

THE SUBMILLIMETER SPECTRUM OF GLYCOLALDEHYDE

P. BRANDON CARROLL¹, BRIAN J. DROUIN², AND SUSANNA L. WIDICUS WEAVER¹

¹ Department of Chemistry, Emory University, Atlanta, GA 30322, USA; pbcarro@emory.edu, susanna.widicus.weaver@emory.edu

² Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91125, USA; brian.j.drouin@jpl.nasa.gov

Received 2010 July 20; accepted 2010 August 26; published 2010 October 15

ABSTRACT

Glycolaldehyde (HOCH₂CHO) is a sugar-related interstellar prebiotic molecule that has been detected in two star-forming regions, Sgr B2(N) and G31.41+0.31. Glycolaldehyde is suspected to form from photodissociation-driven ice chemistry, and therefore can be used to trace complex organic chemistry in interstellar environments. The relative abundance of glycolaldehyde to its structural isomers, methyl formate (HCOOCH₃) and acetic acid (CH₃COOH), can be used to constrain astrochemical models. Given its central role in the complex chemistry of the interstellar medium, glycolaldehyde has been suggested as a prime molecular target for upcoming high-frequency molecular line searches using new far-infrared observatories. In particular, glycolaldehyde is a target for the *Herschel Space Observatory* HEXOS Key Program, which is conducting spectral line surveys of the Sgr B2(N) and Orion KL star-forming regions across the entire HIFI band. Laboratory investigation of glycolaldehyde in the HIFI frequency range is required before its lines can be identified in these spectra. We have therefore acquired the laboratory spectrum of glycolaldehyde in selected frequency ranges across the submillimeter range. We present here the laboratory spectral analysis of the ground vibrational state of glycolaldehyde up to 1.2 THz.

Key words: ISM: molecules – line: identification – methods: data analysis – methods: laboratory – molecular data – molecular processes – radio lines: ISM

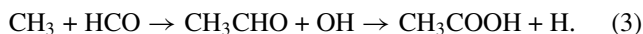
Online-only material: machine-readable table

1. INTRODUCTION

Glycolaldehyde (HOCH₂CHO) is the simplest α -hydroxy aldehyde and is therefore a sugar-related species. Because it is the simplest molecule in this class of biologically relevant compounds, it plays a central role in interstellar prebiotic chemistry. Glycolaldehyde has been detected in two interstellar sources, Sgr B2(N) (Hollis et al. 2000; Halfen et al. 2006) and G31.41+0.31 (Beltrán et al. 2009). In addition to its importance as a prebiotic species, glycolaldehyde is also a key reaction intermediate in the complex organic chemistry of interstellar clouds. Its large spatial distribution and low rotational temperature in Sgr B2(N) (Hollis et al. 2004) indicate that glycolaldehyde originated from condensed-phase ice chemistry in the extended envelope, rather than from the warm gas-phase chemistry in the hot core. Astrochemical models suggest that the dominant formation route for glycolaldehyde is the grain surface radical–radical reaction



which involves photodissociation products arising from CH₃OH and H₂CO (Garrod et al. 2008). Glycolaldehyde has two structural isomers that are also known interstellar molecules, methyl formate (HCOOCH₃) and acetic acid (CH₃COOH). Similar radical–radical reactions are proposed to form these two isomers, through channels involving other methanol photodissociation products (Garrod et al. 2008):



The relative abundance of glycolaldehyde to its two structural isomers therefore provides important constraints on astrochemical models of complex organic chemistry.

Glycolaldehyde has only been detected in cold, extended envelopes rather than in the warm chemistry of hot cores, which is surprising given its chemical link to the highly abundant hot core molecule methyl formate. Additional high-sensitivity studies that probe the warm and compact gas near newly formed stars may yet reveal warmer glycolaldehyde chemistry, and such observations would provide important constraints for models of complex chemistry. In light of its importance in interstellar organic chemistry and its interest as a prebiotic molecule, glycolaldehyde is a target molecule for upcoming spectral line searches using new far-infrared (far-IR) observational facilities. In particular, glycolaldehyde is considered an interstellar “flower” for the ongoing *Herschel Space Observatory* HEXOS Key Program (E. A. Bergin 2010, private communication). This program is conducting deep spectral surveys of the Orion KL and Sgr B2(N) star-forming regions across the entire HIFI spectral range (480 GHz–1.25 THz, 1.41–1.91 THz). Much of the laboratory work dedicated thus far to supporting *Herschel* Key Programs has focused on the so-called interstellar weeds—molecules in high abundance with dense spectra across the submillimeter band. Very little submillimeter spectral information is available for less abundant interstellar molecules like glycolaldehyde, and this limits the search for these “flowers” amongst the “weeds.”

