



---

Universidad de Valladolid



TESIS DOCTORAL:

**FROM INTERSTELLAR SYSTEMS TO  
TERRESTRIAL ORGANIC AND  
BIOMOLECULES: A SYNERGETIC  
THEORETICAL AND EXPERIMENTAL  
ROTATIONAL STUDY**

Presentada por Miguel Sanz Novo para optar al  
grado de  
Doctor por la Universidad de Valladolid

Dirigida por:  
José Luis Alonso Hernández y Carmen  
Barrientos Benito



The following financial support is gratefully acknowledge:

### Ministerio de Ciencia e Innovación



Grants: CTQ2016- 76393-P,  
PID2019-111396GB-I00,  
ID2020-117742GB-I00  
and predoctoral FPU Grant (FPU17/02987)

European Research Council under the European Union's Seventh  
Framework Programme

**nanocosmos**

(FP/2007–2013)/ERC-2013-SyG, Grant Agreement n. 610256 NANOCOSMOS

Junta de Castilla y León



Grants VA077U16 and VA244P20



## Agradecimientos

---

Esta tesis no hubiera sido posible sin el gran apoyo de mis directores de tesis José Luis Alonso y María Carmen Barrientos. Infinitas gracias por acogerme, enseñarme y guiarme durante estos últimos años en un mundo tan apasionante como es la Química-Física, así como brindarme una oportunidad insuperable para desarrollar un trabajo tan completo y multidisciplinar como el que hemos conseguido presentar.

Gracias a todos los miembros del GEM y del CCG, especialmente a Íker, Lucie, Elena, Santi y Pilar, así como a los recién incorporados Raúl, Sergio y Diego, cuya gran humanidad y ayuda ha enriquecido enormemente mi formación como científico y también como persona.

Quiero agradecer el apoyo incondicional de mi familia, en especial a mis padres y a Anita. Gracias por servirme de ejemplo y sobre todo por estar conmigo en los momentos buenos, pero también en aquellos no tan buenos.

A los profesores José María Andrés y Rafael Pedrosa por permitirme acercarme muy tempranamente a la labor de investigación científica.

A Jesús Nieto, por darle “color” a la Memoria de la tesis con sus impresionantes fotografías de diversos objetos astronómicos.

Gracias también al prof. Karl Menten y, de forma especial, al Dr. Arnaud Belloche por acogerme durante varios meses en su grupo en el Max Planck de Radioastronomía en tiempos tan difíciles como estos de pandemia, abriéndome las puertas e instruyéndome en el campo de la radio astronomía. Gracias por recibirme y hacerme sentir, verdaderamente, como en mi propia casa.

Finalmente, quiero agradecer el apoyo de mis amigos y de todas las personas que me he ido encontrando en el camino, sin las cuales estos últimos años mi vida habría sido, con total certeza, mucho menos interesantes y amenos.



*A mi hermana pequeña*





## List of publications related to this Doctoral Thesis

---

**“Formation of protonated glycine isomers in the interstellar medium”**

Miguel Sanz-Novo, Antonio Largo, Pilar Redondo, and Carmen Barrientos, *ACS Earth and Space Chemistry*, 2019, 3 (7), 1170-1181

**“The shape of the archetypical oxocarbon squaric acid and its water clusters”**

Miguel Sanz-Novo, Elena R. Alonso, Iker León, José L. Alonso, *Chem. Eur. J.* 2019, 25, 1–9.

**“Amino acetaldehyde conformers: structure and spectroscopic properties”**

Pilar Redondo, Miguel Sanz-Novo, Antonio Largo and Carmen Barrientos, *MNRAS*, 2019, 492, 2, 1827–1833.

**“Interstellar glycolamide: A comprehensive rotational study and an astronomical search in Sgr B2(N)”**

Miguel Sanz-Novo, Arnaud Belloche, José L. Alonso, Lucie Kolesníková, Robin T. Garrod, Santiago Mata, Holger. S. P. Müller, Karl. M. Menten and Yang Gong. *A&A*, 2020, 639, A135

**“Formation of interstellar cyanoacetamide: a rotational and computational study”**

Miguel Sanz-Novo, Íker León, José L. Alonso, Antonio Largo and Carmen Barrientos, *A&A*, 2020, 644, A3, 10.

**“Laboratory detection of cyanoacetic acid: a jet-cooled rotational study”**

Miguel Sanz-Novo, Íker León, Elena R. Alonso, Lucie Kolesníková and José L. Alonso, *APJ*, 2021, 915:76 (8pp).

**“Toward the limits of complexity of interstellar chemistry: rotational spectroscopy and astronomical search for n- and i-butyraldehyde”**

Miguel Sanz-Novo, Arnaud Belloche, Victor M. Rivilla, Robin T. Garrod, José L. Alonso, Pilar Redondo, Carmen Barrientos, Lucie Kolesníková, J. C. Valles, Lucas Rodríguez-Almeida, Izaskun Jimenez-Serra y Jesús Martín-Pintado, Holger S. P. Müller and Karl M. Menten under review.

**“Unleashing the shape of L-DOPA at last”**

Miguel Sanz-Novo, Íker León, Elena R. Alonso and José L. *Phys. Chem. Chem. Phys.*, 2021, Accepted Manuscript.

**“A rotational study of interstellar acetohydroxamic acid: a glycine isomer”**

Miguel Sanz-Novo, Íker León, Santiago Mata and José L. Alonso in preparation.

**“Shape-shifting molecules: unveiling the valence tautomerism phenomena in barbaralones”**

Miguel Sanz-Novo, Mauro Mato, Íker León, Antonio Echavarren and José L. Alonso under review.



