.15) kpc; 3 kpc ring, yellow; Norma Outer arm, red; Scutum-Centaurus arm, blue; Sagittarius-Carina arm, purple; Local arm, cyan; Perseus arm, black; the “long” bar is indicated by a shaded ellipse after Wegg et al. (2015). The source names are designated using alphabetic letters (as given in Table 2) for clarity, five of which are located in the outer Galaxy. The sources labeled as l and m in this plot are depicted using bars due to uncertainties in their kinematic distances. In this paper, we discuss sources labeled w, x, and y, respectively.

of  $\pm 15 \text{ km s}^{-1}$ . A  $15 \text{ km s}^{-1}$  shift in the signal band corresponds to a  $30 \text{ km s}^{-1}$  relative shift in the image band. This observing strategy therefore results in a net image band offset of  $60 \text{ km s}^{-1}$  among the different setups, which is sufficient to separate features arising in the signal band from those in the image band. Furthermore, for the LFA observations only data collected from the central pixels of both polarizations are used because the continuum flux is much weaker or almost undetectable in the outer pixels which are separated from the central pixel by  $\sim 31.8''$ .

The data are processed using the KOSMA kalibrate program (Guan et al. 2012) and then converted to main-beam temperature scales,  $T_{\text{MB}}$ , assuming a forward efficiency,  $\eta_{\text{F}}$ , of 0.97 and main-beam efficiencies,  $\eta_{\text{MB}}$ , listed in Table 1. The calibrated spectra are subsequently processed using the GILDAS-CLASS software<sup>19</sup> and low order ( $\leq 2$ ) polynomial baselines are removed. The spectra are further resampled to a uniform channel width of  $0.5 \text{ km s}^{-1}$  (except for the OH<sup>+</sup> line toward NGC 7538 IRS1, for which the spectra are displayed with a velocity resolution of  $1.5 \text{ km s}^{-1}$  for clarity).

As of 2021 September, 39.4% ( $\approx 32 \text{ hr}$ ) of the total planned observations for the HyGAL Legacy program have been performed. Table 3 summarizes the current status of the observations toward each source. The post-processed data products delivered to the SOFIA Science Center by the

GREAT team will be publicly accessible via the NASA/IPAC InfraRed Science Archive<sup>20</sup> under Plan ID 08\_0038.

### 3.3. Ancillary Data Sets

In this section, we briefly summarize the ancillary data collected using various ground-based facilities, which have proven to be useful in interpreting previous observations of the FIR and submillimeter transitions of hydrides. Since many of the molecular ions targeted under the HyGAL Legacy program originate in diffuse atomic clouds where the H<sub>2</sub> fraction is small, absorption-line measurements of the 21 cm H I line have proven essential for their interpretation and the determination of their abundances within such clouds (see for example the use of such data as demonstrated in Winkel et al. 2017). Therefore, in order to exploit the diagnostic powers of these species and provide essential information for our modeling efforts, we have carried out observations of the H I 21 cm line using the Karl G. Jansky Very Large Array (JVLA) for the northern sources (Project ID: 20A-519, 21A-287; PI: M. R. Rugel) and archival Australia Telescope Compact Array (ATCA) observations for the southern sources, respectively. Simultaneously with the JVLA H I observations we observed all four hyperfine-structure transitions of the OH ground state near 18 cm. These OH observations, together with the 2.5 THz rotational transitions of OH observed with SOFIA, will provide a unique and valuable data set for this species. The reliable column densities determined from the SOFIA observations will aid in determining the excitation conditions of the former, which are usually not in thermal equilibrium.

For the interpretation of SH, observations of other sulfur-bearing species such as CS, H<sub>2</sub>S, and SO obtained with the Instituto de Radioastronomía Milimétrica (IRAM) 30 m telescope have proven very valuable (Neufeld et al. 2015). Moreover the 3 mm wavelength band covering these species will also cover transitions of other molecular gas tracers, like HCO<sup>+</sup> and C<sub>2</sub>H. Therefore, the SOFIA HyGAL observations will be analyzed alongside ancillary 2 mm and 3 mm wavelength observations carried out with the IRAM 30 m telescope (for the northern HyGAL sources) targeting primarily diatomic and triatomic species, like SO, CS, C<sub>2</sub>H, HCO<sup>+</sup>, HOC<sup>+</sup>, HCN, HNC, and H<sub>2</sub>S (Project ID: 003-20; PI: D. A. Neufeld). For the southern sources, we use archival data collected as a part of the ALMA Three-millimeter Observations of Massive Star-forming regions (ATOMS; Liu et al. 2020) survey (Project ID: 2019.1.00685.S; PI: Tie Liu). Detailed descriptions for each of the supplementary data sets discussed above and the results obtained will be presented in forthcoming papers.

Absorption-line studies of CO, in combination with Herschel archival observations of [C I] and [C II] observations obtained as a part of the SOFIA HyGAL project, will allow us to better understand the C<sup>+</sup>/C/CO (H I-to-H<sub>2</sub>) transition and the carbon budget (fraction of gas-phase carbon) in diffuse and translucent clouds. To reliably estimate CO column densities, we have utilized the excellent sensitivity provided by the Northern Extended Millimeter Array (NOEMA) to observe the  $J = 1 - 0$  lines of CO and its isotopologues; <sup>13</sup>CO, C<sup>17</sup>O, and C<sup>18</sup>O (Project IDs: S21AQ, W21AE; PI: W.-J., Kim). Interferometric observations filters out extended emission, allowing the isolation of absorption features toward the background point source in a clean way. As discussed by Liszt & Lucas (1998),

<sup>19</sup> Software package developed by IRAM, see <https://www.iram.fr/IRAMFR/GILDAS/> for more information regarding GILDAS packages.

<sup>20</sup> <https://irsa.ipac.caltech.edu/applications/sofia/>



**Table 3**  
Status of HyGAL SOFIA Survey as of 2021 September

Source	Species							
	ArH <sup>+</sup>	p-H <sub>2</sub> O <sup>+</sup>	OH <sup>+</sup>	SH	OH	CH	C II	O I
W3(OH)	...	–(3)	x(3)	x	x	x(6)	x	x
W3 IRS5	–(1)	–(3)	x(3)	x	x	x(6)	x(8)	x
NGC 7538 IRS1	...	...	x	x	x	x	x	x
DR21	...	...	...	...	...	x	x(8)	x
IRAS 16060–5146	x'(2)	x'	x'	x'	x'(5)	x'(7)	x'	x'(5)
IRAS 16164–5046	x'(2)	x'	x'	x'	x'(5)	–(7)	...	x'(5)
NGC 6334 I	x'	x'(3)	x'(3)	x'	x'	x'	x'	x
G10.47 + 0.03	–(2)	...	...	...	–(5)	–(2)	x'	x'(5)
G29.96–0.02	x'	x'(3)	x'(3)	x'(4)	x'	x'	x'	x'
G32.80+0.19	x'	x'	x'	x'	x'	x'	x'	x'
G45.07+0.13	x'	x'	x'	x'	x'	x'	x'	x'

**Note.** The columns demarcated using the x symbol denote observations collected as a part of SOFIA’s Cycle 8 northern deployment from Köln-Bonn, Germany in 2021 February–March, while those marked with an x' correspond to data collected during the Cycle 9 southern deployment from Papeete, Tahiti in 2021 July–August. Previous detections: (1) unpublished data observed under the Water In Star-forming regions with Herschel (WISH) survey using Herschel/HIFI (Observation IDs: 1342201530 and 1342201531); (2) Jacob et al. (2020b); (3) Indriolo et al. (2015); (4) Neufeld et al. (2015); (5) Wiesemeyer et al. (2016); (6) Wiesemeyer et al. (2018); (7) Jacob et al. (2019); (8) Gerin et al. (2015).

the combination of absorption and emission measurements provide the necessary information to probe the CO excitation and the CO column densities. The value of CO absorption-line measurements in probing CO in diffuse gas has previously been demonstrated by Liszt & Lucas (1998, 2004); Liszt et al. (2019).

#### 4. First Results Toward W3(OH), W3 IRS5, and NGC 7538 IRS1

This section presents a first impression of the quality and analysis of the data collected as a part of the HyGAL survey, through examples toward three of the targeted sources in our sample—W3(OH), W3 IRS5, and NGC 7538 IRS1.

##### 4.1. Spectra and Line-of-sight Properties

Located in the outer Galaxy, W3(OH), W3 IRS5, and NGC 7538 IRS1 were observed as a part of SOFIA’s Cycle 8 northern campaign from Köln-Bonn, Germany. The data were collected during five flights on 2021 February 9, 10, 17, 18, and 19 with flight ids: 695, 696, 698, 699, and 700, respectively. The average leg time (including overheads) was 148 min per source for both the LFA+HFA and 4GREAT+HFA configurations (see Section 3 for a complete description of the observational setups used) with an average system temperature of 3370, 5540, 2750, 4240, and 13500 K for the LFA, HFA, 4G2, 4G3, and 4G4 receivers, respectively. At the time when these observations were carried out, the 4G1 channel was not functional and the ArH<sup>+</sup>  $J=1-0$  and p-H<sub>2</sub>O<sup>+</sup>  $N_{K_a, K_c} = 1_{1,0} - 1_{0,1}$  lines were therefore not observed. Data products for all the other targeted lines were obtained for these three sources. Toward W3 IRS5, the ArH<sup>+</sup>  $J=1-0$  and p-H<sub>2</sub>O<sup>+</sup>  $N_{K_a, K_c} = 1_{1,0} - 1_{0,1}$  lines had been collected as a part of the guaranteed time key program WISH (van Dishoeck et al. 2011; observation IDs: 1342201530 and 1342201531) and were retrieved from the Herschel science archive.<sup>21</sup> Similarly, we present p-H<sub>2</sub>O<sup>+</sup> data toward W3(OH) collected as a part of Herschel open time programs with observation IDs: 1342268579-81 and published in Indriolo et al. (2015). The

resulting spectra of all observed species toward each of these sources are displayed in Figures 4–6.

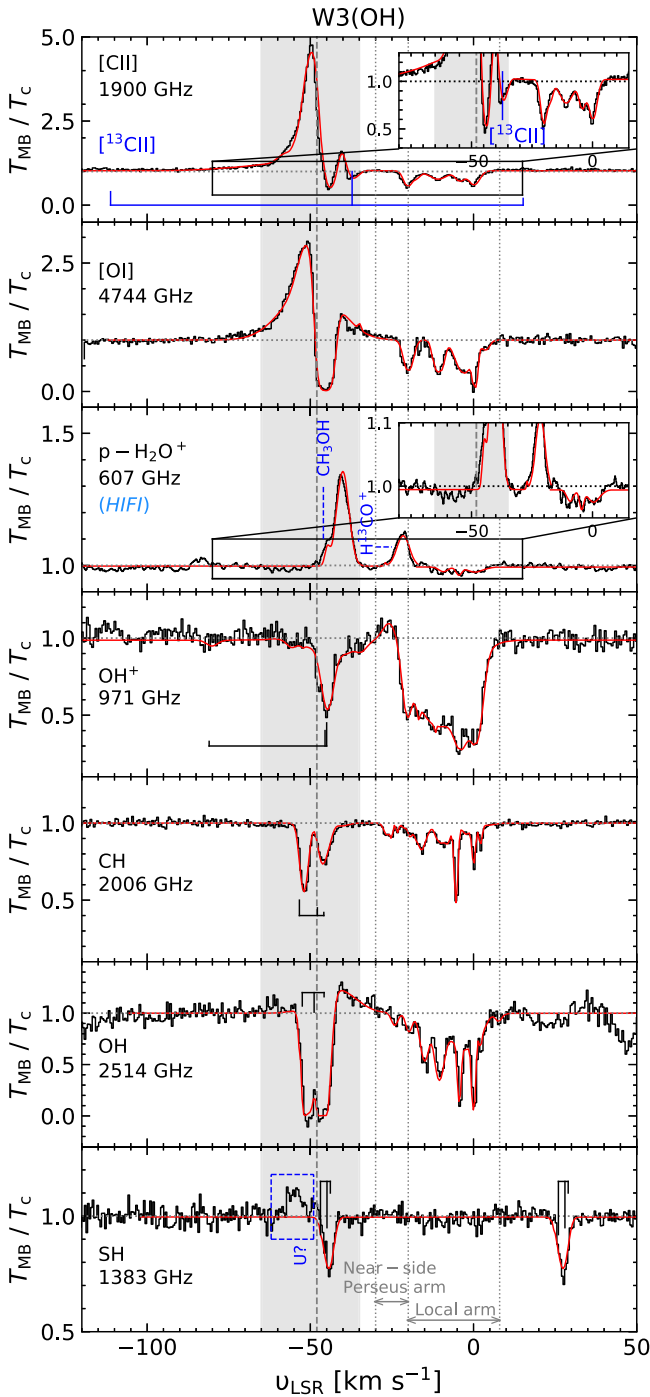
The following paragraphs describe the individual line profiles and line-of-sight features toward W3(OH), W3 IRS5, and NGC 7538 IRS1, while the properties and expected line-of-sight features (as determined from observations of other molecular species) for the remaining sources are given in Appendix appendix. Table 4 lists the single-sideband (SSB) continuum levels ( $T_c$ ) on main-beam temperature scales toward this subsample of sources at each observed frequency, along with the rms noise in the corresponding spectra.

The spectra observed along the line of sight toward hot cores often show emission from a plethora of molecules, including complex organic species (see, e.g., Belloche et al. 2013). Displaying a rich line spectrum at millimeter and submillimeter wavelengths, the emission from these species, many of which remain unidentified are often referred to as “weeds” as they contaminate the absorption profiles. In order to gauge the true depth of the absorption features, we identify and model contributions from the contaminating species. The contamination observed in each spectral line profile, if any, toward each individual source is discussed in more detail in the following sections.

##### 4.1.1. W3 IRS5 and W3(OH)

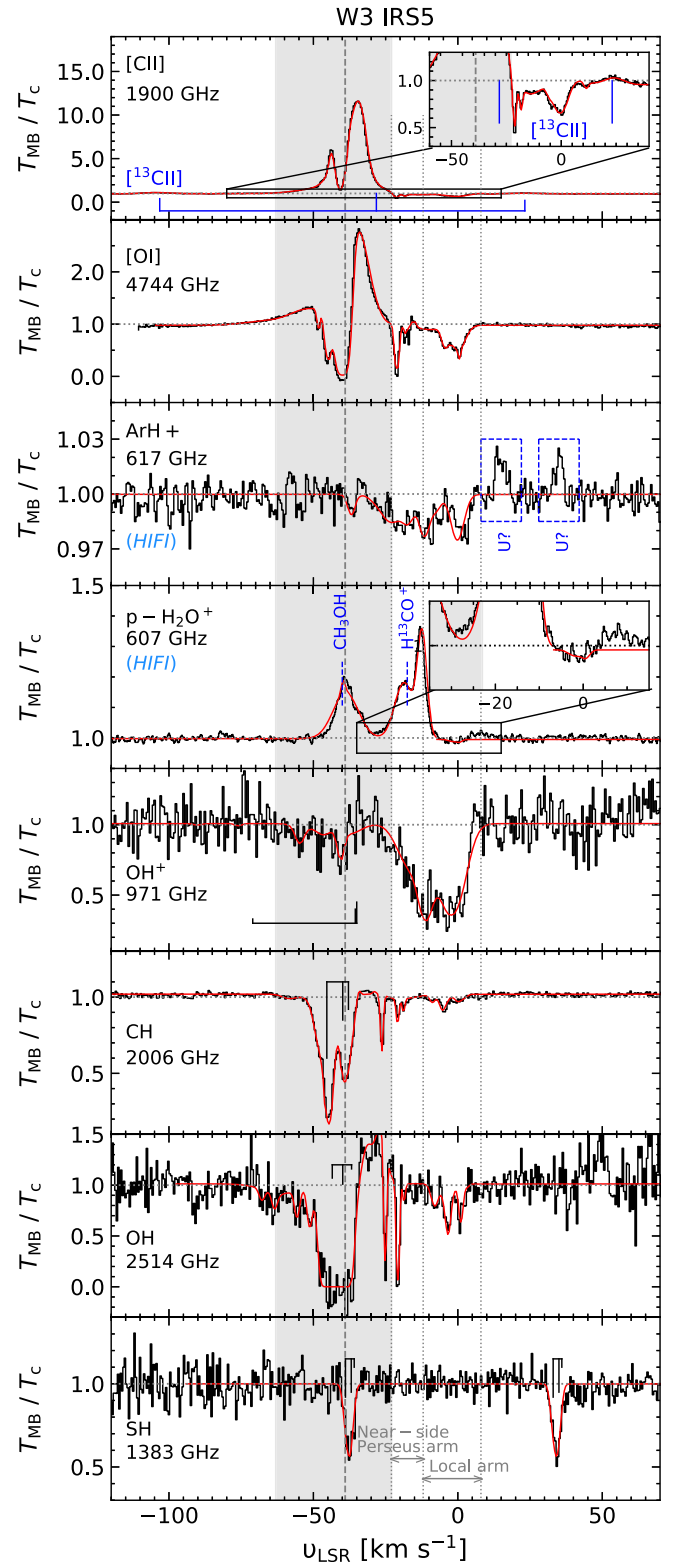
The W3 molecular cloud complex is located in the second quadrant of the Galaxy in the Perseus arm and consists of two main high-mass star-forming subregions, W3(OH) and W3 Main (for an overview see, e.g., Megeath et al. 2008), with a projected separation of  $\sim 10$  pc at the distance of the background sources ( $\approx 16.6$  apart on sky). The archetypical ultracompact H II region W3(OH) (e.g., Qin et al. 2016) is located at a distance of  $\sim 2$  kpc (Hachisuka et al. 2006; Xu et al. 2006) and is thought to harbor a massive young star, while W3 Main contains several luminous infrared sources, of which the W3 IRS5 protocluster is the brightest (Wang et al. 2013). Moreover, the W3 IRS5 cluster system has a well known double IR source at the center of an embedded cluster of a few hundred low mass stars (Megeath et al. 1996). The velocity distribution toward W3(OH) shows two major components, of