The rotational spectrum of glycolaldehyde has been studied extensively in the microwave and millimeter regions. Marstokk & Møllendal (1970, 1973) obtained the pure rotational spectrum of glycolaldehyde from 12 to 26 GHz, measured the electric dipole moment, and assigned lines in the ground and three excited vibrational states. Butler et al. (2001) reported millimeter and submillimeter assignments for the ground vibrational state up to 354 GHz. A follow-up study by Widicus Weaver et al. (2005) included assignments for the ground state and three excited vibrational states up to 354 GHz and a partition function analysis. No pure rotational spectral information is

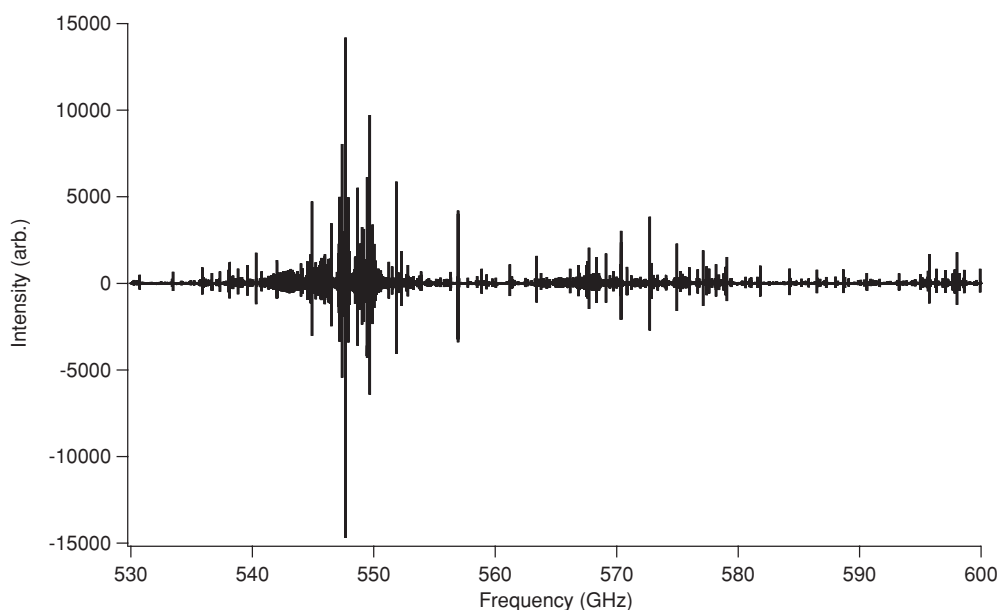


Figure 1. Glycolaldehyde spectrum from 530 to 600 GHz.

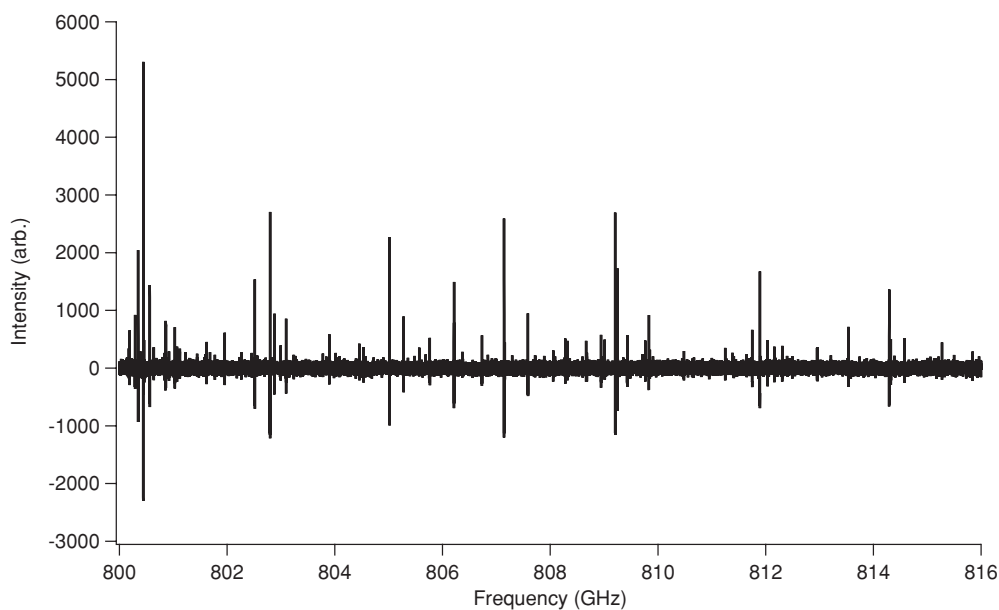


Figure 2. Glycolaldehyde spectrum from 800 to 816 GHz.