<b>CHAPTER I: Summary – Resumen</b> .....	1
<b>CHAPTER IIa: Introduction</b> .....	5
1. Background.....	6
2. Interstellar medium (ISM): composition and relevance of the grains.....	8
3. Laboratory Astrochemistry or Astrophysics.....	8
4. Computational Astrochemistry.....	10
5. Molecules detected in the ISM: complex organic molecules (COMs).....	11
5.1 Origin of life: “glycine, the missing molecule”.....	15
5.2 Cyanides in the interstellar medium.....	19
5.3 Aldehydes and branched-chain species in the interstellar medium.....	20
6. The shape of biomolecules, fluxional systems and organic synthons.....	22
7. Thesis outline.....	25
8. References.....	27
<b>CHAPTER IIb: Introducción</b> .....	33
1. Antecedentes.....	34
2. Medio interestelar (ISM): composición e importancia de los granos.....	36
3. Astroquímica o Astrofísica de laboratorio.....	37
4. Astroquímica computacional.....	38
5. Moléculas descubiertas en el ISM: moléculas orgánicas complejas (COMs).....	40
5.1 Origen de la vida: “glicina, la molécula perdida”.....	44
5.2 Nitrilos en el medio interestelar.....	48
5.3 Aldehídos y especies de cadena ramificada en el medio interestelar.....	49
6. Estructura tridimensional de biomoléculas, sistemas fluxionales y sintones orgánicos.....	51
7. Contenido de la Memoria.....	54
8. Referencias.....	55
<b>CHAPTER III: Methodology</b> .....	61
1. Experimental techniques: rotational spectroscopy.....	62
1.1 Laser ablation time domain Fourier transform microwave spectroscopy.....	63
1.2 Frequency modulated (FM) millimeter and submillimeter-wave spectroscopy.....	73
2. General work strategy.....	75
3. Computational methods.....	77
4. Theoretical tools for rotational analysis.....	89
5. References.....	96

**CHAPTER IV: Experimental and theoretical study of interstellar glycine isomers.....101**

1. Introduction.....	102
2. Results and discussion.....	105
2.1 Theoretical study of neutral and protonated glycine isomers.....	105
2.2 Experimental rotational study of acetohydroxamic acid.....	121
2.3 Experimental rotational study of glycolamide.....	129
2.4 Search for glycolamide toward Sgr B2(N2) with ALMA.....	136
2.5 Search for glycolamide toward Sgr B2(N) with Effelsberg.....	142
2.6 Astrochemical modelling.....	147
3. Conclusions.....	150
4. Experimental section.....	152
5. References.....	154

**CHAPTER V: Hunting complex cyanides in the interstellar medium.....159**

1. Introduction. ....	160
2. Results and discussion. ....	163
2.1 Rotational and computational study of cyanoacetamide. ....	163
2.2 Rotational study of cyanoacetic acid. ....	173
3. Conclusions. ....	181
4. Experimental section. ....	182
5. References.....	184

**CHAPTER VI: Toward the limits of complexity of interstellar chemistry: rotational spectroscopy and astronomical search of n- and i-butyraldehyde. ....187**

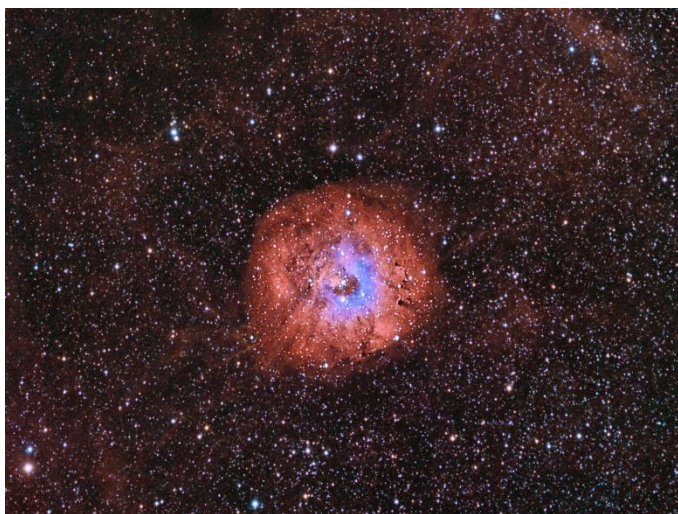
1. Introduction.....	188
2. Results and discussion.....	190
2.1 Molecular modeling.....	190
2.2 Millimeter and submillimeter-wave spectra of n-PrCHO.....	192
2.3 Millimeter and submillimeter-wave spectra of i-PrCHO.....	196
2.4 Search for n- and i-butyraldehyde toward Sgr B2(N1) with ALMA.....	200
2.5 Search for n- and i-butyraldehyde toward G+0.693-0.027 molecular cloud.....	204
2.6 Discussion: Interstellar chemistry of acetaldehyde, propionaldehyde and butyraldehyde.....	206
3. Conclusions. ....	210
4. Experimental section. ....	211
5. References.....	212

<b>CHAPTER VII: The shape of biomolecules, fluxional systems and organic synthons.....</b>	<b>215</b>
CHAPTER VIIa: A deep insight into the shape of L-DOPA.....	215
CHAPTER VIIb: The shape of the archetypical oxocarbon squaric acid and its water clusters.....	233
CHAPTER VIIc: Shape-shifting molecules: Unveiling the valence tautomerism phenomena in barbaralones.....	251
<b>CHAPTER VIII: Conclusions and Future Trends. ....</b>	<b>263</b>
APPENDIX A: Appendix to Chapter IV.....	269
APPENDIX B: Appendix to Chapter V. ....	299
APPENDIX C: Appendix to Chapter VI. ....	307
APPENDIX D: Appendix to Chapter VII. ....	317
APPENDIX I: Amino acetaldehyde conformers: structure and spectroscopic properties.	



## CHAPTER I. ABSTRACT - RESUMEN

---



The hunt for new complex organic molecules (COMs) in the interstellar medium (ISM) based on the synergy between radio astronomy and rotational spectroscopy is among the most coveted targets in current Astronomy and Astrophysics. Recent improvements in sensitivity, bandwidth, and spatial

resolution in instrumentation at existing telescopes, together with the construction of new observational facilities, have opened a new window of possibilities for studying rich astronomical sources, awaiting the detection of entirely new species. Furthermore, understanding their formation pathways and their corresponding intricate chemical networks remains some of the key challenges for Astrochemistry.

