EPJ Web of Conferences **265**, 00019 (2022) https://doi.org/10.1051/epjconf/202226500019
Multi-line Diagnostics of the Interstellar Medium

The LEGO Large Program: Constraining the Physics of Line Emission in Galaxy Observations

Jens Kauffmann[1,*] and *Ashley* Barnes[2]

[1]Haystack Observatory, Massachusetts Institute of Technology, 99 Millstone Road, Westford, MA 01886, USA
[2]Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany

Abstract. The IRAM Large Program LEGO studies molecular clouds in the Milky Way to constrain the physics controlling molecular line emission in galaxies. This is done by imaging two dozen clouds with setups that provide continuous spectral coverage of about 85–115 GHz. To give one example, research in this area permits to "calibrate" extragalactic observations of molecules like HCN and N_2H^+ to explore how the star formation activity in galaxies depends on their dense gas contents (i.e., Gao & Solomon relation, [1]). Interestingly, LEGO and other studies now reveal a substantial cloud-to-cloud variation in line ratios. Once understood properly, this diversity can be used to constrain the properties of extragalactic molecular clouds at great detail. Here we outline the LEGO sample and describe the project status.

1 Introduction

Investigations into the formation of stars over cosmic time continue to constitute one of the central research areas of astronomy. One important area in this field of study is how the star formation (SF) rate per unit gas mass depends on the properties of the molecular clouds. Gao & Solomon [1], e.g., argued in an influential paper that the SF rate of a galaxy is directly proportional to the mass of dense gas that could support star formation. That said, current modeling indicates that the outcome of the SF process sensitively depends on the physical structure of the star-forming clouds (e.g., [2]), and it is known that molecular cloud structure varies substantially inside and between galaxies (e.g., [3]). Detailed research into the connection of molecular cloud properties and SF activity thus continues to be important.

Unfortunately it is typically impossible to resolve the substructure of extragalactic clouds. To give one example, ALMA is limited to resolutions $\vartheta_{\rm beam} \gtrsim 1''$ at useful sensitivities (e.g., [4]), corresponding to spatial scales $\ell_{\rm beam} = 5~{\rm pc} \cdot (\vartheta_{\rm beam}/1'') \cdot (d/1~{\rm Mpc})$ in nearby galaxies that just marginally resolve individual molecular clouds of ~ 10 pc. Research into extragalactic SF therefore routinely relies on indirect means to characterize cloud substructure. Molecular emission lines at ~ 100 GHz frequency are particularly useful for such assessments. To give examples, emission from the ^{12}CO molecule is generally used to sense the total molecular gas reservoir, while it has been reasoned that the luminosity in the HCN (1–0) line is a probe of the enclosed mass of dense gas (i.e., at $\gg 10^4~{\rm cm^{-3}}$; e.g., [1]). Molecules like CH_3OH, which prefer to stick in non-radiating states on dust grains unless disturbed by energetic events, often indicate the presence of shocks (e.g., [5]). Some of the emission lines

*e-mail: jens.kauffmann@mit.edu

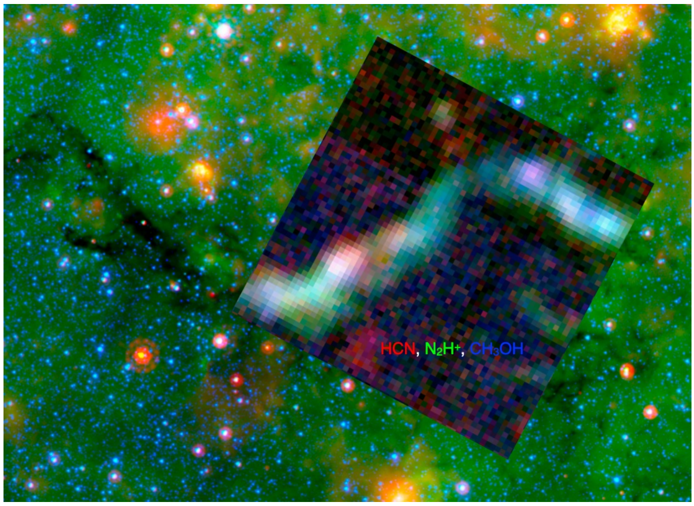

Figure 1. Example map of G11.11−0.12, a.k.a. the "Snake". LEGO data are shown as an inset on top of a Spitzer image obtained at 3.6–24 μm wavelength. A remarkable aspect is the relatively faint emission of HCN in this cloud of high gas density. Another noteworthy aspect is the substantial emission from CH_3OH in a cloud largely devoid of star formation feedback.

are sensitive to the impact of star formation feedback. For example, it has been observed that the HCN-to-HNC line ratio is sensitive to the temperature of the gas heated by stars (e.g., [6]), while UV-radiation from hot young stars impacts the HCN-to-HCO^+ line ratio (e.g., [7]).