available in the literature for glycolaldehyde at frequencies above 354 GHz.

Before searches for glycolaldehyde lines in *Herschel* spectra can be conducted, a laboratory study overlapping with the *Herschel* HIFI frequency range is required. To this end, we have obtained the spectrum of glycolaldehyde in selected frequency ranges across the submillimeter band to provide rest frequencies for comparison to astronomical spectra. Here we present the details of our experimental work and the assignment of the ground vibrational state for glycolaldehyde up to 1.2 THz.

2. EXPERIMENTAL DETAILS

Spectra were taken in three windows between 520 and 608 GHz, 800 and 900 GHz, and 1.07 and 1.2 THz. Spectra from 530 to 600 GHz, 800 to 816 GHz, and 1.08 to 1.16 THz are shown in Figures 1–3, respectively. In all cases, the spectra

were acquired using a 99% pure glycolaldehyde sample from Sigma–Aldrich. The 520–608 GHz and 1.07–1.2 THz spectra were obtained at the Jet Propulsion Laboratory (JPL) molecular spectroscopy lab. The details of the spectrometer have been previously described by Drouin et al. (2005), and will not be described in further detail here. The solid glycolaldehyde sample was loaded in a flask and attached to a 3 m long, 10 cm diameter, static gas cell. Glycolaldehyde vapor was then delivered to the cell by back-pumping ~ 10 mTorr of sample into the cell. The data were acquired as previously described over a period of four days, during which time the sample pressure changed by less than 5 mTorr.

The 800–900 GHz spectra were obtained at Emory University using a single-pass direct absorption flow-cell spectrometer. The spectrometer consists of a 2.1 m long, 5.08 cm inner-diameter PVC cell with the sample input and vacuum lines on opposite ends. Cell pressure is monitored with a KJL-6000 thermocouple

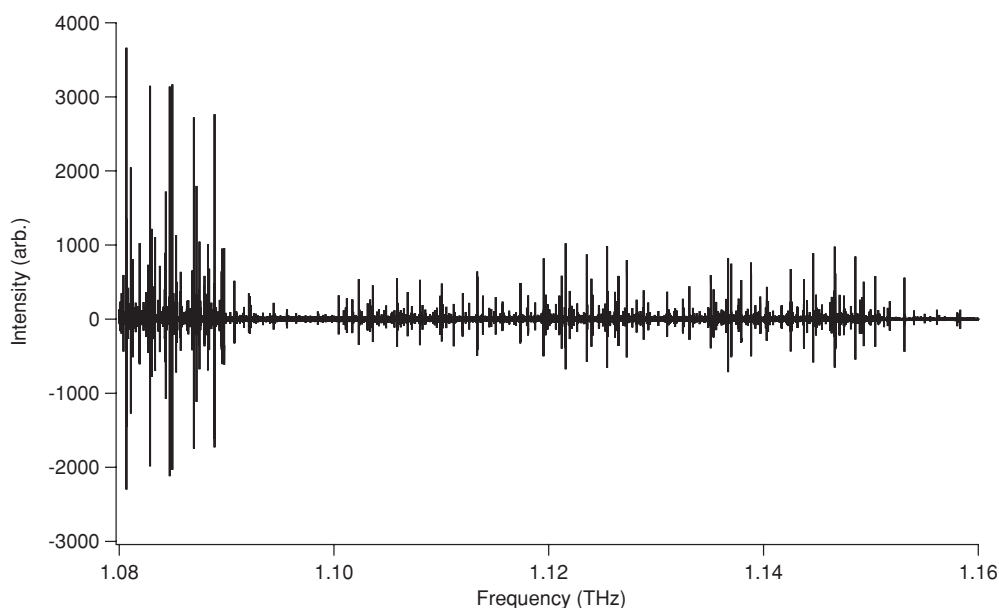


Figure 3. Glycolaldehyde spectrum from 1.08 to 1.16 THz.