In this context, the scientific community is constantly pushing the limits of the complexity of interstellar chemistry, highlighting the recent detection of systems of unparalleled degrees of complexity, such as cyanonaftalene or ethanolamine. Nevertheless, many molecules that are reasonable interstellar candidates remain unexplored even in the laboratory because they are solid and labile molecules with very low vapor pressure, which usually decompose during conventional experiments. To overcome the urgent need for rotational data, in the first part of the present Thesis, we have carried out a combinational theoretical and laboratory rotational study of several COMs as a prerequisite to search for them in the ISM. Hence, we have used a battery of state-of-the-art rotational spectroscopic techniques to characterize the rotational spectrum of several relevant glycine isomers: acetohydroxamic acid and glycolamide; two cyano bearing compounds: cyanoacetamide and cyanoacetic acid; as well as two aldehydes: *n*- and *i*-butyraldehyde, together with a purely computational study of amino acetaldehyde. Moreover, we have explored the potential energy surfaces (PES) on the gas-phase formation of several protonated glycine isomers along with protonated cyanoacetamide, finding for each case a feasible interstellar formation process.

Our precise laboratory measurements have been subsequently used to search for some of the studied molecular systems toward the giant molecular cloud complex Sagittarius B2 (Sgr B2) using different ALMA line surveys and toward the cold molecular cloud G+0.693-0.027 employing IRAM 30-m and Yebes 40-m observations. Regardless of the differences in their physical-chemical properties, both astronomical sources stand among the richest chemical inventories in the ISM. To date, we have not achieved any positive identification of the studied molecules. Nevertheless, the experimental data reported in this thesis will also enable eventual searches for the yet undetected species in the ISM using new and more sensitive astronomical line surveys. These future identifications in space should establish a road map that will help to understand the levels of chemical complexity reached in the ISM, as well as to decipher the link between interstellar chemistry and the rich chemical reservoir found in comets and meteorites.

In the second part of this Thesis, we have taken advantage of the excellent sensitivity and specificity of our laser ablation Fourier transform microwave techniques to unleash the conformational panorama and molecular shape of several organic and biomolecules for the first time. Herein, we report a comprehensive microwave study of the amino acid L-DOPA and unveil the three-dimensional structure of two organic synthons of significant relevance: squaric acid and its water clusters and a “shape-shifting” (fluxional) substituted barbaralone. The latest exhibit an interesting valence tautomerism, which has been conclusively deciphered in the gas phase using rotational spectroscopy. We also analyze the role of the intramolecular interactions that modulate the conformational behavior of these systems using a combination of precise experimental data, such as the analysis of the  $^{14}\text{N}$  hyperfine structure in the case of N-bearing species (i.e., L-DOPA) and quantum chemical computations.



La búsqueda de nuevas moléculas orgánicas complejas (COMs) en el medio interestelar (ISM) basada en la sinergia existente entre la radio astronomía y la espectroscopía de rotación es uno de los objetivos más anhelados dentro de la Astronomía y Astrofísica actual. Las últimas mejoras en sensibilidad, ancho de banda y resolución especial de la instrumentación de los telescopios ya existentes, así como la construcción de nuevas instalaciones observacionales ha abierto una ventana para el estudio de fuentes astronómicas, esperando la detección de especies completamente nuevas. Además, la comprensión de sus procesos y mecanismos de formación, junto con su correspondiente intrincada red química, continúa siendo uno de los retos fundamentales para la Astroquímica.

En este contexto, la comunidad científica se encuentra continuamente sobrepasando los límites de complejidad de la química interestelar, destacando las detecciones recientes de sistemas de un grado de complejidad sin parangón, tales como el cianonaftaleno y la etanolamina. No obstante, muchas de las moléculas que son candidatas interestelares razonables permanecen aún sin explorar incluso en el laboratorio, ya que muchas son sólidos lábiles y de baja presión de vapor, que normalmente se descomponen durante la realización de los experimentos convencionales. Para abordar la necesidad imperiosa de datos de rotación, en la primera parte de la presente Memoria hemos realizado un estudio combinado teórico y de laboratorio de rotación sobre diversos COMs como prerrequisito para su búsqueda en el ISM. Por consiguiente, hemos empleado una batería de técnicas espectroscópicas punteras para caracterizar el espectro de rotación de varios isómeros relevantes de la glicina: el ácido acetohidroxámico y la glicolamida; dos nitrilo-derivados: la cianoacetamida y el ácido cianoacético; además de dos aldehídos de gran tamaño: el *n*- y el *i*-butiraldehído; junto con un estudio puramente computacional del amino acetaldehído. Asimismo, hemos explorado las superficies de energía potencial para la formación de varios isómeros de la glicina protonada, así como de la cianoacetamida protonada, encontrando un proceso posible de formación interestelar para cada sistema.

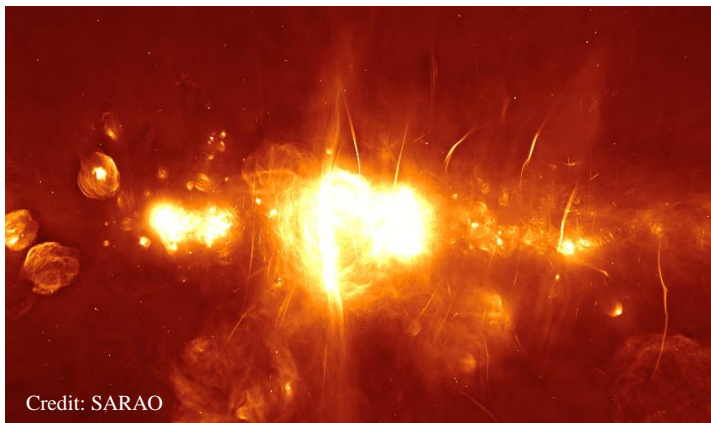
Nuestras medidas precisas de laboratorio se han utilizado posteriormente para buscar algunos de los sistemas moleculares estudiados en el complejo de nubes moleculares gigantes Sagittarius B2 (Sgr B2) empleando diferentes “*line surveys*” registrados con el interferómetro ALMA (Atacama Large Millimeter/submillimeter Array), y en la nube molecular fría G+0.693-0.027 basándonos en observaciones de los radio telescopios IRAM 30-m y Yebes 40-m. Ambas fuentes astronómicas,

independientemente de las diferencias en sus propiedades físico-químicas, se consideran como algunos de los inventarios químicos más ricos del ISM. Hasta la fecha no hemos logrado ninguna identificación positiva de las moléculas bajo estudio. Sin embargo, los datos experimentales presentados en esta Tesis también permitirán realizar eventualmente búsquedas en el ISM de las especies aún sin detectar utilizando nuevos y todavía más sensibles registros astronómicos. Estas futuras identificaciones en el espacio permitirán establecer una hoja de ruta que ayudará a comprender los niveles de complejidad química alcanzados en el ISM, así como a descifrar el vínculo entre la química interestelar y el rico inventario químico que se encuentra en los cometas y meteoritos.