<sup>21</sup> <https://archives.esac.esa.int/hsa/whsa/>



**Figure 4.** From top to bottom: spectra in  $T_{\text{MB}}$  scales normalized with respect to the continuum temperature,  $T_c$ , showing transitions of  $\text{C}^+$ , O,  $\text{p-H}_2\text{O}^+$  (observed with Herschel/HIFI and taken from Indriolo et al. 2015),  $\text{OH}^+$ , CH, OH, and SH, respectively, toward W3(OH), in black, with the XCLASS fits in red. The inset in the top panel zooms-in on the  $\text{C}^+$  absorption features over a range of  $v_{\text{LSR}}$  between  $-80$  and  $15 \text{ km s}^{-1}$ . The vertical gray dashed line and gray shaded regions mark the systemic velocity and the typical velocity dispersion of the source. The relative intensities of the hyperfine-structure components of the  $\text{OH}^+$ , CH, and OH transitions are shown in black and  $^{13}\text{C}^+$  in blue above/below their respective spectra. Contaminating emission features that are identified are labeled in blue and their contributions modeled and removed in all subsequent analysis, while those that remain unidentified are marked (U?).

which one is attributed to the target itself and its envelope at velocities between  $-70$  and  $-35 \text{ km s}^{-1}$ , and the other one to the local spiral arm between  $-20$  and  $7 \text{ km s}^{-1}$ . Despite having



**Figure 5.** Same as Figure 4 but toward W3 IRS5 with the addition of the  $\text{ArH}^+ J=1-0$  spectrum observed using Herschel/HIFI.

line-of-sight velocity components largely similar to those of W3(OH), the sight line toward W3 IRS5 with a systemic cloud velocity of  $-39 \text{ km s}^{-1}$  shows relatively weaker absorption at velocities corresponding to the local spiral arm. This is particularly seen in the comparison of the CH and OH spectral line profiles toward both sources. While the deepest absorption

**Table 4**  
SSB Continuum Levels and rms Noise of the Spectra, and Column Density Sensitivity

Frequency (GHz)		Source		
		W3(OH)	W3 IRS5	NGC 7538 IRS
971	$T_c$ [K]	1.53	1.24	1.10
	$T_{\text{rms}}$ [K]	0.04	0.08	0.14
	$(dN(\text{OH}^+)/dv)_{\text{rms}}$	$2.0 \times 10^{10}$	$4.8 \times 10^{10}$	$9.3 \times 10^{10}$
1383	$T_c$	2.81	2.55	1.96
	rms	0.06	0.14	0.07
	$(dN(\text{SH})/dv)_{\text{rms}}$	$1.6 \times 10^{11}$	$4.0 \times 10^{11}$	$0.6 \times 10^{11}$
1900	$T_c$	5.20	5.57	3.14
	rms	0.07	0.06	0.06
	$(dN(\text{C}^+)/dv)_{\text{rms}}$	$1.0 \times 10^{15}$	$0.7 \times 10^{15}$	$1.2 \times 10^{15}$
2006	$T_c$	5.30	5.48	3.26
	rms	0.04	0.04	0.06
	$(dN(\text{CH})/dv)_{\text{rms}}$	$3.3 \times 10^{10}$	$3.2 \times 10^{10}$	$8.1 \times 10^{10}$
2514	$T_c$	6.33	7.85	4.42
	rms	0.24	0.73	0.27
	$(dN(\text{OH})/dv)_{\text{rms}}$	$5.8 \times 10^{10}$	$14.0 \times 10^{10}$	$9.3 \times 10^{10}$
4744	$T_c$	1.71	6.71	3.01
	rms	0.03	0.04	0.03
	$(dN(\text{O})/dv)_{\text{rms}}$	$6.3 \times 10^{14}$	$2.1 \times 10^{14}$	$3.6 \times 10^{14}$

**Note.** The rms noise level on the  $T_{\text{MB}}$  scale is quoted for a spectral resolution of  $0.5 \text{ km s}^{-1}$ .  $(dN/dv)_{\text{rms}}$  represents the column density corresponding to  $T_{\text{rms}}$  in units of  $[\text{cm}^{-2} / (\text{km s}^{-1})]$ . For those species with hyperfine-structure splitting,  $(dN/dv)_{\text{rms}}$  is computed using the spectroscopic parameters of the strongest hyperfine-structure transition.

features in the foreground material for CH and OH are observed at velocities between  $-7$  and  $+5 \text{ km s}^{-1}$  for W3(OH), those for W3 IRS5 lie between  $-32$  and  $-12 \text{ km s}^{-1}$ , which is absorption attributed to the near side of the Perseus arm. The spatial separations for the foreground clouds (gas which is not locally associated with the source) traced by both sources range from  $10 \text{ pc}$  for clouds located near the sources to  $<1 \text{ pc}$  further away from the background sources at positive velocities. Therefore, the HyGAL project also offers the possibility of studying spatial variations in the abundances of the species studied and in turn the gas properties traced along those sight lines in the survey that are spatially nearby.

As discussed in Section 3 the  $\text{p-H}_2\text{O}^+$  spectra observed toward both W3(OH) and W3 IRS5 are contaminated by high-lying  $\text{H}^{13}\text{CO}^+$  and  $\text{CH}_3\text{OH}$  lines and contributions of these emission features are taken into account when modeling the  $\text{p-H}_2\text{O}^+$  spectral line profiles. The  $\text{ArH}^+$  spectrum along the line of sight toward W3 IRS5 also shows two emission features at  $v_{\text{LSR}} > 10 \text{ km s}^{-1}$ , both corresponding to the same transition and arising from the image band, which we were unable to identify. However, lying at velocities beyond the typical velocity range over which absorption features are observed (for example in comparison to the  $\text{OH}^+$  spectrum displayed in Figure 5), we do not expect any significant contributions from these lines. Similarly, the SH absorption feature at the systemic velocity of W3(OH) blends with an emission feature not seen in previous observations of SH toward other sight lines and showing only weak hints in the spectrum toward W3 IRS5 at  $v_{\text{LSR}} \sim -55 \text{ km s}^{-1}$ . The nondetection of this feature toward other sight lines and its weaker presence toward W3 IRS5 at the same velocity suggests that this line arises from the image band. Covering only a narrow bandwidth and observed at only a single tuning frequency we are at this time unable to attribute this feature to any specific species. While

this renders the true depth of the absorption-line profile uncertain, the column densities that are subsequently derived toward the systemic velocity of the source (as will be discussed in more detail in Section 4.2) only represent lower limits.

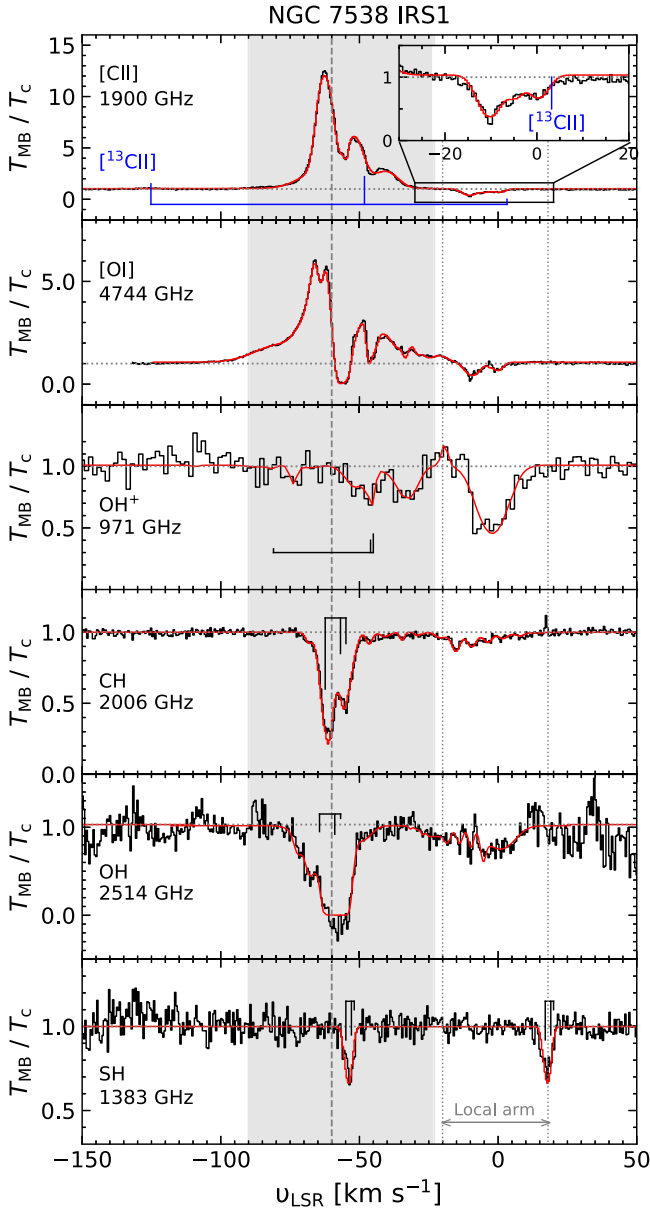
#### 4.1.2. NGC 7538 IRS1

The vicinity of the H II region NGC 7538, located at a distance of  $2.65 \text{ kpc}$  (Moscadelli et al. 2009) in the Perseus arm, is host to multiple infrared sources of which the brightest ones are NGC 7538 IRS 1, IRS 9, and South (S). The most massive of these, NGC 7538 IRS 1, an embedded O-type star, is a remarkably rich maser source at the heart of a young stellar cluster. Moreover, studies by Klaassen et al. (2009) on the dynamics of the warm molecular gas surrounding this hypercompact H II region is indicative of rotation. With a systemic velocity of  $-59 \text{ km s}^{-1}$ , most of the absorption at  $v_{\text{LSR}} < -40 \text{ km s}^{-1}$  is associated with material from the background source, while the absorption features seen at  $v_{\text{LSR}} > -20 \text{ km s}^{-1}$  arise from foreground material in the Local Arm. The self-absorption features seen in the [C II] and [O I] spectra are offset from the systemic velocity of the IRS1 source near  $v_{\text{LSR}} \sim -55 \text{ km s}^{-1}$ , in agreement with the complex kinematics of IRS1 discussed by Beuther et al. (2013).

#### 4.2. Analysis

The observed spectral line features are modeled using an enhanced version of the eXtended CASA Line Analysis Software Suite (XCLASS;<sup>22</sup> Möller et al. 2017), which solves the 1D radiative transfer equation for an isothermal source assuming that the level populations are governed by Boltzmann statistics at a common temperature,  $T_{\text{rot}}$ , such that the excitation

<sup>22</sup> See <https://xclass.astro.uni-koeln.de/> for further information.

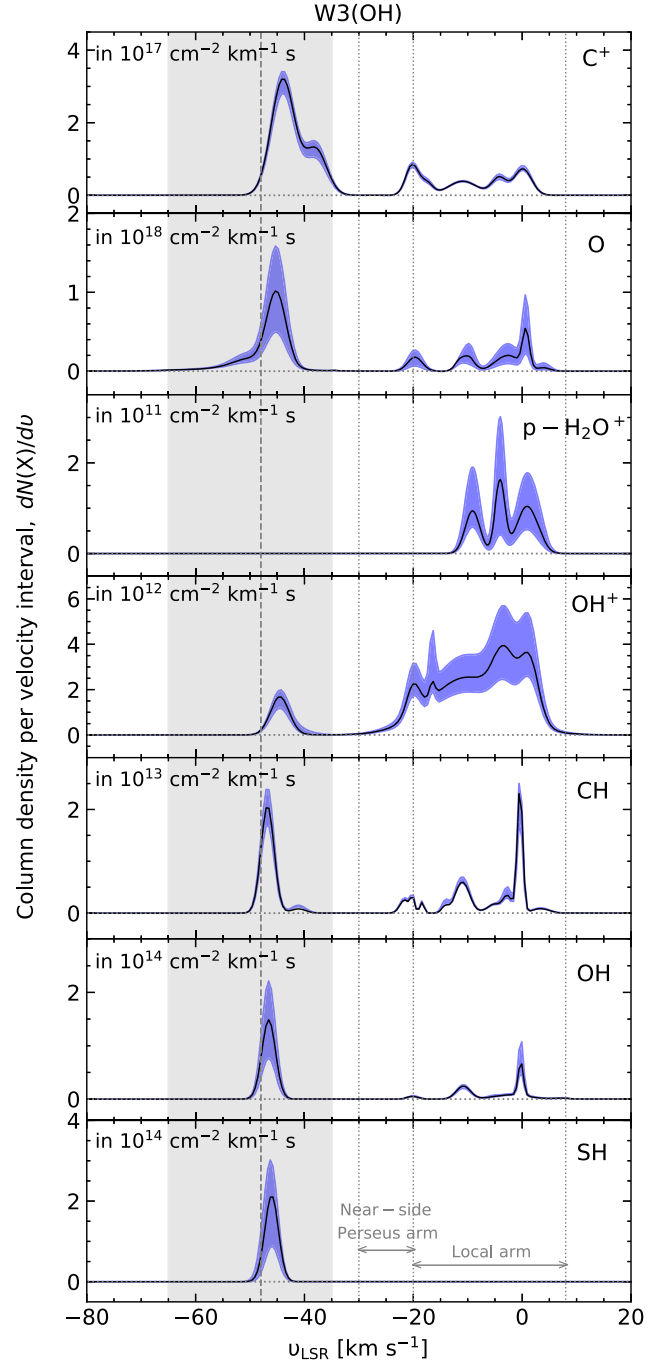


**Figure 6.** Same as Figure 4 but toward NGC 7538 IRS1, with only the SOFIA/GREAT observations.

temperature,  $T_{\text{ex}}$ , is the same for all pairs of states. The spectral line profiles are fitted with multiple Gaussian components, taking into account the effects of optical depth on the line shape for a finite source size and dust attenuation using the automated fitting routine provided by MAGIX (Möller et al. 2013). From the resultant modeled optical depth profile, we determine the column densities per velocity interval,  $dN/dv$ , using

$$\left(\frac{dN}{dv}\right)_i = \frac{8\pi\nu_i^3}{c^3} \frac{Q(T_{\text{rot}})}{g_u A_{u,l}} e^{E_u/T_{\text{ex}}} \left[ \exp\left(\frac{h\nu_{u,l}}{k_B T_{\text{ex}}}\right) - 1 \right]^{-1} \tau_i \quad (3)$$

for each velocity channel,  $i$ . For a given hyperfine-structure splitting transition, the spectroscopic parameters  $g_u$  (the upper level degeneracy),  $E_u$  (the upper level energy), and  $A_{u,l}$  (the Einstein A coefficient between any two given levels  $u$  and  $l$ ) remain constant, except for the partition function,  $Q$ , which itself is a function of the rotation temperature,  $T_{\text{rot}}$ . We further



**Figure 7.** Top to bottom: channel-wise column density profile ( $dN/dv$ ) of  $\text{C}^+$ ,  $\text{O}$ ,  $\text{p-H}_2\text{O}^+$ ,  $\text{OH}^+$ ,  $\text{CH}$ ,  $\text{OH}$ , and  $\text{SH}$  for the foreground absorbing clouds along the line of sight toward W3(OH). The blue shaded region represents the uncertainties ( $2\sigma$  confidence interval), while the gray dashed line and shaded region mark the systemic velocity of the source and highlight the velocity dispersion of the source, respectively.

assume that almost all the hydrides and hydride ions in this study occupy the ground rotational state where the excitation temperature,  $T_{\text{ex}}$ , is determined by the CMB radiation temperature,  $T_{\text{CMB}}$ , of 2.73 K. This is a valid assumption for the conditions that prevail in diffuse clouds where the gas densities, ( $n(\text{H}) \leq 100 \text{ cm}^{-3}$ ), are sufficiently low that collisional excitation becomes unimportant. While it has been shown through the comparison of CH column densities derived