pressure gauge situated near the pump line. Sample was added to a round-bottom flask and was introduced to the evacuated cell through a needle valve, which was used to control the pressure of the sample in the cell. Sample pressure was held at 10 ± 3 mTorr by gently heating the sample flask to 45 ± 2 °C. Radiation was coupled into and out of the cell through Teflon lenses. An Agilent E8257D microwave synthesizer providing a 1–50 GHz output frequency was coupled with a Virginia Diodes frequency multiplier chain, providing output frequency coverage over the range 50 GHz–1 THz. The signal from the synthesizer was frequency modulated using the synthesizer's internal modulation source set to a frequency of 15 kHz. The output radiation was detected with a QMC Instruments QFI/2BI InSb hot electron bolometer detector. The signal from the bolometer was processed with a Stanford Research SR830 lock-in amplifier using 2f detection. The resultant signal was then digitized and recorded by a computer using custom data acquisition software.

3. DATA ANALYSIS

The spectra were analyzed using the SPFIT and SPCAT programs (Pickett et al. 1998). The parameters for the ground vibrational state determined by Widicus Weaver et al. (2005), including the rotational constants, distortion constants, and dipole moment, were used with an asymmetric top Hamiltonian in the I -representation and Watson-A reduction in SPFIT/SPCAT to create a spectral prediction for the glycolaldehyde ground vibrational state transitions up to 1.3 THz. This prediction was then used to assign the observed ground vibrational state lines. Newly assigned transitions were added to the data set, which included the assigned transitions from the previous studies. The Hamiltonian parameters were then optimized and the quantum assignments were modified if necessary. The Hamiltonian and data set were then used to create a prediction from the modified fit parameters, and this prediction was used to assign higher frequency transitions. This process was performed iteratively until all possible ground vibrational state assignments were made within the spectral ranges covered. The primary parameter used to evaluate the quality of the fit was the microwave rms, cal-

Table 1
Spectral Parameters Determined for the Ground Vibrational State of Glycolaldehyde

Parameter	Widicus Weaver et al. (2005)	This Work	Units
A	18446.26074(43)	18446.25696(134)	MHz
B	6525.996379(161)	6525.99578(51)	MHz
C	4969.235801(149)	4969.23452(50)	MHz
Δ_J	6.222339(55)	6.21986(47)	kHz
Δ_{JK}	-20.397978(222)	-20.39838(219)	kHz
Δ_K	47.72338(47)	47.6959(36)	kHz
δ_J	1.8337838(135)	1.833997(112)	kHz
δ_K	8.87889(41)	8.8676(45)	kHz
Φ_J	-6.465(42)	-7.710(170)	mHz
Φ_{JK}	0.15657(108)	0.17086(168)	Hz
Φ_{KJ}	-0.7721(34)	-0.8256(57)	Hz
Φ_K	1.05703(292)	1.0609(65)	Hz
ϕ_J	-2.5042(167)	-2.2752(253)	mHz
ϕ_{JK}	-12.98(82)	-25.42(105)	mHz
ϕ_K	0.1909(107)	0.3096(174)	Hz
L_J	...	0.0776(222)	μ Hz
L_{JK}	...	0.211(203)	μ Hz
L_{JK}	...	-3.90(65)	mHz
L_{KKJ}	...	26.65(154)	μ Hz
L_K	...	-21.82(252)	μ Hz
J_{\max}	66	63	...
K_{\max}	29	32	...
No. of Lines	1657	1412	...
Fit rms	0.110	0.061	MHz

Note. 1σ errors are listed in parentheses in units of last significant figure.

culated by averaging the rms difference between the predicted frequency of the transitions and their observed frequencies.