En la segunda parte de esta Tesis, hemos empleado técnicas espectroscópicas de microondas con transformada de Fourier acopladas con ablación laser, de gran sensibilidad y especificidad, para desvelar por primera vez el panorama conformacional y caracterizar la estructura de varias moléculas orgánicas y de interés biológico. En este contexto se presenta un estudio exhaustivo del aminoácido L-DOPA, así como de la estructura tridimensional de dos sintones orgánicos de gran relevancia: el ácido escuárico y sus complejos con agua, y una barbaralona sustituida (molécula fluxional). Este último sistema presenta un interesante tautomerismo de valencia, que ha sido al fin descifrado de manera concluyente en fase gas empleando espectroscopía de rotación. También se ha realizado un análisis del papel que desempeñan las interacciones intramoleculares que regulan el comportamiento conformacional de cada uno de los sistemas. Para ello se ha empleado una combinación de datos experimentales de gran precisión, tales como el análisis de la estructura hiperfina de  $^{14}\text{N}$  para moléculas que contengan nitrógeno (por ejemplo, la L-DOPA), y cálculos mecanocuánticos.

## CHAPTER II.a INTRODUCTION

---



Astrochemistry or Interstellar Chemistry is among the most exciting research fields in modern Physical Chemistry. On the one hand, laboratory Astrochemistry plays a crucial role in the understanding of our molecular universe. It not only

supplements astronomical observations but also assists astrochemists in building up theoretical models (Tielens, 2013). This field comprises several areas of physics relevant to Astronomy and Astrophysics (Savin et al. 2012). Among them, laboratory rotational studies stand out, whose symbiosis with radio astronomy has enabled the detection of many interstellar molecules (Herbst & van Dishoeck 2009, Jorgensen 2020). Nevertheless, the general conditions of the Interstellar Medium (ISM) imply that interstellar chemistry presents several peculiarities, which are different compared to those of Earth. Sometimes, interstellar compounds are very difficult to generate and isolate in the laboratory. Hence, Computational Chemistry could also play a key role in this field. In this Chapter a series of fundamental topics within laboratory Astrophysics and Computational Astrochemistry will be presented, looking forward to accomplishing a first approach to the area. Also, a review of the most interesting and recent bibliographic references will be carried out. We will focus on the latest detections of interstellar molecules and the study and search for glycine, the simplest amino acid, in the ISM.

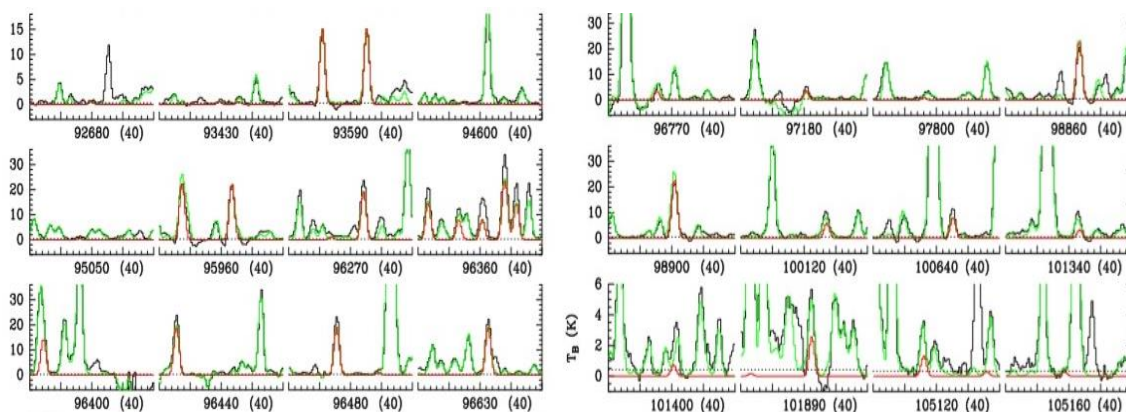
## 1. Background

The present Doctoral Thesis is within the framework of a FPU predoctoral grant (FPU17 / 02987) funded by the “Ministerio de Ciencia, Innovación y Universidades”, associated with the projects of the National Plan CTQ2016-76393-P and PID2019-111396GB-I00. It provided funding to perform theoretical-experimental studies of molecules of astrophysical relevance, as well as to study the conformational behavior and three-dimensional structure of biomolecules during the 2018-2021 period. The work collected in this Dissertation has been carried out in the “Laboratorio de Espectroscopía y Bioespectroscopía” of the “Grupo de Espectroscopía Molecular” (GEM) from the University of Valladolid (UVa), which is an Associated Unit of CSIC, located in the “Edificio Quifima” of the UVa Scientific Park, as well as in the computing facilities of the Computational Chemistry Group (CCG) of the University of Valladolid.

In the last decade, state-of-the-art instrumentation which combines time-domain microwave spectroscopy with supersonic expansions and laser ablation systems to vaporize solid samples has been developed. As it will be explained in Chapter III, these techniques are known by the acronyms LA-MB-FTMW (Laser Ablation Molecular Beam Fourier Transform Microwave) and LA-CP-FTMW (Laser Ablation Chirped Pulse Fourier Transform Microwave). It, along with conventional heating techniques, has enabled the first spectroscopic studies, in particular rotational characterizations in the microwave region, of candidate molecules to be detected in the ISM, as well as the study of several relevant organic and biomolecules under the isolation conditions of the gas-phase, which are granted by the supersonic expansion.

The identification of new molecules in the ISM usually requires a preliminary spectroscopic characterization. In this context, laboratory measurements at low frequencies will allow eventual interstellar searches at centimeter wavelengths carried out with radio telescopes such as the Yebes 40-m (Yebes, Guadalajara, Spain), Effelsberg 100-m (Bad Münstereifel, Germany) and GBT 100-m (Green Bank Telescope, West Virginia, US). Afterward, the experimental data will be extended, if possible, into the millimeter- and submillimeter-wave regions (mm-submm), which will further allow a direct comparison with Line Surveys from different radio observatories. In this regard, the ALMA interferometer (Atacama Large Millimeter Array, Chile) and the IRAM 30-m (Institute for Radio Astronomy in the Millimeter Range, Pico Veleta, España)

radiotelescopes stand out above the rest. In Figure 1, a portion of the ALMA Survey named EMOCA (Exploring Molecular Complexity with ALMA) of Sgr B2(N2) including the detected transitions of **acetaldehyde** ( $\text{CH}_3\text{CHO}$ ) is shown as an example.