**Table 5**  
Synopsis of Derived Column Densities and Inferred Results

Source Designation	$v_{\text{LSR}}$ range (km s <sup>-1</sup> )	$N(\text{ArH}^+)$ (10 <sup>12</sup> cm <sup>-2</sup> )	$N(\text{p-H}_2\text{O}^+)$ (10 <sup>12</sup> cm <sup>-2</sup> )	$N(\text{OH}^+)$ (10 <sup>13</sup> cm <sup>-2</sup> )	$N(\text{OH})$ (10 <sup>14</sup> cm <sup>-2</sup> )	$N(\text{SH})^{\text{a}}$ (10 <sup>13</sup> cm <sup>-2</sup> )	$N(\text{CH})$ (10 <sup>13</sup> cm <sup>-2</sup> )
W3(OH)	-51 to -39 <sup>d</sup>	...	...	>0.73	>4.67	>6.70	>6.48
	-25 to -8	...	0.21 <sup>+0.25</sup> <sub>-0.13</sub>	3.03 <sup>+1.60</sup> <sub>-1.10</sub>	0.90 <sup>+0.12</sup> <sub>-0.12</sub>	<0.10	3.14 <sup>+0.30</sup> <sub>-0.30</sub>
	-8 to +9	...	0.91 <sup>+0.80</sup> <sub>-0.55</sub>	3.86 <sup>+2.30</sup> <sub>-1.70</sub>	1.41 <sup>+0.57</sup> <sub>-0.28</sub>	<0.10	4.60 <sup>+0.70</sup> <sub>-0.66</sub>
W3 IRS5	-55 to -35 <sup>d</sup>	>0.01	...	>0.06	>9.02	>1.10	>9.60 <sub>s</sub>
	-28 to -8	0.51 <sup>+0.16</sup> <sub>-0.19</sub>	...	2.22 <sup>+1.43</sup> <sub>-0.97</sub>	0.75 <sup>+0.05</sup> <sub>-0.05</sub>	<0.09	2.67 <sup>+1.22</sup> <sub>-1.22</sub>
	-8 to +10	0.27 <sup>+0.08</sup> <sub>-0.08</sub>	0.21 <sup>+0.22</sup> <sub>-0.14</sub>	3.48 <sup>+2.00</sup> <sub>-1.64</sub>	0.51 <sup>+0.10</sup> <sub>-0.07</sub>	<0.10	1.50 <sup>+1.40</sup> <sub>-1.00</sub>
NGC 7538 IRS1	-65 to -39 <sup>d</sup>	...	...	>0.93	>10.44	>22.00	>55.68
	-39 to -31	...	...	0.35 <sup>+0.04</sup> <sub>-0.03</sub>	...	<0.05	0.02 <sup>+0.02</sup> <sub>-0.01</sub>
	-31 to -20	...	...	0.28 <sup>+0.05</sup> <sub>-0.05</sub>	0.14 <sup>+0.10</sup> <sub>-0.02</sub>	<0.07	0.83 <sup>+0.60</sup> <sub>-0.40</sub>
	-17 to -3	...	...	1.01 <sup>+0.40</sup> <sub>-0.31</sub>	0.43 <sup>+0.22</sup> <sub>-0.16</sub>	<0.13	5.96 <sup>+1.60</sup> <sub>-1.60</sub>
	-3 to +18	...	...	1.86 <sup>+0.62</sup> <sub>-0.54</sub>	0.45 <sup>+0.08</sup> <sub>-0.05</sub>	<0.09	2.82 <sup>+0.37</sup> <sub>-0.26</sub>

Source Designation	$v_{\text{LSR}}$ range (km s <sup>-1</sup> )	$N(\text{H})^{\text{b}}$ (10 <sup>21</sup> cm <sup>-2</sup> )	$N(\text{H}_2)^{\text{c}}$ (10 <sup>21</sup> cm <sup>-2</sup> )	$N(\text{C}^+)$ (10 <sup>17</sup> cm <sup>-2</sup> )	$N(\text{O})$ (10 <sup>18</sup> cm <sup>-2</sup> )
W3(OH)	-51 to -39 <sup>d</sup>	1.30 <sup>+0.02</sup> <sub>-0.02</sub>	>1.85	>23.00	>5.57
	-25 to -8	1.39 <sup>+0.03</sup> <sub>-0.03</sub>	0.90 <sup>+0.54</sup> <sub>-0.37</sub>	5.51 <sup>+0.40</sup> <sub>-0.40</sub>	1.26 <sup>+0.68</sup> <sub>-0.63</sub>
	-8 to +9	0.95 <sup>+0.03</sup> <sub>-0.03</sub>	1.31 <sup>+0.81</sup> <sub>-0.55</sub>	4.80 <sup>+0.57</sup> <sub>-0.42</sub>	2.00 <sup>+1.56</sup> <sub>-1.08</sub>
W3 IRS5	-55 to -35 <sup>d</sup>	>2.60	>2.74	>5.96	>2.21
	-28 to -8	1.20 <sup>+0.26</sup> <sub>-0.26</sub>	0.76 <sup>+0.57</sup> <sub>-0.46</sub>	6.75 <sup>+2.43</sup> <sub>-2.04</sub>	0.83 <sup>+0.46</sup> <sub>-0.34</sub>
	-8 to +10	0.97 <sup>+0.21</sup> <sub>-0.21</sub>	0.42 <sup>+0.46</sup> <sub>-0.32</sub>	8.00 <sup>+3.06</sup> <sub>-2.40</sub>	1.15 <sup>+0.63</sup> <sub>-0.52</sub>
NGC 7538 IRS1	-65 to -39 <sup>d</sup>	...	>16.00	>0.69	>8.24
	-39 to -31	...	0.01 <sup>+0.01</sup> <sub>-0.01</sub>	...	0.05 <sup>+0.02</sup> <sub>-0.02</sub>
	-31 to -20	...	0.23 <sup>+0.21</sup> <sub>-0.14</sub>	...	...
	-17 to -3	2.96 <sup>+0.03</sup> <sub>-0.03</sub>	1.70 <sup>+1.12</sup> <sub>-0.82</sub>	9.80 <sup>+2.70</sup> <sub>-2.00</sub>	1.01 <sup>+0.82</sup> <sub>-0.70</sub>
	-3 to +18	1.99 <sup>+0.05</sup> <sub>-0.05</sub>	0.81 <sup>+0.50</sup> <sub>-0.33</sub>	3.33 <sup>+0.71</sup> <sub>-0.54</sub>	0.42 <sup>+0.27</sup> <sub>-0.11</sub>

**Notes.**

<sup>a</sup> Represents the  $3\sigma$  upper limits on  $N(\text{SH})$  obtained toward line-of-sight material not associated with the background source for which we do not detect SH.

<sup>b</sup> The H I data toward W3(OH) and W3 IRS5 are extracted from the Canadian Galactic Plane Survey (CGPS; Taylor et al. 2003), and that toward NGC 7538 IRS1 is taken from Lebrón et al. (2001).

<sup>c</sup>  $N(\text{H}_2)$  values were derived using  $N(\text{CH})$  following the relationship determined by Sheffer et al. (2008),  $N(\text{CH})/N(\text{H}_2) = 3.5^{+2.1}_{-1.4} \times 10^{-8}$ .

<sup>d</sup> Indicates the velocity intervals corresponding to the molecular cloud.

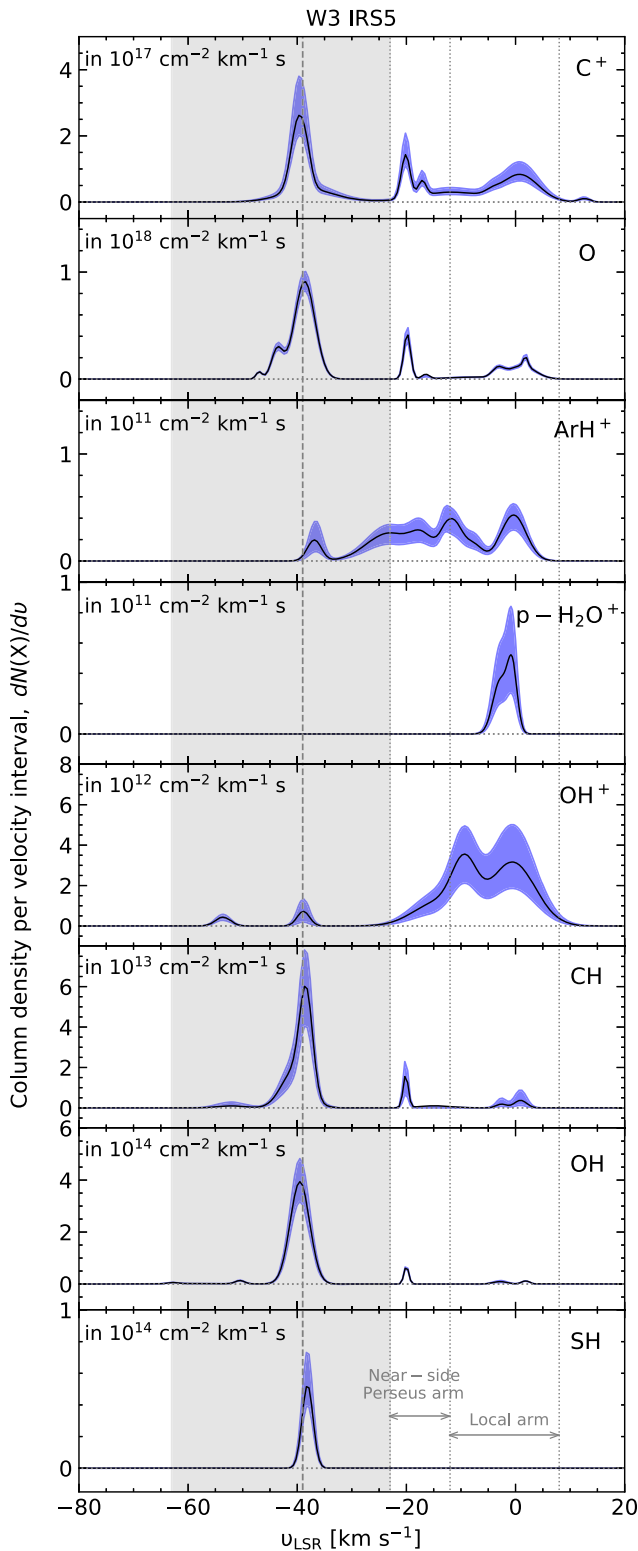
from its 532 GHz and 2007 GHz transitions that weak pumping by the ambient far-infrared radiation field can increase the excitation temperature slightly above the CMB value, this effect only has a small impact on the derived column densities (Wiesemeyer et al. 2018). Changing the assumed  $T_{\text{ex}}$  from 3.1 to 10 K only results in a  $\sim 10\%$  increase in the derived column density. Similarly, by estimating the abundance of  $\text{H}_2\text{O}$  in its excited levels, Flagey et al. (2013) were able to constrain  $T_{\text{ex}} \approx 5$  K, for the ground-state transitions of its ortho-spin state. However, the assumption of a low excitation temperature, equivalent to the CMB value may not be valid for those features that arise from the dense envelopes of the molecular clouds in which collisions can dominate, resulting in higher excitation temperatures (Emprechtinger et al. 2013). Therefore, the column densities derived by integrating Equation (3) over velocity intervals associated with the molecular cloud components only represent lower limits to the column density values.

The spectra of certain species (such as those of C II and O I; see the top panels of Figures 4 to 6), which have lower critical densities and higher optical depths than the hydride species, are more complex, containing both absorption and emission features near the envelope of the molecular cloud/hot core. Here, the emission features corresponding to the core are

modeled with higher excitation temperatures,  $T_{\text{ex}} \leq 120$  K, corresponding to gas temperatures typically found in the envelope material surrounding the central molecular cloud. Both the emission and absorption features are fitted simultaneously by assuming different cloud layers. The former corresponds to a core component that probes the dense, hot cores associated with the studied molecular clouds, while the modeling of the latter is implemented using foreground layers which mimics the source envelopes and line of sight.

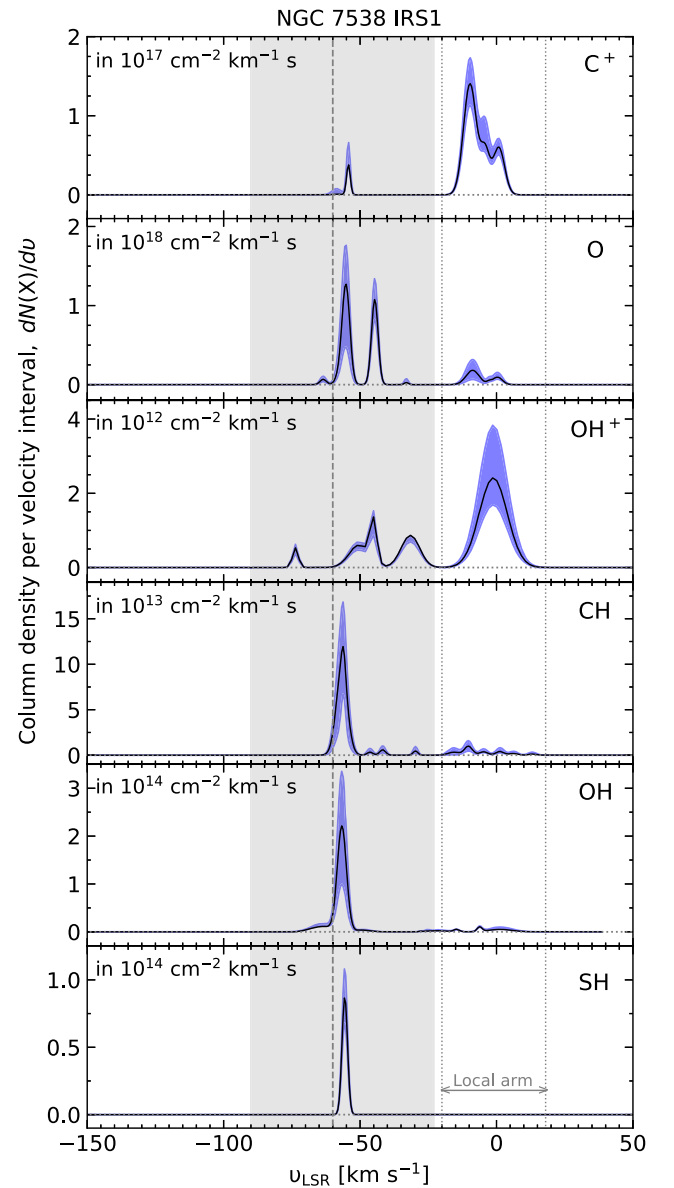
Contributions from the column densities subsequently determined for the emission features are removed from the total distribution revealing only the column density distribution of the underlying absorption from foreground material not associated with the source. The resulting column density profiles are displayed in Figures 7–9. The total column densities estimated over specific velocity intervals, each of which broadly correspond to a spiral arm crossing are summarized in Table 5. The tabulated column densities represent the total columns, which for those species with  $\Lambda$ -doublet splitting contain contributions from both sets of doublets.

The reliability of the column densities derived here are mainly limited by uncertainties present in the continuum level and not by the absolute temperature calibration. We compute



**Figure 8.** Same as Figure 7 but toward W3 IRS5 with the addition of the ArH<sup>+</sup> column density profile.

the errors using Bayesian methods to sample a posterior distribution of the optical depths ( $\propto \ln(T_1/T_0)$ ). We sampled a set of model spectra, each generated by iteratively adding a pseudorandom noise contribution to the absorption spectra, for each of the different species prior to running the XCLASS fits. The number of samplers is set by evaluating the posterior

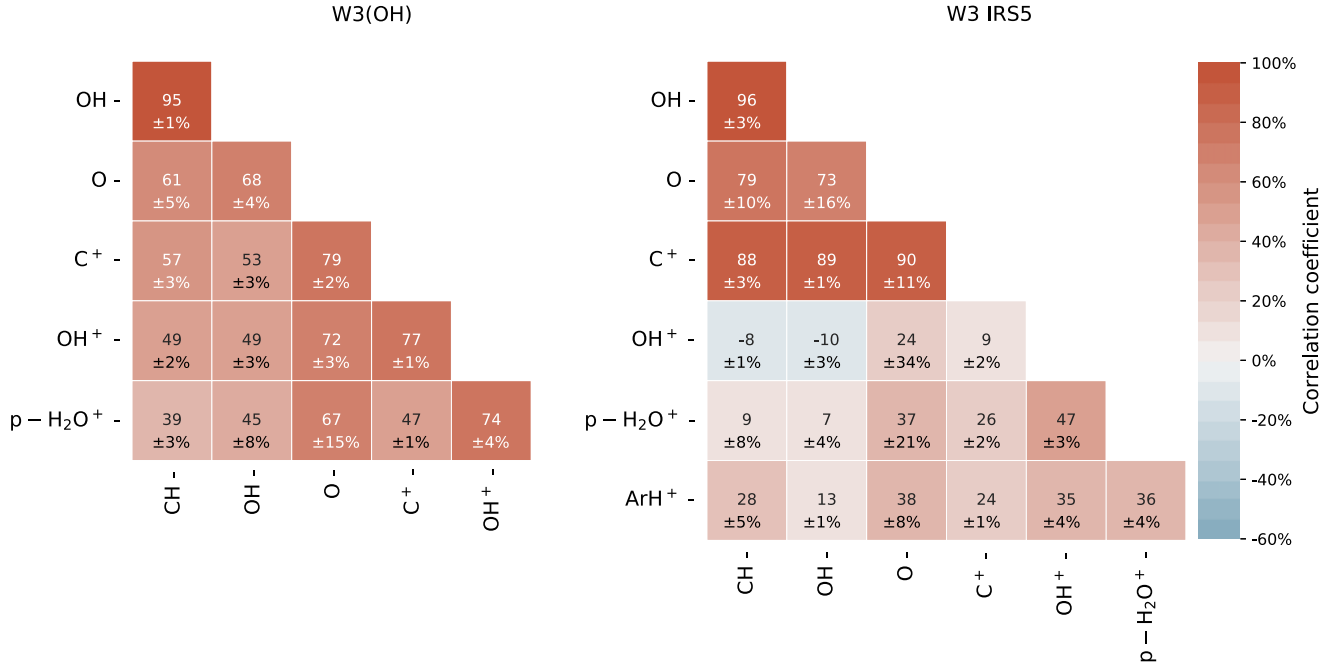


**Figure 9.** Same as Figure 7 but toward NGC 7538 IRS1 but without p-H<sub>2</sub>O<sup>+</sup>.

precision and variance, while the standard deviation in the additive noise is fixed to be the same as that in the line-free part of each spectrum. The best fit curves generated by XCLASS for each spectra samples a point in the channel-wise column density distribution for any given source along any given sight line. The 5% and 95% percentiles obtained from the distributions are then used to quantify the  $2\sigma$  confidence intervals.