Interestingly, while molecular line ratios are frequently employed to study extragalactic star formation, very little observational work has been done in the Milky Way to "calibrate" emission lines and their ratios as probes of molecular cloud structure. This work is important: recent observational work by LEGO and other efforts shows for example that the HCN (1–0) transition is *not* a tracer of high-density gas at $\gg 10^4~\mathrm{cm}^{-3}$ as reasoned before (e.g., [1]), but is a probe of gas at rather modest densities $\sim 10^3~\mathrm{cm}^{-3}$ (e.g., [8]).

The LEGO ("Line Emission in Galaxy Observations") project was therefore specifically designed to advance the interpretation of extragalactic molecular line emission via the observational "calibration" of emission line properties via detailed studies of well-characterized molecular clouds in the Milky Way. This is enabled by imaging of large fields of $30' \times 30'$ and more size with the IRAM 30m-telescope in spectral setups that cover a large number of molecular emission lines. To give examples, LEGO can determine how a specific molecular emission line responds to variations in gas temperature, density, or star formation activity. This is possible because all of these physical properties can be gleaned from rich reference data, such as dust emission maps from Herschel or infrared surveys for young stars.

This article focuses on providing an overview of LEGO's targets, the observational setup, and the processing strategy (Sec. 2). It also outlines selected initial results, describes the project's status and schedule, as well the expected long-term impact of the project on the field (Sec. 3). A summary is presented in Sec. 4.

2 Project Overview

2.1 Sample Selection

LEGO is designed to obtain data for a comprehensive sample of targets. The resulting list is outlined in Table 1. Some targets are covered with large maps of $30' \times 30'$ size or larger. Maps of this scale are produced from data obtained at very high telescope scan speeds, resulting in

Table 1. Summary of molecular clouds targeted by the LEGO survey.

Targets	Map Sizes
nearby clouds: study of detailed physics	
Orion ONC, Orion L1641, IC5146, Rosette	$\geq 30' \times 30'$
luminous complexes: impact of star formation	
W49A, W43, G45.1+0.1	$30' \times 30'$
galactic disk: variation between inner and outer Milky Way	
G6.67−0.20, G8.15+0.27, G11.11−0.12, G25.34−0.20, G28.34+0.06, G28.67+0.08, G35.18−0.77, G45.48+0.06, G51.38−0.02, G125.6+2.1, G144.8+0.4, G149.0+0.4, G160.2+0.8, G168.6+1.0, G170.6−0.3	$10' \times 10'$
low metallicity: impact of low metallicity	
Sh 2–208, Sh 2–266, Sh 2–284	$30' \times 30'$

an effective angular resolution of $\vartheta_{\rm beam} \sim 60''$. Other targets are covered with smaller maps of $10' \times 10'$ size that are processed to $\vartheta_{\rm beam} = 35''$ resolution.

The project covers selected nearby clouds out to $d \sim 1.5$ kpc with maps of $30' \times 30'$ size or larger. These data have sub-parsec physical resolutions of $\ell_{\rm beam} = 0.29~{\rm pc} \cdot (d/{\rm kpc}) \cdot (\vartheta_{\rm beam}/60'')$. Such data permits to moderately resolve the physical substructure of molecular clouds, so that the dependence of molecular line emission on the density and temperature structure of a molecular cloud (e.g., warm and diffuse at the surface, and cold and dense in the center) can be investigated. The sample further includes some of the most luminous star-forming regions in the Milky Way (taken from [9]). Data on these regions can be combined with observations of less luminous clouds in order to probe the impact of star formation and stellar luminosity on molecular line emission. The high-luminosity complexes are of large angular size, and so they are covered with large maps of $30' \times 30'$ size. A first LEGO study of a high-luminosity complex is presented in [10]. Similarly, several clouds believed to have metallicities as low as 1/10 of the solar value are observed in the outer Milky Way (see [11–13]), in order to probe the impact of metallicity on molecular line emission. Finally, 15 clouds at galactic longitudes of $\ell = 6°$–$170°$ are sampled with maps of $10' \times 10'$ to probe the dependence of line emission on the galactic environment. The selection of this latter sample heavily builds on [14].