**Figure 1.** a) Collection of transitions for the ground vibrational state,  $v=0$ , of acetaldehyde ( $\text{CH}_3\text{CHO}$ ), covered by the ALMA Survey named *EMOCA* of Sgr B2(N2). The synthetic spectrum under the local thermodynamic equilibrium (LTE) approximation of acetaldehyde, employed to derive the upper limit to its column density (in red) is overlaid with the spectrum of Sgr B2(N2) (in black). The synthetic spectrum labeled in green contains the contributions from all the species that have been identified so far in this survey. For each panel, the central frequency and width is labeled in MHz in the x-axis, while the y-axis corresponds to the brightness temperature (K) [Figure adapted from Sanz-Novo et al. (2019)].

In this context, of significant importance is the participation of our Associated Unit CSIC in the European project ERC-SyG-2013 Grant Agreement n. 610256 Nanocosmos, as well as our close collaboration with the Max Planck Institute for Radioastronomy, one of the world's leading institutions in the field of Astrophysics, highlighting its role in the detection and identification of quite a few molecules in the ISM.

Furthermore, concurrent to the experimental studies, an analysis of the possible formation processes in the gas phase - from molecules already detected in the ISM - has been carried out for some of the target systems, along with several high-level computational calculations. The astrochemical interpretation of those should be essential to elucidate the model of the molecular universe (Tielens 2013), the formation pathways of the different interstellar molecular systems, and reveal the chemical processes that may have led to the life's origin.

## 2. *Interstellar medium (ISM): composition and relevance of the grains*

Briefly, the matter between the stars, or interstellar medium (ISM), consists of a diluted mix of ions, atoms, molecules, and dust (about 99 % gas and 1 % dust particles in mass percentage, Spitzer, 1978), which is bathed in an electromagnetic radiation field.

Interstellar gas is mainly composed of hydrogen, the most abundant element in the universe (~70% of the ISM total mass), which can be found in different states. Approximately 60% of interstellar hydrogen is in its atomic form (H), ~20% generates diatomic or molecular hydrogen (H<sub>2</sub>), and the rest is present in its ionized state (H<sup>+</sup>). Based on the properties of interstellar gas and depending on the state of hydrogen, the ISM can be classified in different phases (Draine, 2011): hot ionized médium (HIM), warm ionized médium (WIM), warm neutral médium (WNM), cold neutral medium, (CNM) and, finally, molecular clouds, (MC). In addition, the ISM also involves a minority content of helium and other heavier elements. Together with gas, another of the states of matter that can be found in interstellar space is cosmic dust (Thaddeus, 2006). It plays a crucial role in Interstellar Chemistry, acting in many cases as a catalyst surface for diverse chemical processes, and the regulation of adsorption-desorption processes controls the chemistry present in many regions of space. The composition of the grains is fundamentally based on silicon, carbon and oxygen. Depending on the particular proportions, silicates or carbonaceous grains are formed. Besides the main structure of the grains, there are many other molecular species in the form of ice such as CO, CO<sub>2</sub>, CH<sub>3</sub>OH, NH<sub>2</sub>, CH<sub>2</sub>. During the evolution of the grain, this material will undergo numerous chemical modifications caused by ionic and UV irradiations, as well as thermal effects (Dartois, 2005). Interstellar dust particles may have been a possible origin of prebiotic atoms such as C, H and N (Chyba & Sagan 1992; Chyba & Hand 2005), pointing to the interconnection between Interstellar Chemistry and the terrestrial origin of life (Jorten, 2006).

## 3. *Laboratory Astrochemistry or Astrophysics*

Our understanding of the molecular universe and the Astrochemistry that underlies strongly depends on the advances in Laboratory Astrochemistry or Astrophysics. As it is hinted at its name, this field has a purely astrophysical motivation,

and allows us to complement astronomical searches and to accompanying astrochemical theoretical modeling (Tielens, 2013). Following the review by de Savin et al. (2012), a breakdown of five major categories can be stated: a) Planetary systems and star formation, b) Stellar evolution, c) Structure of the Milky Way and other neighbor galaxies, d) Evolution of galaxies over time e) Cosmology and fundamental physics. Their study entails the development of several areas of physics and physical-chemistry, highlighting atomic, molecular and chemical physics (Kallman & Palmeri 2007; Herbst & van Dishoeck 2009), condensed matter physics (Draine 2003), plasma physics (Remington et al. 2006; Yamada et al. 2010), nuclear physics (Käppeler et al. 2011; Wiescher et al. 2010) and particle physics (Aprile & Profumo 2009).

One of the main goals of Laboratory Astrochemistry or Astrophysics is to detect of prebiotic and biological molecules in diverse astrochemical environments, looking forward to unveil a complete census for each spectral survey (ideally with uniform sensitivity) of a particular astronomical object (Jørgensen 2020). In the present Thesis, the work has been focused primarily on molecular and chemical physics, since (theoretical and experimental) laboratory data based on rotational spectroscopy has allowed the detection of a vast amount of interstellar molecules (Herbst & van Dishoeck 2009, Jørgensen 2020).

In this context, the rotational spectra of **methyl mercaptane**  $\text{CH}_3\text{CH}_2\text{SH}$  (Kolesníková et al. 2014) y **hydromagnesium isocyanide**  $\text{HMgNC}$  [Cabezas et al 2013(a)], **ethyl methyl ether**  $\text{CH}_3\text{CH}_2\text{OCH}_3$  (Tercero et al. 2015) y **methyisocyanate**  $\text{CH}_3\text{NCO}$  (Cernicharo et al. 2016) have been studied in our laboratory at the University of Valladolid, which have enabled their unequivocal identification in the ISM, What is more, the spectra of **acrolein**  $\text{CH}_2\text{CHCHO}$  (Daly et al. 2015), **methyl vinyl ether**  $\text{CH}_2\text{CH}_2\text{OCH}_3$  (Daly y et al. 2014), **vinyl acetate**  $\text{CH}_2\text{CHCO}_3\text{H}$  (Alonso et al. 2015), **methoxyamine** (Kolesníková et al. 2017), **lactaldehyde**  $\text{CH}_3\text{CH}(\text{OH})\text{CHO}$  (Alonso et al. 2019) and **propiolamide**  $\text{HCCC}(\text{O})\text{NH}_2$  (Alonso et al. 2021) among many other species, have been recorded, making for some of them tentative detections.