## 5. Discussion

A full statistical investigation of the diffuse gas properties traced using hydrides (as outlined in Section 2) will require a combination of the complete HyGAL SOFIA data set with the ancillary IRAM and JVLA observations. The goal of the discussion section in the present paper is more modest: here we merely illustrate the analysis techniques that can be used to explore the scientific questions addressed by the HyGAL project. In Sections 5.1 and 5.2, we discuss the column density



**Figure 10.** Pearson product-moment correlation coefficients as computed using Equation (4) (represented here as percentages) between the column densities determined toward only those features which are not associated with the background source for W3(OH) (left) and W3 IRS5 (right).

**Table 6**  
Range of Abundances Derived in the Foreground Absorbing Clouds

Species	$N(X)/N(H)$	$N(X)/N(H_2)$	$N(X)/N_H$
ArH <sup>+</sup>	$6.73 - 8.42 \times 10^{-10}$	$2.13 - 5.57 \times 10^{-10}$	$0.92 - 2.10 \times 10^{-10}$
p-H <sub>2</sub> O <sup>+</sup>	$0.15 - 1.00 \times 10^{-9}$	$1.00 - 3.16 \times 10^{-10}$	$0.37 - 1.36 \times 10^{-10}$
OH <sup>+</sup>	$0.34 - 4.06 \times 10^{-8}$	$0.60 - 3.55 \times 10^{-8}$	$0.16 - 1.88 \times 10^{-8}$
OH	$0.14 - 1.48 \times 10^{-7}$	$2.53 - 6.01 \times 10^{-8}$	$0.67 - 2.02 \times 10^{-8}$
C <sup>+</sup>	$1.67 - 8.23 \times 10^{-4}$	$1.58 - 8.14 \times 10^{-4}$	$0.68 - 2.73 \times 10^{-4}$
O	$0.21 - 2.10 \times 10^{-3}$	$0.47 - 8.31 \times 10^{-3}$	$0.12 - 0.39 \times 10^{-3}$

ratios implied by our absorption-line observations and how they vary. Although the goals and analysis of the HyGAL survey concentrate on the foreground diffuse clouds along the sight lines toward bright sources, the SOFIA HyGAL data set provides a wealth of information on the background sources as well. We demonstrate this in Section 5.3 through the analysis of the complex line profiles of the 158  $\mu\text{m}$  [C II] emission, using constraints from the [<sup>13</sup>C II] line also covered in the same sideband.

### 5.1. Column Density Ratios

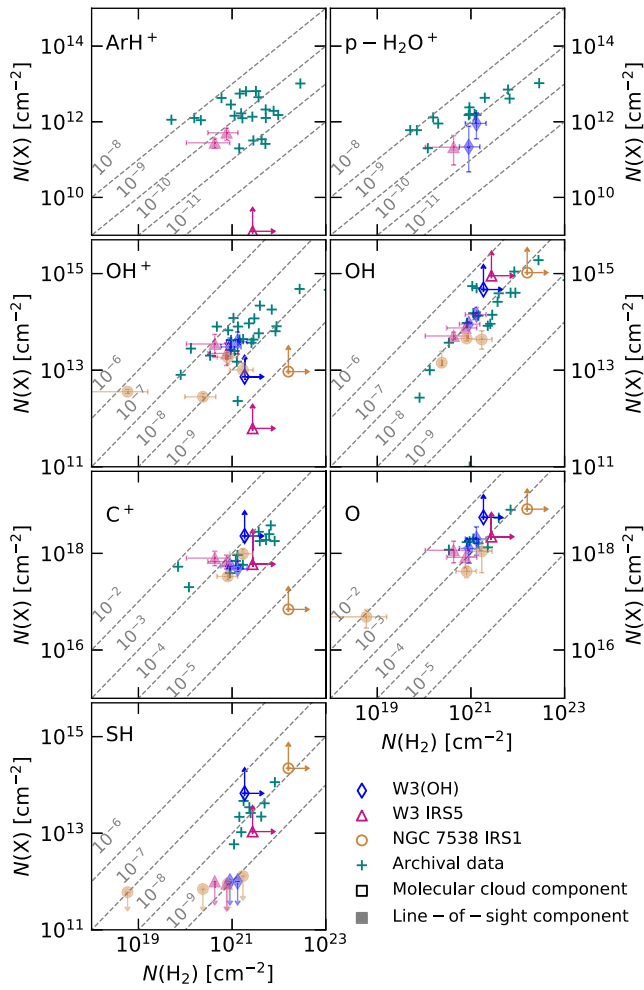
From Section 4.2 it is clear that the line profiles of some of the studied species are much more similar in shape across any given sight line than others. For example, the spectral line profiles of the diffuse molecular gas tracers in our survey, CH and OH, show similar absorption-line features across the range of velocities probed. Their spectra reveal overall narrower features ( $\sim 0.5 \text{ km s}^{-1}$ ) in comparison to the broader and more continuous profiles traced by the diffuse atomic gas tracers studied like OH<sup>+</sup>. Moreover, the trends in absorption-line strengths observed—and in turn the abundances derived—across a sight line differ from one hydride to another. Therefore, in this section we present a quantitative discussion of the observed line-of-sight properties using the column

densities derived as discussed in Section 4. This is best illustrated in Figure 10, which presents cross correlations between the column densities of different pairs of species studied. The correlations presented are computed using the Pearson product-moment correlation coefficient,  $r$ , which describes the strength of the linear relationship between any given pair of spectral lines by using the standard deviation of each data set and the covariance between them. The correlation coefficient between any two species  $X$  and  $Y$  is computed as follows:

$$r = \frac{\int dv \left( \frac{dN(X)}{dv} - \langle \frac{dN(X)}{dv} \rangle \right) \left( \frac{dN(Y)}{dv} - \langle \frac{dN(Y)}{dv} \rangle \right)}{\sqrt{\int dv \left( \frac{dN(X)}{dv} - \langle \frac{dN(X)}{dv} \rangle \right)^2} \sqrt{\int dv \left( \frac{dN(Y)}{dv} - \langle \frac{dN(Y)}{dv} \rangle \right)^2}}, \quad (4)$$

where  $dN/dv$  and  $\langle dN/dv \rangle$  represent the column densities computed per velocity channel and the average of the column densities computed over all relevant velocity channels, respectively.

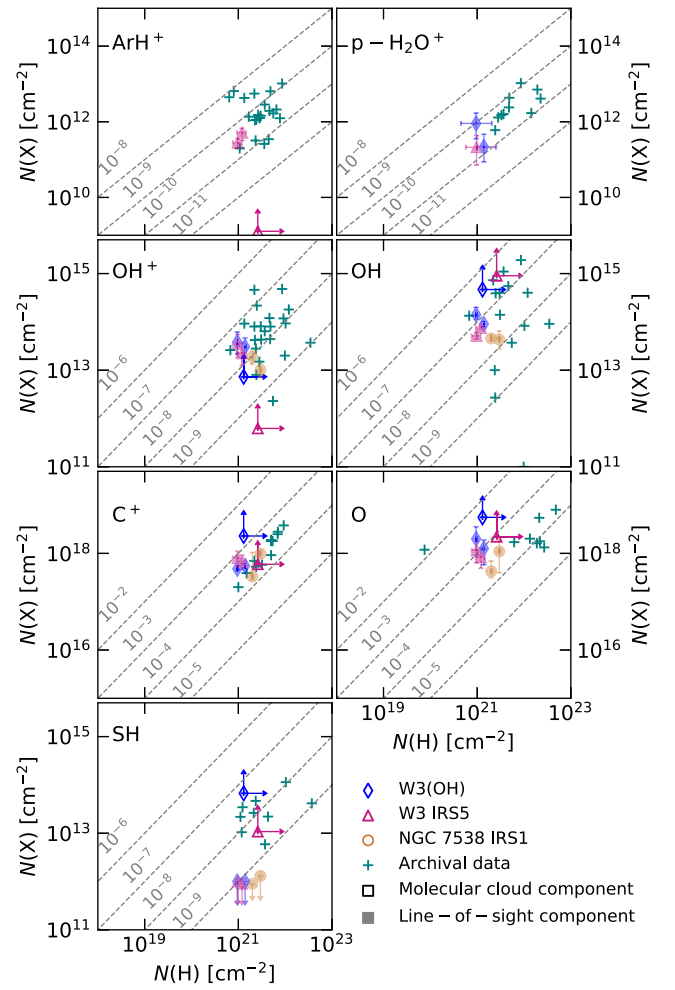
In Figures 11 to 13, we compare the column densities of all the studied species with that of molecular hydrogen, atomic hydrogen and the total gas column,  $N_H = N(H) + 2 \times N(H_2)$  derived over specific velocity intervals, respectively. For



**Figure 11.** Comparison of column densities of the key HyGAL species with  $N(\text{H}_2)$  determined using  $N(\text{CH})/N(\text{H}_2) = 3.5 \times 10^{-8}$  (Sheffer et al. 2008) and integrated over the velocity intervals discussed in Table 5. The data points toward W3(OH), W3 IRS5, and NGC 7538 IRS1 are plotted using blue diamonds, pink triangles, and dark orange circles, respectively where the filled and unfilled markers demarcate the column densities determined toward line-of-sight components (features that are not associated with the background molecular cloud) and molecular cloud components as discussed in Table 5. The dashed grey lines represent lines of constant abundances  $N(\text{X})/N(\text{H}_2)$  for each studied species X as labeled on the top left-hand corner of each panel. The data represented using teal crosses present results from previous observations as discussed in the text.

comparison, we also include the column density results for the same species obtained from previous Herschel/HIFI and SOFIA/GREAT observations carried out toward other star-forming regions in the Galaxy. The data used in this comparison are primarily taken from Schilke et al. (2014) and Jacob et al. (2020b) for  $\text{ArH}^+$ , Indriolo et al. (2015) and Jacob et al. (2020b) for  $\text{p-H}_2\text{O}^+$ , Indriolo et al. (2015) for  $\text{OH}^+$ , Wiesemeyer et al. (2016) for OH and atomic oxygen, Gerin et al. (2015) for  $\text{C}^+$ , and Neufeld et al. (2015) for SH.

Widely established and used as a tracer of CO-dark molecular gas (Federman 1982; Sheffer et al. 2008; Gerin et al. 2010b; Wiesemeyer et al. 2018; Jacob et al. 2019), CH is used as a proxy for  $\text{H}_2$  in our analysis, using the relationship determined by Sheffer et al. (2008),  $N(\text{CH})/N(\text{H}_2) = 3.5^{+2.1}_{-1.4} \times 10^{-8}$ . Since we

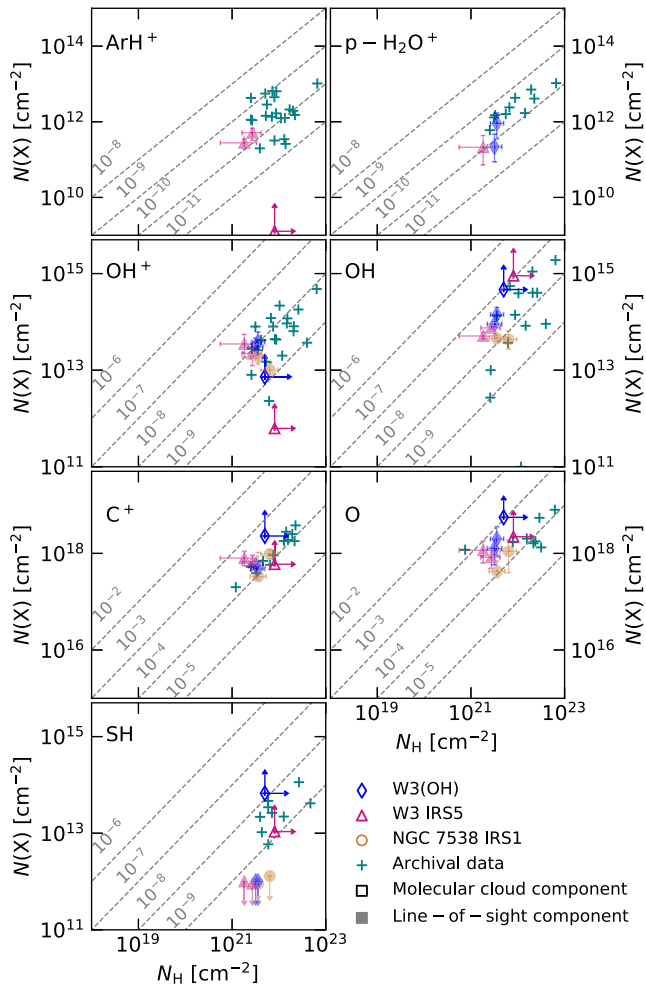


**Figure 12.** Same as Figure 11 but a comparison of column densities with respect to  $N(\text{H})$ .

determine  $N(\text{H}_2)$  using the  $N(\text{CH})$  values derived, we do not include CH in our abundance analysis.

Limited to a small sample of data points, a clear correlation is only discerned for the oxygen-bearing species. As discussed in Section 2.2, the formation of  $\text{OH}^+$  is initiated by cosmic rays, involving reactions with atomic and molecular hydrogen, while the formation of OH greatly relies on the  $\text{H}_2$  fraction. This is well reflected in the column density comparisons presented in Figures 11–13. While both OH and atomic oxygen show positive correlations with increasing  $N(\text{H}_2)$  or molecular content,  $\text{OH}^+$  shows no clear correlation with a larger scatter. The tight correlation observed between OH and CH (and in turn  $\text{H}_2$ ) has also been explored at UV wavelengths by Weselak et al. (2009) and Mookerjee (2016), who derive a  $N(\text{OH})/N(\text{CH})$  ratio of  $\sim 2.6$ , equivalent to a  $N(\text{OH})/N(\text{H}_2)$  ratio of  $\sim 10^{-7}$  (as shown in Figures 11 and 8 of Jacob et al. (2019)). At the Galactocentric distances of these sources the average atomic oxygen abundance derived with respect to  $N_{\text{H}}$  of  $\sim 4 \times 10^{-4}$  is consistent with what is expected when using the Galactic [O I] gradient as estimated by Korotin et al. (2014), who determined a slope of  $-0.058 \text{ dex kpc}^{-1}$  using observations of the near-infrared O I triplet toward a large sample of Cepheids.





**Figure 13.** Same as Figure 11 but a comparison of column densities with respect to  $N_{\text{H}}$ .

While no clear trend is discernible for the atomic gas tracers,  $\text{ArH}^+$  and  $\text{p-H}_2\text{O}^+$  (even with respect to  $N(\text{H})$ ; see Figure 12) where the usefulness of the latter is compromised by contaminating features, the abundances of the line-of-sight components are consistent with what was derived in Indriolo et al. (2015), Bialy et al. (2019), and Jacob et al. (2020b), of a few times  $10^{-10}$ . The lack of a clear correlation with  $N(\text{H})$  also suggests that only a small fraction of the  $\text{H I}$  gas traces the same cloud population as those traced by  $\text{ArH}^+$  and  $\text{p-H}_2\text{O}^+$ .

For  $\text{SH}$ , the derived abundances roughly span over 1 order of magnitude between  $4 \times 10^{-9}$  and  $3.5 \times 10^{-8}$ , consistent with the range of abundances derived in Neufeld et al. (2015). However, unlike the  $\text{SH}$  spectra observed toward the sight lines studied in Neufeld et al. (2015), the  $\text{SH}$  spectra presented here do not show any foreground absorption from line-of-sight clouds, which are not associated with the sources themselves. The upper limits on the  $\text{SH}$  column densities we derive at a  $3\sigma$  rms level for these features vary between  $\sim 1 \times 10^{12}$  and  $3.8 \times 10^{12} \text{ cm}^{-2}$ , which correspond to abundance limits between  $5.1 \times 10^{-10}$  and  $5.6 \times 10^{-9}$  relative to  $\text{H}_2$ .