A total of 907 new ground vibrational state lines were assigned at frequencies ≥ 500 GHz. The assigned frequencies varied from the prediction based on the parameters from Widicus Weaver et al. (2005), with frequency offsets between the prediction and the observed spectral lines as large as 15 MHz at frequencies above 600 GHz. The new assignments were combined with the ground vibrational state assignments from Widicus Weaver et al. (2005) to yield the final fit constants shown in Table 1. All of the ground vibrational state assignments

Table 2
Assigned Transitions for the Ground Vibrational State of Glycolaldehyde

J'	K'_a	K'_c	J''	K''_a	K''_c	Frequency (MHz)	Obs. – Calc. (MHz)	Uncertainty (MHz)	Intensity Weighting for Blended Lines
20	12	8	19	11	9	522012.72010	–0.03424	0.10000	0.5000
20	12	9	19	11	8	522012.72010	–0.03397	0.10000	0.5000
27	9	18	26	8	19	522460.79100	0.05889	0.10000	...
18	13	5	17	12	6	524363.98420	0.00235	0.10000	0.5000
18	13	6	17	12	5	524363.98420	0.00235	0.10000	0.5000
46	7	39	45	8	38	525042.70060	–0.03478	0.10000	...
47	6	41	46	7	40	526025.10330	–0.01189	0.10000	...
47	7	41	46	6	40	526194.12870	0.05471	0.10000	...
48	5	43	47	6	42	526334.99210	0.01866	0.10000	...
48	6	43	47	5	42	526341.79470	–0.04073	0.10000	...

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

for glycolaldehyde included in this analysis are listed in Table 2. Blended transitions were fitted with intensity weighting, and this weighting is given in Table 2. The full line catalog from this analysis, including all assignments and other predicted rotational line frequencies, is available in the JPL millimeter and submillimeter molecular line catalog (Pickett et al. 1998)³.

During the assignment and analysis process, it was noted that the observed–calculated residuals for many of the line assignments from Butler et al. (2001) and included in the analysis of Widicus Weaver et al. (2005) were significantly higher than the estimated line uncertainties, and often 3–5 times larger than the microwave rms of the fit. In addition, a handful of the highest- J line assignments from the original microwave work (Marstokk & Møllendal 1970, 1973) had residuals larger than the estimated uncertainties. These issues had been present in earlier studies, but there was little independent information as to the validity of the assignments and/or calibrations. The higher frequencies measured in this work include independent measurements (b-type transitions) of many of the same J , K levels measured in the work of Butler et al. (2001; a-type transitions). This allows the two data sets to directly compete in the nonlinear least-squares-fitting procedure. Our tests indicate that the submillimeter work is of higher accuracy, perhaps due to calibration issues in the earliest versions of the FASSST spectrometer used by Butler et al. (2001). Our effort is focused on providing the most accurate and precise prediction for astronomical purposes; therefore, it is best to exclude potentially inaccurate data from the analysis. Several changes to the data set were made with this motivation in mind. First, all estimates of frequency uncertainties were set to 0.1 MHz so as to eliminate any weighting that favored the microwave assignments over the other assignments. Then all of the assignments from Butler et al. (2001) were removed because the majority of these assignments had biased residuals indicating a possible problem with experimental frequency calibration. Likewise, any line assignment that resulted in an observed–calculated residual ≥ 0.2 MHz (twice the estimated experimental uncertainty) was removed. Removal of lines from the analysis was performed in an iterative process, beginning with the lines that had the highest residuals, and working down to the 0.2 MHz threshold.

After this initial removal of lines, it was noted that many of the removed lines at frequencies >500 GHz were in the range $K_a \sim 28$ –31, and there was a clear trend of increasing residual as a function of the K_a quantum number. A set of octic

centrifugal distortion constants were added to the Hamiltonian, and these lines were added back into the line list for inclusion in the analysis. The residuals for these transitions fell below the 0.2 MHz cutoff threshold upon inclusion of the higher order constants.

A total of 54 assignments were ultimately removed from the data set in addition to removing all assignments from 120 to 357 GHz. Only 18 of the 54 suspicious assignments were included in previous studies, which is $\sim 1\%$ of the total number of lines assigned for the ground vibrational state in the Widicus Weaver et al. (2005) analysis. Most of the lines removed from the fit were high J or high K_a lines that are likely bad assignments, weak lines that were close to the spectral noise level, or severely blended lines without absolute center frequency determination. The quality of the fit was reassessed after these changes were implemented, and the microwave rms was reduced from 0.110 to 0.061 MHz, which is a significant improvement over previous studies. While the removed lines may actually be correct assignments, the data quality makes these assignments inadequate for inclusion in the analysis, and their removal allows the determination of more reliable parameters on which extrapolations can be based. The parameters determined from this final analysis are those listed in Table 1.