As we will see in detail in Chapter III (Methodology), besides using conventional heating methods, recently laser ablation sources have been coupled with high-resolution Fourier transform microwave spectroscopic techniques. Many solid compounds with high melting points present very low vapor pressures, being out of the scope of conventional

gas-phase studies in the centimeter- and millimeterwave regions. In the last couple of years, taking advantage of the laser ablation techniques, diverse species of astrophysical relevance have been investigated as a prerequisite for their eventual astronomical identification, such as **hydantoin** (Alonso y et al. 2017), el **hydantoic acid** (Kolesniková et al. 2019) y la **glycinamide**  $\text{CH}_2(\text{NH}_2)\text{C}(\text{O})\text{NH}_2$  (Alonso et al. 2018), a plausible interstellar precursor of glycine.

It is worth noticing that, in order to carry out these projects, a close and thorough collaboration between a very diverse community of laboratory astrophysicists and astrochemists, together with important groups of radio astronomers and theoretical chemists is required. In this regard, Astrochemistry is on the rise and gradually transitioning to a new era of extremely high-precision (experimental or radio astronomical) measurements and high-fidelity chemical modeling (Garrod 2013).

#### 4. *Computational Astrochemistry*

Another of the main goals for Astrochemistry is the elucidation of the formation and transformation mechanisms of interstellar molecular systems (Vazart y col. 2016), which are still to date an object of debate. Consequently, a new research field known as Computational Astrochemistry arises, which implements quantum chemical methods to study systems of astrochemical relevance (Biczysko y col. 2018). This discipline has become a very useful tool to face different chemical problems from a global and rational perspective (Barone y col. 2015). Moreover, the field of Interstellar Chemistry constitutes a perfect medium for applying theoretical methods, and it is also beneficial when the experimentation is directly not accessible.

The chemical complexity and diversity of the distinct astronomical sources can be ascribed to gas, grain and gas-grain interactions. Taken as a whole, we will be able to build intricate chemical networks that encompass thousands of chemical processes and reactions. Thus, complex astrochemical models will be assembled, such as the *MAGICKAL* model (*Model for Astrophysical Gas and Ice Chemical Kinetics and Layering*, Garrod 2013), mentioned in different chapters of the present thesis.

It is well known that the study of surface processes has recently brought the attention of many experimental and theoretical researches (Willis & Garrod 2017),



looking forward to comprehend the composition of the dust-grain mantles, which are generated by accretion and accumulation of gaseous species in the grains. These processes are in some cases an alternative to gas-phase reactions, and explain the formation of some interstellar molecules in icy mantles, such as **methyl formate** ( $\text{CH}_3\text{OCHO}$ ). However, **dimethyl ether** ( $\text{CH}_3\text{OCH}_3$ ) and **formic acid** ( $\text{HCOOH}$ ) are most likely originated from both grain-surface and gas-phase chemistry (Garrod & Herbst 2006). Nevertheless, gas-phase ion-molecule reactions still play a fundamental role in synthesizing countless interstellar organic molecules (Petrie et al. 2007, Larsson et al. 2012, Geppert & Larsson 2013; Barrientos et al. 2012; Redondo et al. 2014, 2015, 2017). The Computational Chemistry Group (CCG) from the University of Valladolid, under the direction of prof. Antonio Largo has been studying the formation of astrophysical relevant species for more than a decade, highlighting those that target the formation of **glycine** under interstellar conditions ( $\text{NH}_2\text{CH}_2\text{COOH}$ ) (Barrientos y col. 2013, Redondo y col. 2014, 2015, 2017).

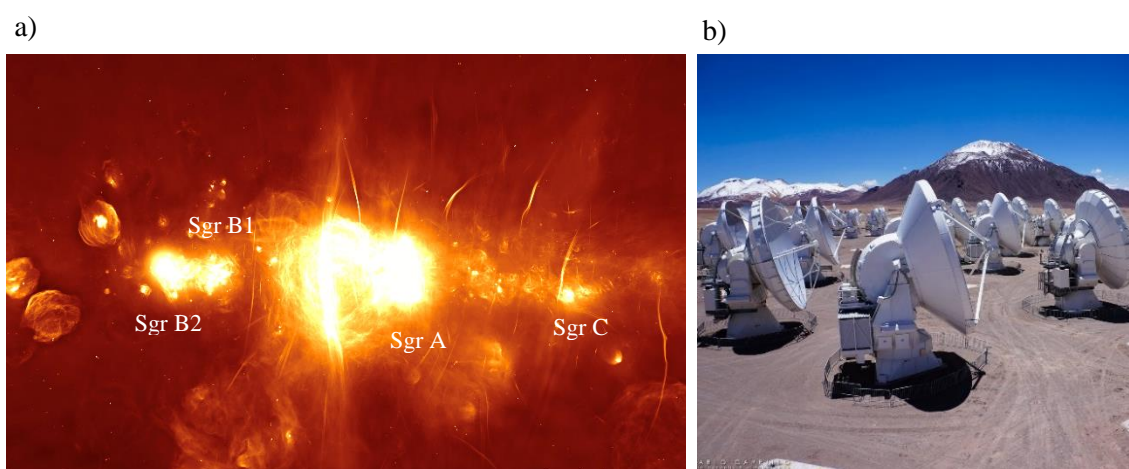
Given the latest advances in Theoretical and Computational Chemistry, it is now possible to perform mechanistic, structural and thermodynamic studies, as well as to predict spectroscopic properties of middle-sized molecules with an outstanding precision (Barone y col. 2015). Among all theoretical approaches, Coupled Cluster Methods and hybrid models stand out above the rest (see [Chapter III](#), Methodology). Finally, it is worth noticing that purely theoretical data have enabled to guide astronomical observations and detections [Cernicharo y col. 2021 (a)], and also to suggest the search for new and undetected molecular systems (Puzzarini y col. 2014), which are essential to rationalize the origin of large interstellar molecules that are in some cases fundamental ingredients of life as we know it on Earth.