We note that some of the upper limits derived here are smaller than the range of  $\text{SH}$  abundances derived for diffuse clouds along the lines of sight toward other star-forming regions by Neufeld et al. (2015). This suggests that the nondetection of  $\text{SH}$  is not solely an issue of sensitivity, but also that the abundance of  $\text{SH}$  along the sight lines presented is

lower than those measured previously for foreground diffuse clouds. As in the case of  $\text{SH}$ , no line-of-sight absorption is seen in the spectra of related sulfur-bearing species such as  $\text{CS}^{23}$ ,  $\text{H}_2\text{S}$  and  $\text{SO}$  observed using the IRAM 30 m telescope discussed in Section 3.3 (the spectra of which will be presented in a forthcoming paper); neither is it seen in the  $\text{SH}^+$  spectra taken from Herschel science archives toward  $\text{W3(OH)}$ ,  $\text{W3 IRS5}$ , and  $\text{NGC 7538 IRS1}$ . However, foreground absorption from clouds not associated with the background sources are observed in the Herschel spectra of  $\text{CH}^+$ , which like  $\text{SH}^+$  traces turbulent dissipation regions. The presence of  $\text{CH}^+$  and absence of  $\text{SH}^+$  (and by extension  $\text{SH}$ ) in these clouds can be explained by the large differences in their endothermicities ( $\Delta E_{(\text{CH}^+)}/k = 4640 \text{ K}$  and  $\Delta E_{(\text{SH}^+)}/k = 9860 \text{ K}$ ) of the ion-molecular reactions that are responsible for their formation. This, in addition to the lower effective temperatures and velocity drifts required, favors the production of  $\text{CH}^+$  over  $\text{SH}^+$  as discussed in Godard et al. (2014). Therefore, in general the lack of foreground absorption in sulfur-bearing species toward these sources might be because their short sight lines do not probe sufficiently shocked or turbulence dissipated environments necessary for the formation of these molecules.

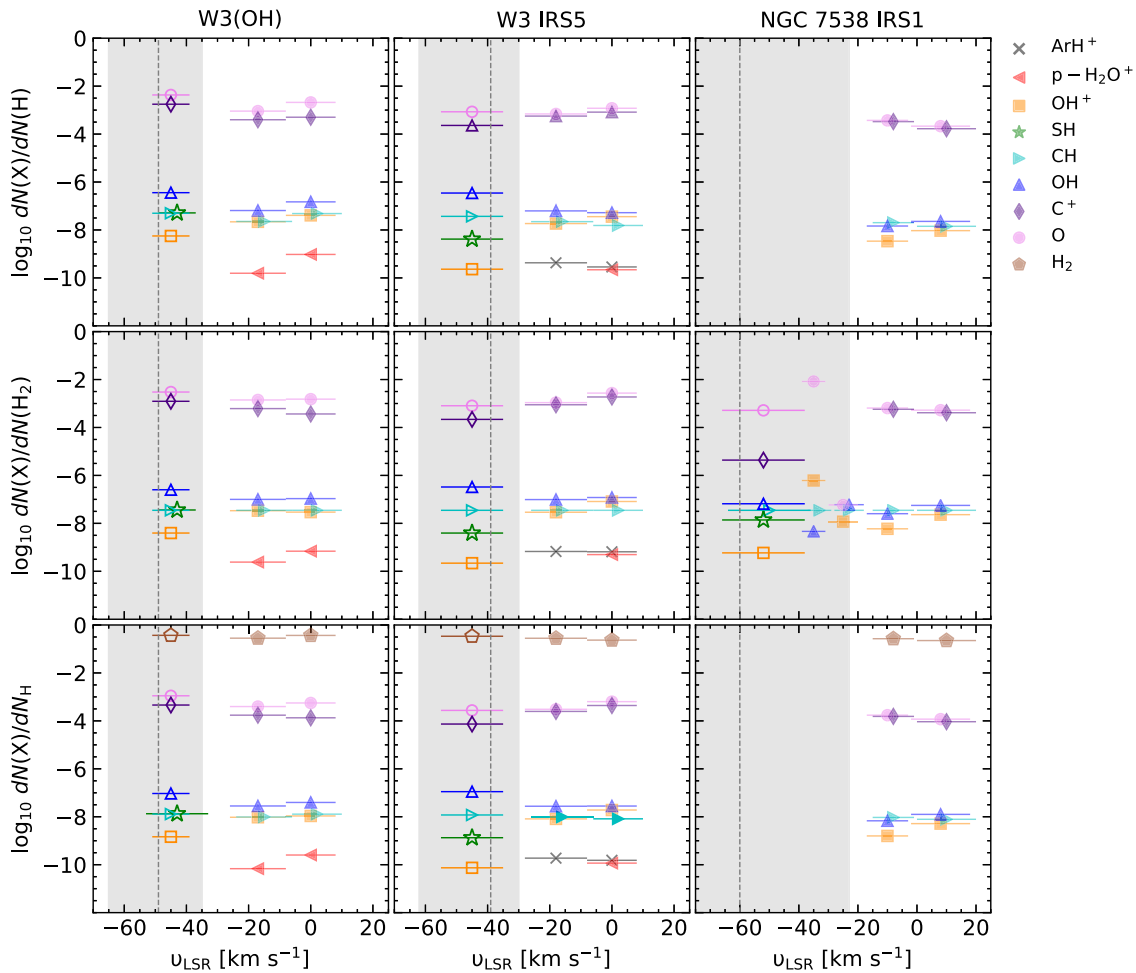
Since the molecular gas column traced along the sight lines discussed is greater than or equivalent to the atomic gas column, the overall trends observed when comparing the column densities of the different species studied with that of molecular hydrogen and the total gas column, are consistent with one another (see Figures 11 and 13). The main difference arises from the fact that there are fewer atomic gas components measured toward  $\text{NGC 7538 IRS1}$ . This is because the  $\text{H I}$  spectrum toward  $\text{NGC 7538 IRS1}$  is extracted from a neighboring region  $\text{G111.61} + 0.37$  (Lebrón et al. 2001), and therefore we can only assume that the foreground material along the lines-of-sight toward both clouds is the same. Eliminating such uncertainties the auxiliary data collected using the JVLA under the umbrella of the HyGAL program (see Section 3.3) will greatly aid in the analysis and interpretation of the targeted hydrides and other small interstellar molecules.

## 5.2. Variations Along the Sight Line

Figure 14 displays the average column density ratios of the targeted species relative to atomic hydrogen and molecular hydrogen and the total gas column computed over the velocity intervals discussed in Table 5. While the ratios computed for the velocity intervals associated with the background source are highly uncertain, the column density ratios even among line-of-sight clouds greatly deviate from one another. The range of abundances derived for each species toward the foreground material is tabulated in Table 6. The atomic gas tracer  $\text{OH}^+$ , unlike the molecular gas tracers  $\text{CH}$  and  $\text{OH}$ , displays the most stark variation in the column density ratios inferred toward the molecular cloud and line-of-sight velocities.

Overall, variations in the column density ratios inferred from the foreground components associated with absorption against the Local Arm are similar for all three sources. Located at comparable distances beyond the solar circle, all three of the sight lines discussed in this paper probe similar foreground material, in particular that toward  $\text{W3(OH)}$  and  $\text{W3 IRS5}$ , both of which are star-forming regions contained in the  $\text{W3}$

<sup>23</sup> Except for  $\text{W3(OH)}$ , along the line of sight toward which we observe weak absorption against the Local Arm.

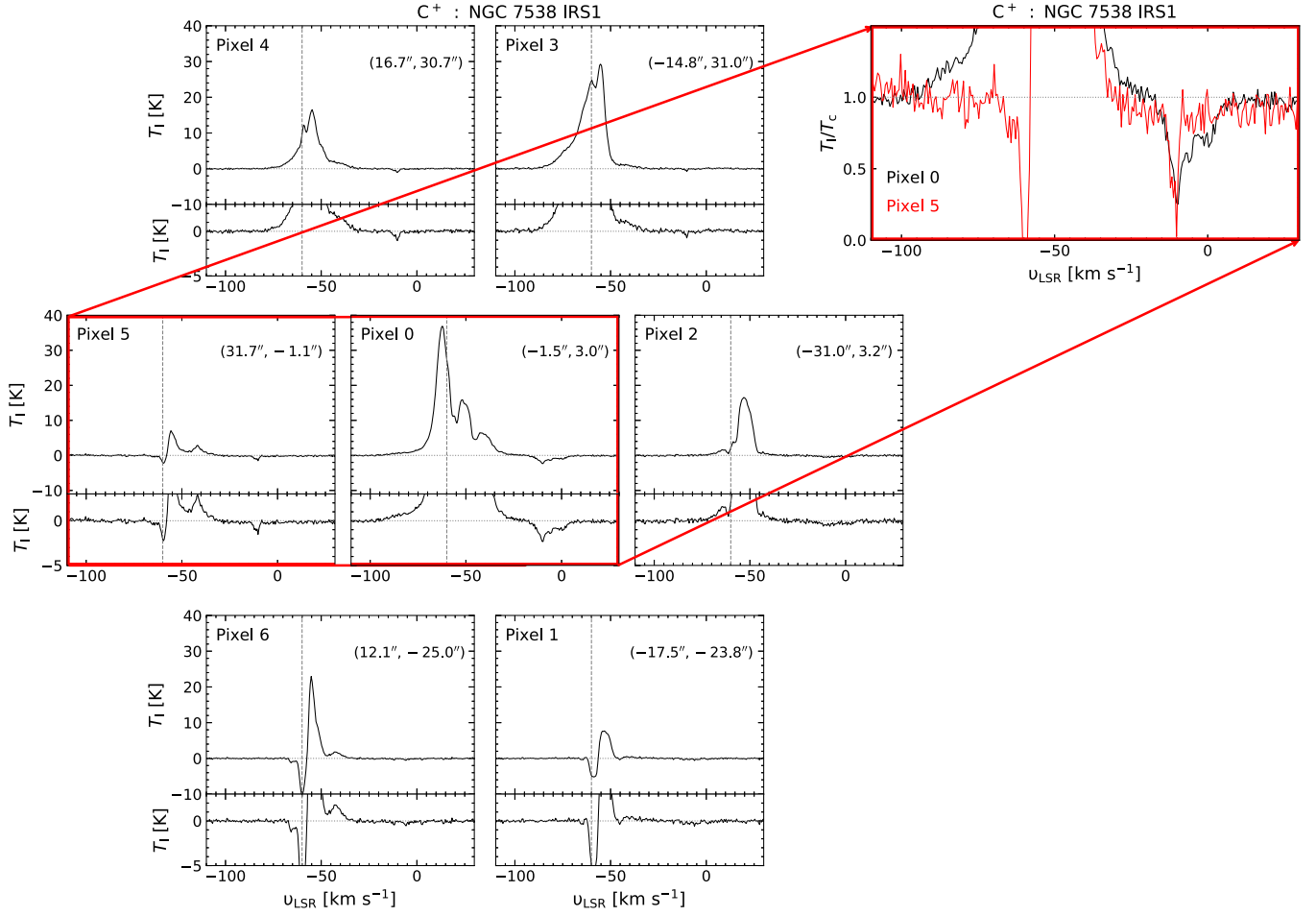


**Figure 14.** From top to bottom: inferred  $N(X)/N(H)$ ,  $N(X)/N(H_2)$ , and  $N(X)/N_H$  ratios over the adopted velocity intervals for W3(OH), W3 IRS5, and NGC 7538 IRS1 from left to right, with the different species denoted using different markers and colors as displayed in the legend. The filled and unfilled markers demarcate the column densities determined toward line-of-sight and molecular cloud features, respectively. The gray dashed lines and shaded regions indicate the systemic velocity and velocity dispersion of the source, respectively, where the column densities determined only form lower limits. Markers lying close to each other are shifted by  $+2 \text{ km s}^{-1}$  for clarity. Covering a large range of abundances, we do not display the corresponding errors in this plot as they coincide with the length scales of the markers.

complex. While the ratios averaged over line-of-sight velocity intervals for both W3(OH) and W3 IRS5 appear consistent with each other, examining their individual spectra shows apparent variations, as discussed in Section 4.1. A similar analysis has been carried out by Welty et al. (2020) at far-ultraviolet wavelengths in an attempt to characterize the properties and spatial structure for sight lines near HD 62542. In addition, for CH,  $C^+$  and O observed using the upGREAT seven pixel array receivers, LFA and HFA, we can measure fluctuations on smaller scales for sufficiently extended sources by examining the spectral line features obtained in the outer pixels; the latter are positioned in a hexagonal pattern around the center with a pixel spacing of  $\sim 31''8$  and  $\sim 13''6$  for the LFA and HFA, respectively. We illustrate this in Figure 15, which displays the [C II] spectral line features obtained in each of the seven LFA pixels toward NGC 7538 IRS1. While we detect [C II] emission in each individual pixel, the continuum brightness is only sufficient in a few of the off-center pixels to detect foreground absorption components (not associated with the molecular cloud). The spectrum obtained in pixel 5, for example, shows narrow absorption features near  $10 \text{ km s}^{-1}$  associated with material in the Local Arm, with an absorption-line strength

(when normalized with respect to the continuum temperature of 1.57 K) that is comparable to that observed in the central pixel. However, with weaker continuum levels of the order of 0.1 K, we do not observe significant ( $\geq 3\sigma$ ) line-of-sight absorption in the [C II] spectra obtained in the other off-center pixels.

The observed fluctuations in the abundance ratios are very useful to constrain properties of interstellar turbulence: its driving scale, driving mode (solenoidal versus compressional) and its intensity (e.g., the sonic Mach number discussed by Bialy et al. (2017, 2019)). The turbulence in cold gas is typically supersonic, thus producing significant density fluctuations in the gas, which are in turn translated into fluctuations in the molecular abundances. Thus larger dispersions in molecular abundances typically correspond to larger turbulence driving scales and higher sonic Mach numbers (Bialy et al. 2017, 2019; Bialy & Burkhardt 2020). For example, in their study Bialy et al. (2019) used Herschel observations of  $OH^+$ ,  $H_2O^+$ , and  $ArH^+$  and found that toward the observed sight lines, a model of turbulence driven on large scale with supersonic velocities is favored (see also Bialy et al. 2017; Bellomi et al. 2020, for an application to H and  $H_2$  data).



**Figure 15.** Left: LFA/upGREAT footprint displaying the [C II] spectral line profiles as observed in each of its seven pixels centered on the background continuum source NGC 7538 IRS1 in the top panels, while the bottom panels zoom-in on the respective spectra to better see the absorption components. Right: overlay comparing the line-of-sight absorption observed in the continuum normalized [C II] spectra obtained from pixel 0 (in black) and pixel 5 (in red).

Furthermore, the column density ratios in turn can be translated into molecular fractions. For example, toward W3 IRS5, using Equation (2) for CH, Equation (12) from Indriolo et al. (2015) for  $\text{OH}^+/\text{H}_2\text{O}^+$ , and Equation (9) from Jacob et al. (2020b) for  $\text{ArH}^+$  in the foreground clouds, we find that CH traces molecular fractions between  $f^N(\text{H}_2) \sim 0.66$  and 0.83, while  $\text{OH}^+$  and  $\text{ArH}^+$  trace gas with  $f^N(\text{H}_2) \sim 2 \times 10^{-2}$  and  $\sim 1-5 \times 10^{-3}$ , respectively, assuming an equilibrium ortho-to-para  $\text{H}_2\text{O}^+$  ratio of 3. The complete HyGAL SOFIA data set will enable us to study variations in the average abundances of or molecular fractions traced by the different species studied and across different spiral arms.

### 5.3. $^{12}\text{C}/^{13}\text{C}$ Isotopic Abundances

Owing to the lower ionization potential of atomic carbon (11.2 eV) in comparison to hydrogen (13.6 eV), ionized carbon,  $\text{C}^+$ , has proven invaluable as a tracer of both H II regions and cold neutral gas using its  $^2P_{3/2} \rightarrow ^2P_{1/2}$  fine-structure transition at  $158 \mu\text{m}$ . As one of the key coolants in the ISM alongside the [O I] line at  $63 \mu\text{m}$ , the [C II]  $158 \mu\text{m}$  line is responsible for the transition between different ISM phases and is therefore an excellent tracer for the warm ionized medium, the diffuse atomic gas, and the diffuse molecular gas.

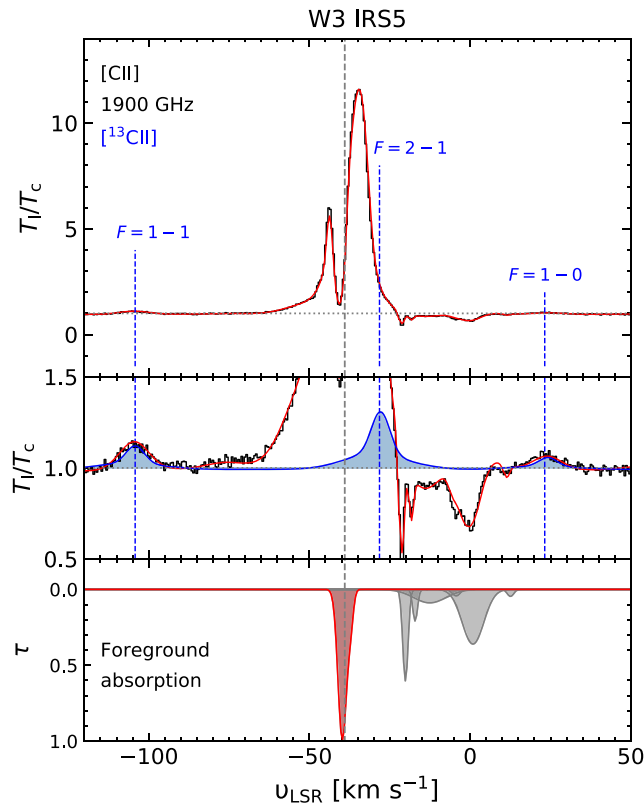
**Table 7**  
Spectroscopic Parameters of the  $^{13}\text{C}^+$  Transition near  $158 \mu\text{m}$

Transition $F' - F''$	Frequency (GHz)	$A_{u,1}$ $10^{-6} \times (\text{s}^{-1})$	$E_u$ (K)	Velocity offset <sup>a</sup> ( $\text{km s}^{-1}$ )	Relative Intensity
1-1	1900.136(10)	0.77	91.2	+63.2	0.125
2-1	1900.466(2)	2.32	91.2	+11.2	0.625
1-0	1900.950(15)	1.55	91.2	-65.2	0.250

**Note.**

<sup>a</sup> The velocity offsets are computed with respect to the [C II] line at 1900.536 GHz.