The key aspect of the LEGO sample is the substantial angular size of the maps. These maps correspond to physical scales $\ell_{\rm map} = 6~{\rm pc} \cdot (d/2~{\rm kpc}) \cdot (\vartheta_{\rm map}/10')$. Given the chosen combinations of target distance d and map size $\vartheta_{\rm map}$, the derived maps probe *large representative sections of molecular clouds, or clouds in their entirety*. LEGO achieves this by making use of substantial time allocations. Including overheads, imaging of one region of $10' \times 10'$ size in two setups at $\vartheta_{\rm beam} = 35''$ resolution for example takes 10.4 h, while such observations of a field of $30' \times 30'$ size with $\vartheta_{\rm beam} = 60''$ take 16.9 h.

2.2 Spectral Setup

LEGO primarily targets the 85–115 GHz spectral window. This can be achieved with two tunings of the EMIR receiver system on the IRAM 30m-telescope. LEGO employs a spectral resolution $\Delta\nu = 200$ kHz, corresponding to a velocity resolution of $0.6~{\rm km/s} \cdot (\Delta\nu/200~{\rm kHz}) \cdot (\nu/100~{\rm GHz})^{-1}$. The survey targets a noise temperature of 0.1 K in the $T_{\rm mb}$-scale, but in practice the sensitivity is typically higher.

The current data reduction pipeline extracts position-position-velocity data cubes for 29 "line groups" of up to 100 MHz width that stem from 15 molecular species (i.e., 24 species if including isotopologues). These groups are listed in Table 2. In selected regions we further extract 5 groups from deuterium-bearing molecules at frequencies in the range 72–86 GHz.

The overall philosophy driving the line selection and spectral setup are described in [10]. To give an example, we cover emission lines detected in the nearby M 51 galaxy (e.g., [15]).

It is very likely that the lines extracted by regular pipeline reduction only constitute a small fraction of the transitions that can be meaningfully studied with the LEGO data. We conclude this from an analysis of published line surveys of other targets. Would these objects be covered by LEGO maps of typical sensitivity, we would detect dozens of emission lines (e.g., assuming appropriate frequency coverage and distance, ≥ 92 transitions in a protostar like NGC 2264 CMM3 [16], ≥ 134 transition groups in an object like Orion KL [17], and ≥ 79 transition groups in a Class 0 object like L 483 [18]). Such rich detections of line emission towards compact objects like embedded young stars will only occur towards small beam-sized areas.

2.3 Data Processing

The LEGO pipeline builds on data that are calibrated by observatory systems. In addition to the observed spectra, the project also retrieves machine-readable logs from the telescope's data systems. A combination of scripts written in the Python and CLASS languages are used to process the data into science-ready data cubes. Python's superior ability to search and manage data is used to orchestrate the data reduction process, while CLASS is used (via the PyGILDAS interface) to perform low-level data processing steps. Science-ready data cubes in FITS format constitute the final product of the pipeline. In a given target, a cube is extracted for each of the transition groups listed in Table 2. A dedicated server at MIT Haystack Observatory is used to process all of the data in a unified environment.

Jupyter notebooks (https://jupyter.org) are used to document the pipeline, the pipeline execution, and the data products. Specifically, the entire pipeline is written in the form of richly-annotated Jupyter notebooks that can be imported as libraries. One notebook per target is used to control and document the pipeline execution. Additional notebooks are used to document the data quality and the overall features of the targets as revealed by the observations. All of the pipeline, as well as the rich automatically generated documentation on data products, can easily be browsed online via GitHub. All of this material will be part of an upcoming public data release. This includes science-ready data cubes for all targets.