### 5. *Molecules detected in the ISM: complex organic molecules (COMs)*

In the last years, the number of molecules detected in the ISM and in circumstellar envelopes, around 250, has grown almost exponentially by virtue of the great effort that has been made by the scientific community [see McGuire 2018 (a) for a global census].

Among the families of molecules of greater astrochemical relevance we find the so-called complex organic molecules (COMs), defined as species containing six or more atoms including at least one carbon atom (Jørgensen 2020). Most of these compounds

were detected for the first time in the millimeter and submillimeter region in hot molecular cores (HMCs,  $T \approx 100\text{-}300\text{ K}$ ), such as Sagittarius B2 (N) [see Figure 2 (a)] and Orion-K (Herbst & van Dishoeck 2009). Afterward, astronomical studies have been extended to different astronomical environments such as cold gas clouds ( $T \approx 10\text{ K}$ ) of pre-stellar cores (Vastel et al. 2014), hot corinos, that is, the lower mass analogues of hot cores (Sakai & Yamamoto 2001, Maury et al. 2014), protoplanetary disks (Walsh et al. 2016, Bergner et al. 2018), and comets (Altwegg et al. 2016), among others.



**Figure 2.** a) Galaxy center and Sgr B2 ( $\sim 390$  light years away), as seen from the MeerKat interferometer [Image credit: MeerKat, SARA0]; b) ALMA project, Atacama desert (Chile) [Image credit: NRAO].

Nowadays, the detection of new complex systems in ISM has been greatly facilitated by the development of modern radio telescopes, together with the use of interferometric methods (Wilson et al. 2013). Radio telescopes allow us to detect and analyze the radiation emitted in the micro- and radio wavelengths, where transitions take place between the different rotational states of the molecules. We find different radio astronomical facilities of interest, remarking the ALMA Project (Atacama Large Millimeter Array) in the Atacama desert (Chile) as one of the most promising observatories [see Figure 2 (b)]. The identification of a molecular species in space usually implies a direct comparison of the recorded emission frequency lines (rest frequencies) with spectroscopic measurements obtained from theoretical information (quantum chemical calculations) and (mainly) experimental data recorded in the laboratory. In general, with a dozen positive rotational identifications (fingerprints) of unblended lines,

a molecule is considered as unequivocally identified in the ISM. Although, new techniques are being used recently, highlighting the so-called Line Stacking approach (Loomis et al. 2021, Lee et al. 2021). These methodologies increase the signal to noise ratio by obtaining an average spectrum that encapsulates the total information of all the observed lines instead of examining each of the lines individually (McGuire et al. 2021). It is very useful for certain surveys of cold dark molecular cloud such as TMC-1 (when the survey is not close to the so-called confusion limit). Moreover, astronomical detections and discoveries of new species usually drive the selection of novel molecular targets.

The development of the modern radio astronomical facilities in the last decades, i.e. the construction of new broadband receivers such as those of the Yebes-40m radio telescope (18+ GHz bandwidth), built in the framework of the Nanocosmos Project, has enabled to acquire spectral surveys of a resolution and sensitivity so high - between 0.25 and 1 mK for the  $Q$  band in the *QUIJOTE* Survey [*Q-band Ultrasensitive Inspection Journey to the Obscure TMC-1 Environment*, Cernicharo 2021(a)] – that plenty of spectral lines have not been yet assigned (Unidentified lines or U-lines). In many cases, this fact is due to the existence of Line Surveys that are very close to the already mentioned “confusion limit”. These rotational transitions may belong to new but unidentified molecules, but could also be attributed to vibrationally excited states of a few molecules which are called the “astrophysical weeds” (Goldsmith et al. 2006; Fortman et al. 2010, Kolesníková et al. 2017), or even to conformers or isotopologues of already detected molecules, whose rotational spectrum is still unknown. This fact is crucial when astronomical observations imply hot regions of the ISM, such as sources related to star-forming regions (hot cores... etc.). New European and American observatories are expected to increase the known chemical reservoir of our molecular universe dramatically.

In Table 1 we present a list of the detected molecules, to date, in the interstellar and circumstellar medium, which was obtained from the Cologne Database for Molecular Spectroscopy<sup>1</sup> (CDMS).