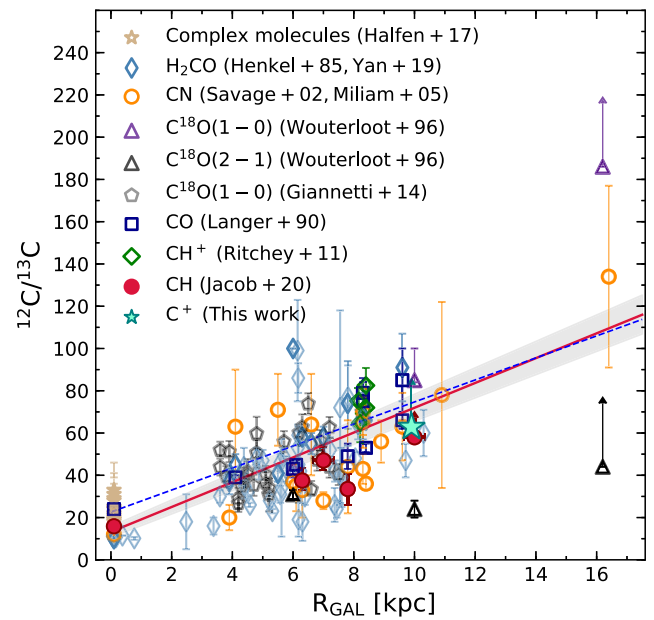
With a frequency separation  $< 1$  GHz (equivalent to a velocity separation of  $130 \text{ km s}^{-1}$ ; see Table 7) between its hyperfine-structure components, the  $^{13}\text{C II}$  line is detected simultaneously with [C II] in the 4 GHz wide band of the LFA channel on upGREAT. The  $^{13}\text{C II}$  signatures were first detected using the Kuiper Airborne Observatory (KAO) by Boreiko et al. (1988), Stacey et al. (1991), Boreiko & Betz (1996), and then by Ossenkopf et al. (2013) using Herschel/HIFI, for example, and more recently by Graf et al. (2012) and Guevara et al. (2020) using LFA/upGREAT. While the strongest  $^{13}\text{C II}$  hyperfine component corresponding to the



**Figure 16.** Demonstration of the XCLASS fits to both the [C II] and [ $^{13}\text{C}$  II] features toward W3 IRS5. Top: the simultaneous fit in red overlaid on the spectrum in black. Middle: a zoom-in on the spectrum and fits, where the blue shaded curves highlight the XCLASS fit to the [ $^{13}\text{C}$  II] hyperfine-structure lines. Bottom: individual [C II] foreground absorption components in gray with the self-absorption feature seen near the systemic velocity of the source highlighted in red.

$F=2-1$  transition blends with the broad emission wing of [C II], the  $F=1-0$  and  $F=1-1$  transitions are detected toward W3 IRS5 and NGC 7538 IRS1. The high spectral resolution provided by the upGREAT instrument allows us to study the  $^{12}\text{C}/^{13}\text{C}$  isotopic abundance ratio, an important diagnostic tool for probing Galactic chemical evolution or simply the nucleosynthesis history of the Galaxy. As a central species in gas-phase astrochemistry, the detection of  $^{13}\text{C}^+$  further presents an opportunity to study optical depth effects in  $\text{C}^+$  as well as fractionation that affects the  $^{12}\text{C}/^{13}\text{C}$  ratio. This has important ramifications on the fractionation of subsequently formed carbon bearing species such as CH and  $\text{CH}^+$ .

Being affected by self-absorption, the [C II] spectra are modeled simultaneously with that of the [ $^{13}\text{C}$  II] hyperfine lines, as shown in Figure 16. Using the  $F=1-0$  transition as a template for the line widths and intensities (which for W3 IRS5 is fit using two components; a narrow feature ( $\Delta v \sim 4.5 \text{ km s}^{-1}$ ) and a broader component at ( $\Delta v \sim 11.2 \text{ km s}^{-1}$ )), the [C II] line emission and self-absorption features at the velocity of the molecular cloud are modeled without using priors for the intensity of the  $\text{C}^+$  emission determined from measurements of the [ $^{12}\text{C}$ ]/[ $^{13}\text{C}$ ] Galactic gradient. From the resulting profiles, we derive a [ $^{12}\text{C}$ ]/[ $^{13}\text{C}$ ] ratio  $>63$ . Partly taking into account the optical depth effects seen in [C II] the [ $^{12}\text{C}$ ]/[ $^{13}\text{C}$ ] discussed above are higher than what they would be if the self-absorption feature was not accounted for. Figure 17 compares the [ $^{12}\text{C}$ ]/[ $^{13}\text{C}$ ] ratios hence derived with that obtained using other species. We find the isotopic



**Figure 17.**  $^{12}\text{C}/^{13}\text{C}$  isotopic abundance ratio as a function of Galactocentric distance,  $R_{\text{GAL}}$  (kpc) adapted from Jacob et al. (2020a). The teal star represents the  $^{12}\text{C}/^{13}\text{C}$  ratio obtained using  $\text{C}^+$  (this paper). The black solid line represents the weighted fit to the data except that from this work with the gray shaded region demarcating the  $1\sigma$  interval of this fit.

abundance ratio derived using  $^{12}\text{C}^+$  and  $^{13}\text{C}^+$  to lie perfectly on the Galactic gradient derived by Jacob et al. (2020a). This result is most likely fortuitous, representing only a lower limit given that the [C II]/[ $^{13}\text{C}$  II] is affected by fractionation effects and that there are uncertainties in the excitation temperatures used. With more detections of the [ $^{13}\text{C}$ ] transitions toward the other targeted sources, this survey will extend our understanding of the optical depth and fractionation effects of  $\text{C}^+$ .

## 6. Summary

In this paper, we introduced the SOFIA Legacy program HyGAL aimed at characterizing the Galactic ISM through absorption-line studies of six key hydrides ( $\text{ArH}^+$ ,  $p\text{-H}_2\text{O}^+$ ,  $\text{OH}^+$ , SH, OH, and CH) and two atomic species  $\text{C}^+$  and O toward 25 bright Galactic background continuum sources. Concerned with the life cycle of molecular material in the universe, the HyGAL program addresses questions related to the formation of molecular clouds and investigates the processes responsible for the phase transition from atomic to molecular gas. This is achieved by making use of the well-established diagnostic powers of the different targeted hydrides, which range from tracers of the atomic gas and cosmic-ray ionization rates such as  $\text{ArH}^+$ ,  $p\text{-H}_2\text{O}^+$ , and molecular and CO-dark  $\text{H}_2$  gas tracers, like CH and OH, to tracers of the dissipation of turbulence such as SH.

These goals are realized by taking advantage of the unique capabilities and high resolution provided by the upGREAT and 4GREAT instruments on board the SOFIA telescope. Here, we presented the observing strategy and data reduction techniques used as well as illustrated the potential analysis through the example of three sources in our sample- W3(OH), W3 IRS5, and NGC 7538 IRS1. For all three sources, the observed hydride spectra reveal widespread absorption not only at the velocities of the molecular cloud but also along its lines of sight, except for SH likely because of the species lower

abundances toward these sight lines. The spectra of the atomic species,  $C^+$  and  $O$ , are more complex showing a combination of emission and absorption features toward the hot core, in addition to the line-of-sight absorption features. Despite local variations in abundances, the column densities and column density ratios with respect to both atomic and molecular hydrogen derived over specific velocity intervals that correspond to spiral arm crossings are similar for all three sources. Extending this analysis to the complete HyGAL data set will provide column density distribution profiles for all the species observed, toward all sources studied to systematically investigate the properties of the diffuse gas sight lines over Galactic scales. The joint analysis of the SOFIA HyGAL data along with the HI data obtained from the JVLA, the 3 mm lines obtained from the IRAM 30 m telescope and chemical models will not only address the questions we set out to answer, but will also provide a wealth of knowledge about the background continuum sources making this survey of valuable legacy for future ISM studies.

This work is based on observations made with the NASA/DLR Stratospheric Observatory for Infrared Astronomy (SOFIA). SOFIA is jointly operated by the Universities Space Research Association, Inc. (USRA), under NASA contract NNA17BF53C, and the Deutsches SOFIA Institut (DSI) under DLR contract 50 OK 0901 to the University of Stuttgart. We are immensely grateful for the outstanding support provided by the SOFIA Operations Team and the GREAT Instrument Team for this large observing effort. A.M.J., D.A.N., and M.G.W. were generously supported by USRA through a grant for SOFIA Program 08–0038. Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. D.S. and S.W. acknowledges support of the Bonn-Cologne Graduate School, which is funded through the German Excellence Initiative as well as funding by the Deutsche Forschungsgemeinschaft (DFG) via the Collaborative Research Center SFB 956 “Conditions and Impact of Star Formation” (subprojects C5 and C6). S.W. acknowledges support by the European Research Council via ERC Starting grant No. 679852 (Radfeedback).

*Facilities:* SOFIA (upGREAT, 4GREAT), Herschel (HIFI).

*Software:* GILDAS/CLASS <https://www.iram.fr/IRAMFR/GILDAS>, matplotlib (Hunter 2007), numpy (Harris et al. 2020), and scipy (Virtanen et al. 2020), XCLASS (Möller et al. 2017).

## Appendix A

### Line-of-sight Properties of Individual Sources

Given that much of this work investigates the trends of various quantities with Galactocentric distance,  $R_{\text{GAL}}$ , the spectra toward each sight line is divided into local standard of rest (LSR) velocity intervals that correspond to absorption features arising from different spiral-arm and inter-arm crossings. Galactic rotation models are used to relate the velocities of the different line-of-sight absorption components to Galactocentric distance, following the parameters determined by Reid et al. (2019). Furthermore, for a general check on all distances, we employed the parallax-based distance calculator accessible from the website of the Bar and Spiral Structure Legacy survey (BeSSeL).<sup>24</sup>

#### A.1. HGAL284.015–00.86

HGAL284.015–00.86 (IRAS 10184–5748 or Dutra-45; Dutra et al. 2003) is a very small irregularly shaped ultracompact H II region (as per the classification by Wood & Churchwell (1989a)) located on the western edge of an optical nebula at a distance of 5.7 kpc on the Carina arm. Extensive observations of the CS  $J=2-1$  transition (Bronfman et al. 1996) toward this region reveals HGAL284.015–00.86 to be a promising region which can be used to investigate mass segregation effects at the birth sites of massive stars. With only a handful of molecular line surveys being carried out toward this region, the velocity range covered by the line-of-sight absorption is unknown.

#### A.2. HGAL285.26–00.05

HGAL285.26–00.05 is an embedded infrared cluster 1.5 wide, commonly referred to as Dutra, Bica, Soares, Barbuy 48—DBSB 48 (Dutra et al. 2003), located in the Hoffleit 18 nebula discovered by Hoffleit (1953). Prior to its detection at infrared wavelengths using the Two Micron All Sky Survey (2MASS), this nebula was identified as an H II region using radio observations by Caswell & Haynes (1987). Located roughly at the center of the Sagittarius–Carina arm at a distance of 4.3 kpc in the fourth quadrant of the Galaxy, the line of sight toward this source does not cross any other spiral arm.

#### A.3. G291.579–00.431

NGC 3603 is a massive OB cluster and one of the most luminous H II regions in the Galaxy (100 times more luminous than the Trapezium cluster in Orion; Goss & Radhakrishnan 1969). Associated with NGC 3603, G291.579–00.431 is located in the Sagittarius–Carina spiral arm at a distance of approximately 8 kpc (Lee et al. 2012), corresponding to one the strongest masing sources of the region. Lying tangential to the Carina arm this source only covers a short sight line with absorption features typically observed between  $-2$  and  $23 \text{ km s}^{-1}$ .

#### A.4. IRAS 12326–6245

The infrared source IRAS 12326–6245, with a  $v_{\text{LSR}} \sim -40 \text{ km s}^{-1}$ , located tangential to the Scutum–Centaurus arm, located 7.2 kpc from the Galactic center (Green & McClure-Griffiths 2011). Deeply embedded in a dense molecular cloud, IRAS 12326–6245 harbors in its center one of the most energetic and massive bipolar molecular outflows known among objects of similar luminosity (Henning et al. 2000). The line of sight crosses the Sagittarius–Carina arm near  $-22 \text{ km s}^{-1}$ .

#### A.5. G327.3–00.60

Located in the Scutum–Centaurus–Crux arm, the G327.3–00.60 star-forming region is at a distance of 3.1 kpc (Wiener et al. 2015) and hosts different evolutionary phases of massive star formation consisting of a bright H II region and a chemically rich hot core in dust condensates within a  $\sim 3 \text{ pc}$  region (Wyrowski et al. 2006). The sight line toward this target shows absorption against the molecular cloud at a  $v_{\text{LSR}} = -47 \text{ km s}^{-1}$  as well as absorption near  $-10 \text{ km s}^{-1}$  coinciding with the crossing of the Sagittarius–Carina arm.

<sup>24</sup> See <http://www.vlbi-astrometry.org/BeSSeL/node/378>.

### A.6. G328.307+00.423

The G328.307+00.423 cluster harbors massive young stellar objects that are highly concentrated within the central region, with intermediate-mass young stellar objects more spread out. Situated in the fourth quadrant, G328.307+00.423 is a bright ATLASGAL (Schuller et al. 2009) 870  $\mu\text{m}$  source at a distance of 5.8 kpc (Urquhart et al. 2018) and located in the Norma spiral arm. The sight line toward this source contains contributions from the molecular cloud and the near side of the Norma arm between  $-110$  and  $-75 \text{ km s}^{-1}$ , the Scutum–Centaurus arm between  $-62$  and  $-50 \text{ km s}^{-1}$  and the Sagittarius arm between  $-45$  and  $-35 \text{ km s}^{-1}$ . In addition to these features, certain molecular species show a narrow absorption feature centered at  $-15 \text{ km s}^{-1}$ , likely associated with local gas.

### A.7. IRAS 16060–5146

Located at a distance of 5.3 kpc (Urquhart et al. 2018), IRAS 16060–5146 (G330.954–00.182) is part of the G331.394–00.125 complex. It is a luminous star-forming region that lies close to G328.307+00.423 in the Norma spiral arm. High angular resolution continuum observations toward this region shows the presence of a compact H II region with a complex and irregular morphology (Walsh et al. 1998). With its peak position coincident with a dust core, the H II region is extremely embedded and therefore undetectable at near-infrared wavelengths. It has an absorption profile that extends from  $-130$  to  $+20 \text{ km s}^{-1}$ . The gas between  $-60$  and  $+20 \text{ km s}^{-1}$  traces the near and far-side crossings of both the Scutum–Centaurus as well as the Sagittarius spiral arms.

### A.8. IRAS 16164–5046

The sight line toward IRAS 16164–5046 (G332.826–00.549), which has a near kinematic distance of 3.6 kpc (Moisés et al. 2011), primarily traces the Scutum–Centaurus and Sagittarius spiral arms. The strongest absorption features are near  $-40 \text{ km s}^{-1}$  and are often observed alongside redshifted absorption that traces infalling material, similar to that seen in other sources like G19.61–0.23.

### A.9. IRAS 16352–4721

At a distance of  $\sim 12.3$  kpc (Green & McClure-Griffiths 2011), IRAS 16352–4721 (G337.704–00.054) is an embedded young stellar object, and one of the most distant Galactic sources in our sample. Absorption features cover a range of velocities from  $-135$  to  $\sim +20 \text{ km s}^{-1}$ . The line-of-sight absorption at negative velocities between  $-140$  and  $-60 \text{ km s}^{-1}$  arises mainly from the 3 kpc and the Norma spiral arms with relatively weaker contributions from the other arms as well.

### A.10. IRAS 16547–4247

IRAS 16547–4257 (G343.12–0.06) at 2.7 kpc (Giannetti et al. 2014) on the Scutum arm has the distinction of being the first reported, and most luminous young stellar object known to harbor a radio jet. The jet, located at the center of the molecular core, drives a highly energetic and collimated bipolar outflow over 1.5 pc (Garay et al. 2003). In addition to the features associated with the molecular cloud, itself at  $v_{\text{LSR}}$  between  $-50$  and  $-30 \text{ km s}^{-1}$ , the line of sight toward this source is also known to show absorption from  $-25$  to  $-15 \text{ km s}^{-1}$  associated with the Sagittarius arm and the Local Arm near  $0 \text{ km s}^{-1}$ .