2.4 Project Team

The project is headed and managed by J. Kauffmann (MIT Haystack), who also served as PI on the IRAM observing proposals. Observations were conducted, and data processing tools developed and applied, by A. Anderson (U. Hawaii), A.T. Barnes (AIfA Bonn), N. Brinkmann (MPIfR Bonn), A. Broadmeadow (U. Maryland), D. Colombo (MPIfR Bonn), A.E. Guzmán (Joint ALMA Observatory), W.J. Kim (U. Köln), A. Patel (MIT Haystack), L. Szűcs (NRAO), and V. Wakelam (U. Bordeaux). The scientific motivation for LEGO was developed, and the exploration of the data supported, by S. Aalto (Chalmers U.), T. Albertsson (MPIfR Bonn), F. Bigiel (AIfA Bonn), N.J. Evans II (U. Texas), S.C.O. Glover (U. Heidelberg), P.F. Goldsmith (JPL), C. Kramer (IRAM), K. Menten (MPIfR Bonn), Y. Nishimura (U. Tokyo, NAOJ), S. Viti (U. Leiden), Y. Watanabe (Shibaura Institute of Technology), A. Weiss (MPIfR Bonn), M. Wienen (MPIfR Bonn, U. Exeter), H. Wiesemeyer (MPIfR Bonn), and F. Wyrowski (MPIfR Bonn).

3 Project Timeline and Expected Impact

LEGO builds on a number of pilot studies that started in 2015. Primary data collection began in fall of 2017, and it was completed in the summer of 2020. Pilot studies of the project were for example presented in [8, 19], while [10] constitutes the first direct output from the LEGO survey. The pandemic required a re-focus of the project: given inefficiencies and hiring restrictions, the project had to concentrate on supporting ongoing student projects. This has

Table 2. Summary of transition groups observed by the LEGO survey.

Species	Transition	Frequency [GHz]
c–C_3H_2	$J_{K_a,K_c} = 2_{1,2}–1_{0,1}$	85.3389060
$H^{13}CN$	1–0	86.3401764
$H^{13}CO^+$	1–0	86.7542880
SiO	2–1	86.8469950
$HN^{13}C$	1–0	87.0908590
CCH	$N = 1-0, J = 3/2–1/2$	87.3169250
CCH	$N = 1-0, J = 1/2–1/2$	87.4020040
HNCO	$J_{K_a,K_c} = 4_{0,4}–3_{0,3}$	87.9252380
HCN	1–0	88.6318473
HCO^+	1–0	89.1885260
HNC	1–0	90.6635640
H	41α (42–41)	92.0344340
N_2H^+	1–0	93.1737770
HC_5N	35–34	93.1881250
HC_5N	36–35	95.8503350
$C^{34}S$	2–1	96.4129500
CH_3OH-E	$J_K = 2_{-1}–1_{-1}$	96.7393630
CH_3OH-A	$J_K = 2_0–1_0$	96.7413770
CS	2–1	97.9809530
SO	$J_K = 3_2–2_1$	99.2999050
HC_3N	12–11	109.1736380
OCS	9–8	109.4630630
$C^{18}O$	1–0	109.7821760
HNCO	$J_{K_a,K_c} = 5_{0,5}–4_{0,4}$	109.9057530
^{13}CO	1–0	110.2013540
$C^{17}O$	1–0	112.3589880
CN	$N = 1-0, J = 1/2–1/2$	113.1913170
CN	$N = 1-0, J = 3/2–1/2$	113.4909820
CO	1–0	115.2712020

resulted in exciting new results. Specifically, studies of IC 5146-W (Anderson et al., in prep.) and G11.11−0.12 (Broadmeadow et al., in prep.) expose additional striking examples of faint HCN emission in dense molecular clouds, just as found in LEGO pilot studies [8].

More generally, the current study of LEGO targets reveals a surprising diversity in line emission properties. Line luminosity ratios L_i/L_j (i.e., for signals integrated over entire clouds) show a strong cloud-to-cloud variation, independent of the specific line pair i, j chosen. At this moment it appears that no line pair can be characterized by a single representative line ratio. Extragalactic line observations typically probe regions of $\gtrsim$ 100 pc size that contain several clouds, and so *the line emission properties observed in galaxies constitute complex averages over diverse cloud populations*. To give an example, line ratios are observed to vary within galaxies (e.g., [20]), and the LEGO data imply that *line ratio variations within galaxies might indicate that cloud populations change between galactic environments*.

LEGO is thus likely to provide new and important new information that will ultimately allow us to disentangle the physical substructure of galaxies. By connecting variations in line ratios to physical cloud properties like the gas temperature and density, *LEGO will provide the tools needed to map out the variation of gas properties in galaxies*.

LEGO is now in a push to deliver science-ready cubes and associated documentation for all transitions and targets covered by LEGO. Delivery is planned for the winter of 2022/23.