4. DISCUSSION

The pure rotational spectrum of the glycolaldehyde ground vibrational state has now been analyzed up to 1.2 THz. The analysis revealed that the spectral prediction based on the Widicus Weaver et al. (2005) study, which is the catalog previously listed in the JPL millimeter and submillimeter molecular line catalog (Pickett et al. 1998, see also footnote 3), is shifted from the observed lines in the submillimeter and THz spectra by factors as large as 15 MHz. Glycolaldehyde is a reasonably rigid molecule, with global sextic distortion constants used for all vibrational states in the millimeter analysis (Widicus Weaver et al. 2005). Yet the prediction to higher frequencies based on earlier millimeter work is not sufficient to guide astronomical searches above 600 GHz, even for such a rigid species. The new centrifugal distortion constants determined in the present study display large changes compared to the previous study, in some cases beyond the reported 1σ uncertainties. This is not surprising given that the previous analysis used a global fit for distortion constants including all vibrational states. Underlying problems with previous assignments were also revealed during the analysis of the higher

³ <http://spec.jpl.nasa.gov>

frequency spectra, and these assignments likely influenced the best-fit parameters and skewed the prediction to higher frequency. This study underscores the importance of submillimeter laboratory spectroscopy to guide high-frequency astronomical searches for complex organic molecules, even for molecules that have extensive lower frequency information available.

This analysis provides a basis for the comparison of the glycolaldehyde spectrum to interstellar spectral line surveys currently being conducted with the *Herschel Space Observatory*, and for future studies with additional far-IR telescopes that are coming online in the next few years, including the Atacama Large Millimeter Array (ALMA) and the Stratospheric Observatory for Infrared Astronomy (SOFIA). While the laboratory experiment does not entirely cover the frequency window of the HIFI spectrometer on *Herschel* (up to 1.9 THz), the analysis presented here can be used to generate a spectral prediction up to higher frequencies. Such a prediction is estimated to be reliable to an accuracy of ≤ 15 MHz for all lines, and for the strongest lines should be reliable to within ≤ 2 MHz. Additional higher frequency laboratory work is required before the full analysis up to 1.9 THz can be completed.

Other strong lines from excited vibrational states are observed in room temperature laboratory spectra, but only the ground vibrational state analysis has been included here. It was found during analysis that the predictions for these excited states are significantly shifted from the observed spectrum, at times by more than 20 MHz, and that the addition of higher order distortion constants did not result in physically meaningful fit parameters. The assignment of these higher vibrational state

lines appears to be more challenging and is beyond the scope of the work presented here.

Portions of this paper present research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This work was supported by SLWW's startup funds provided by Emory University. We are grateful to Geoffrey Blake for providing lodging for P.B.C. during his visit to Pasadena. We thank the JPL Microwave, Millimeter, and Submillimeter Spectroscopy Group, particularly Shanshan Yu, for their help during data collection. We also gratefully acknowledge the support and services rendered by the Emory and JPL staff.

REFERENCES

- Beltrán, M. T., Codella, C., Viti, S., Neri, R., & Cesaroni, R. 2009, *ApJ*, **690**, L93
- Butler, R. A. H., De Lucia, F. C., Petkie, D. T., Møllendal, H., Horn, A., & Herbst, E. 2001, *ApJ*, **134**, 319
- Drouin, B. J., Maiwald, F. W., & Pearson, J. C. 2005, *Rev. Sci. Instrum.*, **76**, 093113
- Garrod, R. T., Widicus Weaver, S. L., & Herbst, E. 2008, *ApJ*, **682**, 283
- Halfen, D. T., Apponi, A. J., Woolf, N., Polt, R., & Ziurys, L. M. 2006, *ApJ*, **639**, 237
- Hollis, J. M., Jewell, P. R., Lovas, F. J., & Remijan, A. 2004, *ApJ*, **613**, L45
- Hollis, J. M., Lovas, F. J., & Jewell, P. R. 2000, *ApJ*, **540**, L107
- Marstokk, K. M., & Møllendal, H. 1970, *J. Mol. Struct.*, **5**, 205
- Marstokk, K. M., & Møllendal, H. 1973, *J. Mol. Struct.*, **16**, 259
- Pickett, H. M., Poynter, R. A., Cohen, E. A., Delitsky, M. L., Pearson, J. C., & Müller, H. S. P. 1998, *J. Quant. Spectrosc. Radiat. Transfer*, **60**, 883
- Widicus Weaver, S. L., Butler, R. A. H., Drouin, B. J., Petkie, D. T., Dyl, K. A., De Lucia, F. C., & Blake, G. A. 2005, *ApJS*, **158**, 188