### A.11. NGC 6334I

NGC 6334 is a filamentary, star-forming cloud on the Sagittarius spiral arm at a distance of 1.35 kpc (Immer et al. 2013). It harbors a string of dense cores at far-infrared wavelengths that are identified using roman numerals from I to VI (McBreen et al. 1979). Source I (also designated as NGC 6334 F) is a prototypical hot molecular core at the head of a cometary ultracompact H II region with a rich molecular inventory. An additional source was identified  $\sim 2'$  north of source I, named “I(N)” (Gezari 1982). While both cores are sites of active star formation, NGC 6334 I is in a relatively more evolved stage of star formation than NGC 6334 I(N). Absorption components observed in the velocity range from  $-17$  to  $3 \text{ km s}^{-1}$  can be ascribed to the absorption against the core itself and foreground components associated with the NGC 6334 molecular cloud. In addition, this source’s line of sight covers foreground absorption arising from two unrelated clouds between 5 and  $10 \text{ km s}^{-1}$  (van der Wiel et al. 2016).

### A.12. G357.558–00.321

G357.558–00.321 is an H II region located in the vicinity of an unusual supernova remnant, the “Tornado nebula”, G357.7–0.1 (MSH 17–39) in the central 1 kpc of the Galaxy (Frail et al. 1996). Spatially, G357.7–0.1 consists of a bright head and a faint tail, as well as a compact source  $30''$  north of the head (known as the eye and not associated with the remnant; Sakai et al. 2014). Owing to difficulties in determining meaningful kinematic distances toward this direction using HI absorption measurements, the distance adopted toward this region has been determined from a single maser measurement near  $v_{\text{LSR}} = -12.4 \text{ km s}^{-1}$ , detected along the western edge of MSH 17–39, just north of the compact source. However, very little is known about G357.558–00.321 itself.

### A.13. HGAL0.55–0.85

The HGAL0.55–0.85 molecular cloud region is located in the Galactic center region (Walsh et al. 1998). A site of active star formation, this region hosts two Class II methanol masers separated by  $\sim 3''$  with associated  $\text{NH}_3$  emission (Hill et al. 2005; Beuther et al. 2009). However, as it has been the target of only a handful of astronomical studies, very little else is known about this region.

### A.14. G09.62+0.19

The G09.62+0.19 complex, at a distance of 5.2 kpc (Sanna et al. 2009), is located in the Norma arm close to the expanding 3 kpc arm and contains a cluster of radio continuum sources, denoted from A to I. Since each of these continuum sources is in a different evolutionary stage, the overall morphology of the complex is suggestive of sequential high-mass star formation progressing from the most evolved H II region, component A—hosting massive star formation—to regions like component E that show signs of a very early phase of massive star formation (Hofner et al. 1994; Liu et al. 2017). Beyond the systemic velocity of the source  $v_{\text{LSR}} = 4.3 \text{ km s}^{-1}$ , the source also shows weaker line-of-sight absorption against the Scutum–Centaurus and Sagittarius spiral arms between 10 and  $40 \text{ km s}^{-1}$ .

## A.15. G10.47 + 0.03

The G10.47 + 0.03 H II region contains three ultracompact H II regions, G10.47 + 0.03 A, B and C, and is located at a distance of 8.55 kpc (Sanna et al. 2014). The sight line toward G10.47 + 0.03 therefore crosses several spiral arms in the inner Galaxy, resulting in a broad absorption spectrum covering a  $v_{\text{LSR}}$  range from  $-35$  to  $180 \text{ km s}^{-1}$ . Absorption at  $v_{\text{LSR}}$  between  $-35$  and  $48 \text{ km s}^{-1}$  arises partly from inter-arm gas and partly from the near-side crossing of the Sagittarius and Scutum–Centaurus arms. Features at  $v_{\text{LSR}} > 80 \text{ km s}^{-1}$  are associated with the 3 kpc arm and the Galactic bar with contributions from the Norma arm at intermediate velocities ( $40 \text{ km s}^{-1} < v_{\text{LSR}} < 80 \text{ km s}^{-1}$ ). Higher velocity components at  $v_{\text{LSR}} > 120 \text{ km s}^{-1}$  likely trace gas that lies beyond the Galactic center belonging to the  $135 \text{ km s}^{-1}$  arm (Sormani & Magorrian 2015).

## A.16. G19.61–0.23

G19.61–0.23 is a high-mass star-forming region, known to harbor several ultracompact H II regions and a hot molecular cloud (Furuya et al. 2011). Rich in young stellar objects, G19.61–0.23 has been the target of several studies aimed at understanding cluster formation and other processes at the early phases of the high-mass star formation. This massive star-forming clump is at a heliocentric distance of 12.7 kpc (Urquhart et al. 2014a) with an absorption spectrum extending from  $-10$  to  $150 \text{ km s}^{-1}$ . The absorption at  $v_{\text{LSR}}$  between  $-10$  and  $20 \text{ km s}^{-1}$  primarily traces inter-arm gas whereas absorption dips between  $35$  and  $74 \text{ km s}^{-1}$  arise from the near- and far-side crossings of the Sagittarius spiral arm, where the systemic velocity of the source is at  $40 \text{ km s}^{-1}$ . Signatures of infalling gas have previously been studied by Furuya et al. (2011), whose observations of the  $J = 3 \rightarrow 2$  transitions of  $^{13}\text{CO}$  and  $^{18}\text{CO}$  show inverse P-Cygni profiles or redshifted absorption alongside a blueshifted emission component. Narrow absorption features at  $80$  and  $113 \text{ km s}^{-1}$  possibly arise from the edge of the Sagittarius–Carina and Scutum–Centaurus spiral arms, respectively.

## A.17. G29.96–0.02

One of the brightest submillimeter continuum sources in the Milky Way galaxy, G29.96–0.02 is a classical example of a compact H II region with a cometary-like morphology (Wood & Churchwell 1989b) caused by its interaction with a molecular core  $\sim 2.''6$  away. While the main exciting source of the compact H II region has been identified in the near-infrared as an O5–O6 star (Watson & Hanson 1997), additional sources have been detected toward the rim of the H II region, indicating an embedded cluster (Pratap et al. 1999). The rich molecular inventory of this region has not only been revealed through the characterization of strong emission lines but also through absorption-line measurements of species such as HF (Kirk et al. 2010) and ortho- $\text{H}_2\text{Cl}^+$ . At a distance of 6.7 kpc (Urquhart et al. 2018), G29.96–0.02 lies along the Galactic bar and on the Norma arm. In addition to absorption at molecular cloud velocities between  $87$  and  $120 \text{ km s}^{-1}$ , the sight line velocity distribution also includes absorption arising from the Sagittarius near- and far-side between  $37$  and  $85 \text{ km s}^{-1}$ , and the Aquila Rift near  $9 \text{ km s}^{-1}$ .

## A.18. G31.41+0.31

A prototypical hot molecular core, G31.41+0.31 is located on the Scutum–Centaurus arm at a distance of 5 kpc (Moisés et al. 2011), with the hot core separated by  $\sim 5''$  from an ultracompact H II region. The high bolometric luminosity ( $\approx 3 \times 10^5 L_{\odot}$ ; Cesaroni et al. (2011) and references therein) and extended halo hints toward the presence of a disk-like structure rotating about young stellar objects. The presence of such a rotating toroid, together with the possibility of outflows from the center of the molecular core complicates the analysis of spectral line features toward this cloud component at  $v_{\text{LSR}} = 98 \text{ km s}^{-1}$ , often distorting the observed line profiles or showing self-absorption (Colzi et al. 2017). The line of sight toward G31.41+0.31 crosses the Sagittarius arm at velocities roughly between  $30$  and  $55 \text{ km s}^{-1}$  and the Local Arm at velocities  $< 20 \text{ km s}^{-1}$ .

## A.19. W43 MM1

W43 MM1 is a high-mass protostellar object located at a distance of 5.5 kpc (Zhang et al. 2014) near the end of the Galactic bar, in the mini-starburst region of W43. While it resembles a young protostar in its main accretion phase, W43 MM1 has already developed a hot core with temperatures higher than 200 K. In spite of its many peculiarities, several absorption studies have been previously carried out toward this source; for example, it was found to be rich in water and its isotopologues in the WISH program (van Dishoeck et al. 2011; Herpin et al. 2012). Located in a crowded part of the Galactic plane, molecular line data toward this source exhibits absorption features between  $0 \text{ km s}^{-1}$  and  $140 \text{ km s}^{-1}$  with extended absorption near  $v_{\text{LSR}} = 98 \text{ km s}^{-1}$  associated with absorption against the continuum background source and narrower features arising from line-of-sight components at  $v_{\text{LSR}} < 85 \text{ km s}^{-1}$ . This includes components near  $40 \text{ km s}^{-1}$ ,  $10 \text{ km s}^{-1}$  and  $0 \text{ km s}^{-1}$  corresponding to the near-side crossing of the Sagittarius arm, the Aquila Rift, and the Local Arm, respectively.

## A.20. G32.80+0.19

The G32.80+0.19 H II region has a complex structure with at least three associated components; a bright compact source labeled A, an extended component with a cometary-like morphology located  $6''$  northeast of A called B, and a weaker component to the northwest which is unresolved in continuum maps at 8.3 GHz, dubbed C (Garay et al. 1994; Gomez et al. 1995). Situated in the Perseus arm 13 kpc away, the sight line toward this H II region consists of several spiral arm components spanning a range of velocities from  $0$  to  $110 \text{ km s}^{-1}$ . Beyond the systemic velocity of the source at  $\sim 15 \text{ km s}^{-1}$ , the line of sight crosses the near- and far-side of the Sagittarius arm between  $25$  and  $65 \text{ km s}^{-1}$ , showing narrow absorption at  $71 \text{ km s}^{-1}$  corresponding to absorption against the Aquila rift and finally the Scutum–Centaurus arm at  $v_{\text{LSR}} > 78 \text{ km s}^{-1}$ .

## A.21. G45.07+0.13





At a distance of 4.3 kpc, G45.07+0.13 is a pair of spherical ultracompact H II regions showing OH,  $\text{H}_2\text{O}$ , and  $\text{CH}_3\text{OH}$  maser emission located in the Sagittarius–Carina spiral arm with at least three continuum sources observed in the

mid-infrared (De Buizer et al. 2005). Furthermore, observational evidence distinguishes the G45.07+0.13 core from neighboring cores such as G45.12 + 0.13 as a relatively young site of massive star formation (Hunter et al. 1997). In addition to absorption against the background continuum source at  $v_{\text{LSR}} = 60 \text{ km s}^{-1}$  and the Local Arm near  $0 \text{ km s}^{-1}$ , some species show narrow absorption components between  $v_{\text{LSR}} = 20$  and  $60 \text{ km s}^{-1}$  tangentially along the Sagittarius arm.

### A.22. DR21

Located in the Cygnus X complex at a distance of 1.5 kpc (Rygl et al. 2012), the DR21 molecular ridge is a dense region of massive star formation at  $v_{\text{LSR}} = -3 \text{ km s}^{-1}$ . This star-forming region comprises of a group of several compact H II regions and is known to host highly energetic molecular outflows. This outflow is likely driven by an atomic jet as evidenced by the detection of high-velocity H I emission (up to  $90 \text{ km s}^{-1}$ ) using the Very Large Array (Russell et al. 1992). In addition to the molecular cloud associated with the DR21 region itself, the filament contains another center of activity corresponding to the OH maser source, DR21(OH), which lies about  $3'$  (or a projected distance of 1.3 pc) north of DR21. Molecular line observations reveal a prominent blueshifted outflow wing at velocities  $< -17 \text{ km s}^{-1}$ , while the corresponding redshifted wing is more difficult to discern (Schneider et al. 2010). Furthermore, in addition to the foreground absorption arising from the Cygnus X complex, there is a foreground component at  $v_{\text{LSR}} \sim 9 \text{ km s}^{-1}$  associated with W75 N star-forming region located northwest of the filament.

### ORCID iDs

A. M. Jacob  <https://orcid.org/0000-0001-7838-3425>  
 D. A. Neufeld  <https://orcid.org/0000-0001-8341-1646>  
 P. Schilke  <https://orcid.org/0000-0003-2141-5689>  
 H. Wiesemeyer  <https://orcid.org/0000-0002-5135-8657>  
 W.-J. Kim  <https://orcid.org/0000-0003-0364-6715>  
 S. Bialy  <https://orcid.org/0000-0002-0404-003X>  
 M. Busch  <https://orcid.org/0000-0003-4961-6511>  
 D. Elia  <https://orcid.org/0000-0002-9120-5890>  
 M. Gerin  <https://orcid.org/0000-0002-2418-7952>  
 R. Higgins  <https://orcid.org/0000-0001-8195-3900>  
 N. Indriolo  <https://orcid.org/0000-0001-8533-6440>  
 D. C. Lis  <https://orcid.org/0000-0002-0500-4700>  
 K. M. Menten  <https://orcid.org/0000-0001-6459-0669>  
 A. Sanchez-Monge  <https://orcid.org/0000-0002-3078-9482>  
 T. Möller  <https://orcid.org/0000-0002-9277-8025>  
 V. Ossenkopf-Okada  <https://orcid.org/0000-0002-8351-3877>  
 D. Seifried  <https://orcid.org/0000-0002-0368-9160>  
 P. Sonnentrucker  <https://orcid.org/0000-0003-2027-5020>  
 S. Walch  <https://orcid.org/0000-0001-6941-7638>  
 M. G. Wolfire  <https://orcid.org/0000-0003-0030-9510>

### References

Allen, R. J., Hogg, D. E., & Engelke, P. D. 2015, *AJ*, 149, 123  
 Belloche, A., Müller, H. S. P., Menten, K. M., Schilke, P., & Comito, C. 2013, *A&A*, 559, A47  
 Bellomi, E., Godard, B., Hennebelle, P., et al. 2020, *A&A*, 643, A36  
 Belov, S. P., Winnewisser, G., & Herbst, E. 1995, *JMoSp*, 174, 253  
 Beuther, H., Linz, H., & Henning, T. 2013, *A&A*, 558, A81  
 Beuther, H., Walsh, A. J., & Longmore, S. N. 2009, *ApJS*, 184, 366