4 Summary

LEGO generates a very detailed picture of how line intensities and ratios vary within and between molecular clouds, and how these variations depend on physical cloud properties like gas temperature and density. The observations reveal strong cloud-to-cloud variations in line ratios. This means that line ratios can be used to uncover important new physical constraints on the physical structure of extragalactic molecular clouds. LEGO is enabled by NSF-AAG grant 1909097.

References

[1] Y. Gao, P.M. Solomon, The Astrophysical Journal **606**, 271 (2004)
[2] M.Y. Grudić, P.F. Hopkins, C.A. Faucher-Giguère, E. Quataert, N. Murray, D. Kereš, Monthly Notices of the Royal Astronomical Society **475**, 3511 (2018)
[3] D. Colombo, A. Hughes, E. Schinnerer, S.E. Meidt, A.K. Leroy, J. Pety, C.L. Dobbs, S. García-Burillo, G. Dumas, T.A. Thompson et al., The Astrophysical Journal **784**, 3 (2014)
[4] J. Sun, A.K. Leroy, A. Schruba, E. Rosolowsky, A. Hughes, J.M.D. Kruijssen, S. Meidt, E. Schinnerer, G.A. Blanc, F. Bigiel et al., The Astrophysical Journal **860**, 172 (2018)
[5] R. Bachiller, S. Liechti, C. Walmsley, F. Colomer, Astronomy & Astrophysics **295**, L51 (1995)
[6] P. Schilke, C. Walmsley, G. Pineau Des Forets, E. Roueff, D. Flower, S. Guilloteau, Astronomy & Astrophysics **256**, 595 (1992)
[7] R. Meijerink, M. Spaans, F.P. Israel, Astronomy and Astrophysics **461**, 793 (2007)
[8] J. Kauffmann, P.F. Goldsmith, G. Melnick, V. Tolls, A. Guzman, K.M. Menten, Astronomy & Astrophysics **605**, L4 (2017)
[9] J.S. Urquhart, C.C. Figura, T.J.T. Moore, M.G. Hoare, S.L. Lumsden, J.C. Mottram, M.A. Thompson, R.D. Oudmaijer, Monthly Notices of the Royal Astronomical Society **437**, 1791 (2014)
[10] A.T. Barnes, J. Kauffmann, F. Bigiel, N. Brinkmann, D. Colombo, A.E. Guzmán, W.J. Kim, L. Szűcs, V. Wakelam, S. Aalto et al., Monthly Notices of the Royal Astronomical Society **497**, 1972 (2020)
[11] C. Yasui, N. Kobayashi, M. Saito, N. Izumi, The Astronomical Journal **151**, 115 (2016)
[12] A. Mehner, W.J. de Wit, J.H. Groh, R.D. Oudmaijer, D. Baade, T. Rivinius, F. Selman, H.M.J. Boffin, C. Martayan, Astronomy & Astrophysics **585**, A81 (2015)
[13] E. Puga, S. Hony, C. Neiner, A. Lenorzer, A.M. Hubert, L.B.F.M. Waters, F. Cusano, V. Ripepi, Astronomy and Astrophysics **503**, 107 (2009)
[14] M. Wienen, F. Wyrowski, K.M. Menten, J.S. Urquhart, T. Csengeri, C.M. Walmsley, S. Bontemps, D. Russeil, L. Bronfman, B.S. Koribalski et al., Astronomy & Astrophysics **579**, A91 (2015)
[15] Y. Watanabe, N. Sakai, K. Sorai, S. Yamamoto, The Astrophysical Journal **788**, 4 (2014)
[16] Y. Watanabe, N. Sakai, A. López-Sepulcre, R. Furuya, T. Sakai, T. Hirota, S.Y. Liu, Y.N. Su, S. Yamamoto, Astrophysical Journal **809** (2015)
[17] B.E. Turner, The Astrophysical Journal Supplement Series **70**, 539 (1989)
[18] M. Agúndez, N. Marcelino, J. Cernicharo, E. Roueff, M. Tafalla, Astronomy and Astrophysics **625**, A147 (2019)
[19] P.F. Goldsmith, J. Kauffmann, The Astrophysical Journal **841**, 25 (2017)
[20] F. Bigiel, A.K. Leroy, M.J. Jiménez-Donaire, J. Pety, A. Usero, D. Cormier, A. Bolatto, S. Garcia-Burillo, D. Colombo, M. González-García et al., The Astrophysical Journal Letters **822**, L26 (2016)