Bialy, S., Belli, S., & Padovani, M. 2022, *A&A*, 658, L13  
 Bialy, S., & Burkhart, B. 2020, *ApJL*, 894, L2  
 Bialy, S., Burkhart, B., & Sternberg, A. 2017, *ApJ*, 843, 92  
 Bialy, S., Neufeld, D., Wolfire, M., Sternberg, A., & Burkhart, B. 2019, *ApJ*, 885, 109  
 Bialy, S., & Sternberg, A. 2016, *ApJ*, 822, 83  
 Boreiko, R. T., & Betz, A. L. 1996, *ApJL*, 467, L113  
 Boreiko, R. T., Betz, A. L., & Zmuidzinas, J. 1988, *ApJL*, 325, L47  
 Bronfman, L., Nyman, L. A., & May, J. 1996, *A&AS*, 115, 81  
 Burgh, E. B., France, K., & McCandliss, S. R. 2007, *ApJ*, 658, 446  
 Burton, M. G., Hollenbach, D. J., & Tielens, A. G. G. 1992, *ApJ*, 399, 563  
 Busch, M. P., Allen, R. J., Engelke, P. D., et al. 2019, *ApJ*, 883, 158  
 Busch, M. P., Engelke, P. D., Allen, R. J., & Hogg, D. E. 2021, *ApJ*, 914, 72  
 Butner, H. M., Charnley, S. B., Ceccarelli, C., et al. 2007, *ApJL*, 659, L137  
 Caselli, P., Walmsley, C. M., Terziewa, R., & Herbst, E. 1998, *ApJ*, 499, 234  
 Caswell, J. L., & Haynes, R. F. 1987, *AuJPh*, 40, 215  
 Caswell, J. L., & Vaile, R. A. 1995, *MNRAS*, 273, 328  
 Cesaroni, R., Beltrán, M. T., Zhang, Q., Beuther, H., & Fallscheer, C. 2011, *A&A*, 533, A73  
 Colzi, L., Fontani, F., Rivilla, V. M., et al. 2017, *MmSAI*, 88, 728  
 Cummings, A. C., Stone, E. C., Heikkilä, B. C., et al. 2016, *ApJ*, 831, 18  
 De Buizer, J. M., Radoski, J. T., Telesco, C. M., & Piña, R. K. 2005, *ApJS*, 156, 179  
 de Graauw, T., Helmich, F. P., Phillips, T. G., et al. 2010, *A&A*, 518, L6  
 Dunham, T. J. 1937, *PASP*, 49, 26  
 Durán, C. A., Güsten, R., Risacher, C., et al. 2021, *ITTTST*, 11, 194  
 Dutra, C. M., Bica, E., Soares, J., & Barbuy, B. 2003, *A&A*, 400, 533  
 Eddington, A. S. 1926, *RSPSA*, 111, 424  
 Elia, D., Merello, M., Molinari, S., et al. 2021, *MNRAS*, 504, 2742  
 Emprechtinger, M., Lis, D. C., Rolfs, R., et al. 2013, *ApJ*, 765, 61  
 Engelke, P. D., & Allen, R. J. 2018, *ApJ*, 858, 57  
 Falgarone, E., Godard, B., Cernicharo, J., et al. 2010, *A&A*, 521, L15  
 Federman, S. R. 1982, *ApJ*, 257, 125  
 Federman, S. R., Glassgold, A. E., & Kwan, J. 1979, *ApJ*, 227, 466  
 Flagey, N., Goldsmith, P. F., Lis, D. C., et al. 2013, *ApJ*, 762, 11  
 Frail, D. A., Goss, W. M., Reynoso, E. M., et al. 1996, *AJ*, 111, 1651  
 Furuya, R. S., Cesaroni, R., & Shinnaga, H. 2011, *A&A*, 525, A72  
 Garay, G., Brooks, K. J., Mardones, D., & Norris, R. P. 2003, *ApJ*, 587, 739  
 Garay, G., Lizano, S., & Gomez, Y. 1994, *ApJ*, 429, 268  
 Gerin, M., Neufeld, D. A., & Goicoechea, J. R. 2016, *ARA&A*, 54, 181  
 Gerin, M., de Luca, M., Black, J., et al. 2010a, *A&A*, 518, L110  
 Gerin, M., de Luca, M., Goicoechea, J. R., et al. 2010b, *A&A*, 521, L16  
 Gerin, M., Ruaud, M., Goicoechea, J. R., et al. 2015, *A&A*, 573, A30  
 Gezari, D. Y. 1982, *ApJL*, 259, L29  
 Giannetti, A., Wyrowski, F., Brand, J., et al. 2014, *A&A*, 570, A65  
 Godard, B., Falgarone, E., & Pineau Des Forêts, G. 2009, *A&A*, 495, 847  
 Godard, B., Falgarone, E., & Pineau des Forêts, G. 2014, *A&A*, 570, A27  
 Gomez, Y., Garay, G., & Lizano, S. 1995, *ApJ*, 453, 727  
 Goss, W. M. 1968, *ApJS*, 15, 131  
 Goss, W. M., & Radhakrishnan, V. 1969, *Astrophys. Lett.*, 4, 199  
 Graf, U. U., Simon, R., Stutzki, J., et al. 2012, *A&A*, 542, L16  
 Green, J. A., & McClure-Griffiths, N. M. 2011, *MNRAS*, 417, 2500  
 Grenier, I. A., Casandjian, J.-M., & Terrier, R. 2005, *Sci*, 307, 1292  
 Guan, X., Stutzki, J., Graf, U. U., et al. 2012, *A&A*, 542, L4  
 Guevara, C., Stutzki, J., Ossenkopf-Okada, V., et al. 2020, *A&A*, 636, A16  
 Gupta, H., Rimmer, P., Pearson, J. C., et al. 2010, *A&A*, 521, L47  
 Güsten, R., Nyman, L. Å., Schilke, P., et al. 2006, *A&A*, 454, L13  
 Habart, E., Abergel, A., Boulanger, F., et al. 2011, *A&A*, 527, A122  
 Hachisuka, K., Brunthaler, A., Menten, K. M., et al. 2006, *ApJ*, 645, 337  
 Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Natur*, 585, 357  
 Henning, T., Lapinov, A., Schreyer, K., Stecklum, B., & Zinchenko, I. 2000, *A&A*, 364, 613  
 Herpin, F., Chavarría, L., van der Tak, F., et al. 2012, *A&A*, 542, A76  
 Hill, T., Burton, M. G., Minier, V., et al. 2005, *MNRAS*, 363, 405  
 Hoffleit, D. 1953, *AnHar*, 0119, 37  
 Hofner, P., Kurtz, S., Churchwell, E., Walmsley, C. M., & Cesaroni, R. 1994, *ApJL*, 429, L85  
 Hunter, J. D. 2007, *CSE*, 9, 90  
 Hunter, T. R., Phillips, T. G., & Menten, K. M. 1997, *ApJ*, 478, 283  
 Immer, K., Reid, M. J., Menten, K. M., Brunthaler, A., & Dame, T. M. 2013, *A&A*, 553, A117  
 Indriolo, N., Geballe, T. R., Oka, T., & McCall, B. J. 2007, *ApJ*, 671, 1736  
 Indriolo, N., & McCall, B. J. 2012, *ApJ*, 745, 91  
 Indriolo, N., Neufeld, D. A., Gerin, M., et al. 2012, *ApJ*, 758, 83  
 Indriolo, N., Neufeld, D. A., Gerin, M., et al. 2015, *ApJ*, 800, 40



- Jacob, A. M., Menten, K. M., Wiesemeyer, H., et al. 2020a, *A&A*, 640, A125
- Jacob, A. M., Menten, K. M., Wiesemeyer, H., et al. 2019, *A&A*, 632, A60
- Jacob, A. M., Menten, K. M., Wyrowski, F., Winkel, B., & Neufeld, D. A. 2020b, *A&A*, 643, A91
- Kennicutt, R. C., & Evans, N. J. 2012, *ARA&A*, 50, 531
- Kirk, J. M., Polehampton, E., Anderson, L. D., et al. 2010, *A&A*, 518, L82
- Klaassen, P. D., Wilson, C. D., Keto, E. R., & Zhang, Q. 2009, *ApJ*, 703, 1308
- Klein, B., Hochgürtel, S., Krämer, I., et al. 2012, *A&A*, 542, L3
- Kolpak, M. A., Jackson, J. M., Bania, T. M., Clemens, D. P., & Dickey, J. M. 2003, *ApJ*, 582, 756
- Korotin, S. A., Andrievsky, S. M., Luck, R. E., et al. 2014, *MNRAS*, 444, 3301
- Krumholz, M. R., & Federrath, C. 2019, *FrASS*, 6, 7
- Langer, W. D., Velusamy, T., Pineda, J. L., Willacy, K., & Goldsmith, P. F. 2014, *A&A*, 561, A122
- Lattanzi, V., Walters, A., Drouin, B. J., & Pearson, J. C. 2007, *ApJ*, 662, 771
- Lebrón, M. E., Rodríguez, L. F., & Lizano, S. 2001, *ApJ*, 560, 806
- Lee, E. J., Murray, N., & Rahman, M. 2012, *ApJ*, 752, 146
- Liszt, H., Gerin, M., & Grenier, I. 2019, *A&A*, 627, A95
- Liszt, H., & Lucas, R. 2004, *A&A*, 428, 445
- Liszt, H. S., & Lucas, R. 1998, *A&A*, 339, 561
- Liu, T., Lacy, J., Li, P. S., et al. 2017, *ApJ*, 849, 25
- Liu, T., Evans, N. J., Kim, K.-T., et al. 2020, *MNRAS*, 496, 2790
- Lucas, R., & Liszt, H. 1996, *A&A*, 307, 237
- Mac Low, M.-M., & Klessen, R. S. 2004, *RvMP*, 76, 125
- Matsushima, F., Matsunaga, M., Qian, G.-Y., et al. 2001, *JMoSp*, 206, 41
- McBreen, B., Fazio, G. G., Stier, M., & Wright, E. L. 1979, *ApJL*, 232, L183
- McCall, B. J., Huneycutt, A. J., Saykally, R. J., et al. 2003, *Natur*, 422, 500
- McGuire, B. A. 2022, *ApJS*, 259, 30
- McKee, C. F., & Ostriker, E. C. 2007, *ARA&A*, 45, 565
- McKellar, A. 1940, *PASP*, 52, 187
- Mège, P., Russeil, D., Zavagno, A., et al. 2021, *A&A*, 646, A74
- Megeath, S. T., Herter, T., Beichman, C., et al. 1996, *A&A*, 307, 775
- Megeath, S. T., Townsley, L. K., Oey, M. S., & Tieftrunk, A. R. 2008, *Handbook of Star Forming Regions*, Vol. 1 (San Francisco, CA: ASP), 264
- Menten, K. M., Wyrowski, F., Belloche, A., et al. 2011, *A&A*, 525, A77
- Moisés, A. P., Damineli, A., Figuerêdo, E., et al. 2011, *MNRAS*, 411, 705
- Molinari, S., Swinyard, B., Bally, J., et al. 2010, *PASP*, 122, 314
- Möller, T., Bernst, I., Panoglou, D., et al. 2013, *A&A*, 549, A21
- Möller, T., Endres, C., & Schilke, P. 2017, *A&A*, 598, A7
- Mookerjee, B. 2016, *MNRAS*, 459, 2822
- Moscadelli, L., Reid, M. J., Menten, K. M., et al. 2009, *ApJ*, 693, 406
- Moseley, E. R., Draine, B. T., Tomida, K., & Stone, J. M. 2021, *MNRAS*, 500, 3290
- Müller, H. S. P., Schlöder, F., Stutzki, J., & Winnewisser, G. 2005, *JMoSt*, 742, 215
- Muller, S., Müller, H. S. P., Black, J. H., et al. 2016, *A&A*, 595, A128
- Myers, A. T., McKee, C. F., & Li, P. S. 2015, *MNRAS*, 453, 2747
- Navarete, F., Galli, P. A. B., & Damineli, A. 2019, *MNRAS*, 487, 2771
- Neufeld, D. A., & Wolfire, M. G. 2016, *ApJ*, 826, 183
- Neufeld, D. A., & Wolfire, M. G. 2017, *ApJ*, 845, 163
- Neufeld, D. A., Goicoechea, J. R., Sonnentrucker, P., et al. 2010a, *A&A*, 521, L10
- Neufeld, D. A., Sonnentrucker, P., Phillips, T. G., et al. 2010b, *A&A*, 518, L108
- Neufeld, D. A., Falgarone, E., Gerin, M., et al. 2012, *A&A*, 542, L6
- Neufeld, D. A., Godard, B., Gerin, M., et al. 2015, *A&A*, 577, A49
- Ossenkopf, V., Röllig, M., Neufeld, D. A., et al. 2013, *A&A*, 550, A57
- Ossenkopf, V., Röllig, M., Simon, R., et al. 2010, *A&A*, 518, L79
- Padoan, P., Federrath, C., Chabrier, G., et al. 2014, in *Protostars and Planets VI*, ed. H. Beuther et al. (Tucson, AZ: Univ. Arizona Press), 77
- Padovani, M., Galli, D., & Glassgold, A. E. 2009, *A&A*, 501, 619
- Padovani, M., Ivlev, A. V., Galli, D., et al. 2020, *SSRv*, 216, 29
- Pan, K., Federman, S. R., Sheffer, Y., & Andersson, B. G. 2005, *ApJ*, 633, 986
- Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, *A&A*, 518, L1
- Pineda, J. L., Langer, W. D., Velusamy, T., & Goldsmith, P. F. 2013, *A&A*, 554, A103
- Pratap, P., Megeath, S. T., & Bergin, E. A. 1999, *ApJ*, 517, 799
- Qin, S.-L., Schilke, P., Wu, J., et al. 2016, *MNRAS*, 456, 2681
- Rachford, B. L., Snow, T. P., Tumlinson, J., et al. 2002, *ApJ*, 577, 221
- Rachford, B. L., Snow, T. P., Destree, J. D., et al. 2009, *ApJS*, 180, 125
- Reid, M. J., Brunthaler, A., Menten, K. M., et al. 2017, *AJ*, 154, 63
- Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2019, *ApJ*, 885, 131
- Risacher, C., Güsten, R., Stutzki, J., et al. 2018, *JAI*, 7, 1840014
- Rosen, A. L., Offner, S. S. R., Sadavoy, S. I., et al. 2020, *SSRv*, 216, 62
- Rugel, M. R., Beuther, H., Bühr, S., et al. 2018, *A&A*, 618, A159
- Russell, A. P. G., Bally, J., Padman, R., & Hills, R. E. 1992, *ApJ*, 387, 219
- Rygl, K. L. J., Brunthaler, A., Sanna, A., et al. 2012, *A&A*, 539, A79
- Sakai, D., Oka, T., Tanaka, K., et al. 2014, *ApJ*, 791, 49
- Sanna, A., Reid, M. J., Moscadelli, L., et al. 2009, *ApJ*, 706, 464
- Sanna, A., Reid, M. J., Menten, K. M., et al. 2014, *ApJ*, 781, 108
- Santangelo, G., Antonucci, S., Nisini, B., et al. 2014, *A&A*, 569, L8
- Savage, B. D., Bohlin, R. C., Drake, J. F., & Budich, W. 1977, *ApJ*, 216, 291
- Schilke, P., Comito, C., Müller, H. S. P., et al. 2010, *A&A*, 521, L11
- Schilke, P., Neufeld, D. A., Müller, H. S. P., et al. 2014, *A&A*, 566, A29
- Schneider, N., Csengeri, T., Bontemps, S., et al. 2010, *A&A*, 520, A49
- Schuller, F., Menten, K. M., Contreras, Y., et al. 2009, *A&A*, 504, 415
- Sheffer, Y., Rogers, M., Federman, S. R., et al. 2008, *ApJ*, 687, 1075
- Shull, J. M., Danforth, C. W., & Anderson, K. L. 2021, *ApJ*, 911, 55
- Snow, T. P., & McCall, B. J. 2006, *ARA&A*, 44, 367
- Sonnentrucker, P., Wolfire, M., Neufeld, D. A., et al. 2015, *ApJ*, 806, 49
- Sormani, M. C., & Magorrian, J. 2015, *MNRAS*, 446, 4186
- Stacey, G. J., Townes, C. H., Poglitsch, A., et al. 1991, *ApJL*, 382, L37
- Stead, J. J., & Hoare, M. G. 2010, *MNRAS*, 407, 923
- Sternberg, A., Le Petit, F., Roueff, E., & Le Bourlot, J. 2014, *ApJ*, 790, 10
- Swings, P., & Rosenfeld, L. 1937, *ApJ*, 86, 483
- Taylor, A. R., Gibson, S. J., Peracaula, M., et al. 2003, *AJ*, 125, 3145
- Truppe, S., Hendricks, R. J., Hinds, E. A., & Tarbutt, M. R. 2014, *ApJ*, 780, 71
- Urquhart, J. S., Figura, C. C., Moore, T. J. T., et al. 2014a, *MNRAS*, 437, 1791
- Urquhart, J. S., Moore, T. J. T., Csengeri, T., et al. 2014b, *MNRAS*, 443, 1555
- Urquhart, J. S., König, C., Giannetti, A., et al. 2018, *MNRAS*, 473, 1059
- Valdivia, V., Godard, B., Hennebelle, P., et al. 2017, *A&A*, 600, A114
- van der Wiel, M. H. D., Naylor, D. A., Makiwa, G., Satta, M., & Abergel, A. 2016, *A&A*, 593, A37
- van Dishoeck, E. F., Kristensen, L. E., Benz, A. O., et al. 2011, *PASP*, 123, 138
- Vastel, C., Ceccarelli, C., Caux, E., et al. 2010, *A&A*, 521, L31
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *NatMe*, 17, 261
- Walch, S., Girichidis, P., Naab, T., et al. 2015, *MNRAS*, 454, 238
- Walsh, A. J., Burton, M. G., Hyland, A. R., & Robinson, G. 1998, *MNRAS*, 301, 640
- Wang, K. S., Bourke, T. L., Hogerheijde, M. R., et al. 2013, *A&A*, 558, A69
- Watson, A. M., & Hanson, M. M. 1997, *ApJL*, 490, L165
- Wegg, C., Gerhard, O., & Portail, M. 2015, *MNRAS*, 450, 4050
- Weinreb, S., Barrett, A. H., Meeks, M. L., & Henry, J. C. 1963, *Natur*, 200, 829
- Welty, D. E., Sonnentrucker, P., Snow, T. P., & York, D. G. 2020, *ApJ*, 897, 36
- Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, *A&AS*, 143, 9
- Weselak, T. 2019, *A&A*, 625, A55
- Weselak, T., Galazutdinov, G., Beletsky, Y., & Krelowski, J. 2009, *A&A*, 499, 783
- Whiteoak, J. B., & Gardner, F. F. 1970, *Astrophys. Lett.*, 5, 5
- Wienen, M., Wyrowski, F., Menten, K. M., et al. 2015, *A&A*, 579, A91
- Wiesemeyer, H., Güsten, R., Heyminck, S., et al. 2016, *A&A*, 585, A76
- Wiesemeyer, H., Güsten, R., Menten, K. M., et al. 2018, *A&A*, 612, A37
- Winkel, B., Wiesemeyer, H., Menten, K. M., et al. 2017, *A&A*, 600, A2
- Wood, D. O. S., & Churchwell, E. 1989a, *ApJS*, 69, 831
- Wood, D. O. S., & Churchwell, E. 1989b, *ApJ*, 340, 265
- Wyrowski, F., Menten, K. M., Güsten, R., & Belloche, A. 2010, *A&A*, 518, A26
- Wyrowski, F., Menten, K. M., Schilke, P., et al. 2006, *A&A*, 454, L91
- Xu, Y., Reid, M. J., Zheng, X. W., & Menten, K. M. 2006, *Sci*, 311, 54
- Young, E. T., Becklin, E. E., Marcum, P. M., et al. 2012, *ApJL*, 749, L17
- Zhang, B., Moscadelli, L., Sato, M., et al. 2014, *ApJ*, 781, 89