**Table 1.** List of molecules detected in the ISM and in circumstellar envelopes.

2 atoms		3 atoms		4 atoms		5 atoms		10 atoms	
H <sub>2</sub>	SH <sup>+</sup>	H <sub>3</sub> <sup>(*)</sup>	OCS	CH <sub>3</sub>	C <sub>3</sub> S	CH <sub>4</sub> <sup>*</sup>	CH <sub>3</sub> COCH <sub>3</sub>		
AlF	SH	CH <sub>2</sub>	SO <sub>2</sub>	NH <sub>3</sub>	HMgNC	CH <sub>2</sub> NH	(CH <sub>2</sub> OH) <sub>2</sub>		
AlCl	HCl <sup>+</sup>	NH <sub>2</sub>	<i>c</i> -SiC <sub>2</sub>	H <sub>3</sub> O <sup>+</sup>	HCCO	SiH <sub>4</sub>	CH <sub>3</sub> CH <sub>2</sub> CHO		
C <sub>2</sub> <sup>**</sup>	TiO	H <sub>2</sub> O	SiCN	C <sub>2</sub> N <sub>2</sub>	CNCN	H <sub>2</sub> COH <sup>+</sup>	CH <sub>3</sub> C <sub>5</sub> N		
CH	ArH <sup>+</sup>	C <sub>2</sub> H	C <sub>2</sub> P	H <sub>2</sub> CN	HONO	<i>c</i> -C <sub>3</sub> H <sub>2</sub>	<b>HC<sub>3</sub>S<sup>+</sup></b>		
CH <sup>+</sup>	N <sub>2</sub>	HCN	AlNC	HCNH <sup>+</sup>	MgC <sub>2</sub> H	<i>l</i> -C <sub>3</sub> H <sub>2</sub>	CH <sub>3</sub> OCH <sub>2</sub> OH		
CN	NO <sup>+</sup> ?	HNC	AlOH	H <sub>2</sub> CO	<b>HCCS</b>	CH <sub>2</sub> CN	<b>H<sub>2</sub>CCCHC<sub>3</sub>N</b>		
CO	NS <sup>+</sup>	HCO	H <sub>2</sub> O <sup>+</sup>	PH <sub>3</sub> ?	<b>HNCN</b>	H <sub>2</sub> C <sub>2</sub> O	<b>C<sub>2</sub>H<sub>5</sub>NCO</b>		
CO <sup>+</sup>	HeH <sup>+</sup>	HCO <sup>+</sup>	CaNC	C <sub>3</sub> H <sup>*</sup>	<b>H<sub>2</sub>NC</b>	H <sub>2</sub> CCN	<b><i>c</i>-C<sub>6</sub>H<sub>4</sub></b>		
CP	CS	HOC <sup>+</sup>	<b>NCS</b>	<i>c</i> -C <sub>3</sub> H	<b>HCCS<sup>+</sup></b>	NH <sub>2</sub> CN			
SiC	HF	N <sub>2</sub> H <sup>+</sup>	NaCN	HC <sub>2</sub> N		HCOOH	<b>11 atoms</b>		
HCl	HD	HNO	C <sub>2</sub> S	HNCO		C <sub>4</sub> H	C <sub>2</sub> H <sub>5</sub> OCHO		
KCl	FeO?	H <sub>2</sub> S	H <sub>2</sub> Cl <sup>+</sup>	HCNO		C <sub>4</sub> H	CH <sub>3</sub> C <sub>6</sub> H		
NH	O <sub>2</sub>	HCP	KCN	HNCN?		HC <sub>3</sub> N	HC <sub>9</sub> N		
NO	CF <sup>+</sup>	CO <sub>2</sub>	HSC	HCO <sub>2</sub> <sup>+</sup>		HC <sub>2</sub> NC	CH <sub>3</sub> OC(O)CH <sub>3</sub>		
NS	SiH?	HCS <sup>+</sup>	NCO	H <sub>2</sub> CS		HNC <sub>3</sub>	CH <sub>3</sub> C(O)CH <sub>2</sub> OH		
NaCl	PO	C <sub>2</sub> O	HO <sub>2</sub>	C <sub>3</sub> H <sup>+</sup>		CNCHO	<b><i>c</i>-C<sub>5</sub>H<sub>6</sub></b>		
OH	AlO	C <sub>3</sub> <sup>*</sup>	C <sub>2</sub> N	C <sub>3</sub> N		C <sub>5</sub> <sup>*</sup>	<b>HOCH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub></b>		
PN	OH <sup>+</sup>	MgNC	Si <sub>2</sub> C	C <sub>3</sub> N <sup>-</sup>		SiC <sub>4</sub>			
SO	CN <sup>-</sup>	MgCN	HS <sub>2</sub>	C <sub>3</sub> O		CH <sub>3</sub> Cl	<b>12 atoms</b>		
SO <sup>+</sup>	SiO	FeCN	NCO	HNCS		NCCNH <sup>+</sup>	<b><i>c</i>-C<sub>6</sub>H<sub>6</sub><sup>*</sup></b>		
SiN	SiS	N <sub>2</sub> O	HNO	<i>c</i> -SiC <sub>3</sub>		<b>C<sub>4</sub>S</b>	<i>n</i> -C <sub>3</sub> H <sub>7</sub> CN		
						<b>HC<sub>3</sub>S<sup>+</sup></b>	C <sub>2</sub> H <sub>5</sub> OCH <sub>3</sub>		
	<b>6 atoms</b>	<b>7 atoms</b>	<b>8 atoms</b>	<b>9 atoms</b>		<b>H<sub>2</sub>C<sub>2</sub>S</b>	<i>i</i> -C <sub>3</sub> H <sub>7</sub> CN		
	C <sub>2</sub> H <sub>4</sub>	CH <sub>3</sub> NH <sub>2</sub>	CH <sub>2</sub> CHCHO	CH <sub>2</sub> CHCH <sub>3</sub>		<b>HC(O)SH</b>	<b>1-<i>c</i>-C<sub>2</sub>H<sub>5</sub>CN</b>		
	CH <sub>3</sub> OH	CH <sub>3</sub> C <sub>2</sub> H	NH <sub>2</sub> CH <sub>2</sub> CN	CH <sub>3</sub> OCH <sub>2</sub>		<b>HC(O)SH</b>	<b>2-<i>c</i>-C<sub>2</sub>H<sub>5</sub>CN</b>		
	CH <sub>3</sub> CN	CH <sub>3</sub> CHO	CH <sub>3</sub> COOH	CH <sub>3</sub> CH <sub>2</sub> CN					
	CH <sub>3</sub> NC?	<i>c</i> -C <sub>2</sub> H <sub>4</sub> O	CH <sub>2</sub> OHCHO	CH <sub>3</sub> CH <sub>2</sub> OH					
	CH <sub>2</sub> CNH	CH <sub>2</sub> CHOH	HC(O)OCH <sub>3</sub>	<b>H<sub>2</sub>C<sub>3</sub>S</b>			<b>&gt; 12 atoms</b>		
	NH <sub>2</sub> CHO	CH <sub>2</sub> CHCN	CH <sub>3</sub> C <sub>3</sub> N	<b>HC(O)SH</b>			C <sub>60</sub> <sup>*</sup>		
	CH <sub>3</sub> SH	C <sub>6</sub> H	CH <sub>2</sub> CCHCN	<b>HC(S)CN</b>			C <sub>70</sub> <sup>*</sup>		
	CH <sub>3</sub> SH	C <sub>6</sub> H <sup>-</sup>	C <sub>7</sub> H	C <sub>8</sub> H			C <sub>60</sub> <sup>+</sup> *		
	C <sub>4</sub> H <sub>2</sub>	HC <sub>3</sub> N	C <sub>6</sub> H <sub>2</sub>	C <sub>8</sub> H <sup>-</sup>			<i>c</i> -C <sub>6</sub> H <sub>5</sub> CN		
	HC <sub>4</sub> N	CH <sub>3</sub> NCO	H <sub>2</sub> C <sub>6</sub>	HC <sub>7</sub> N			HC <sub>11</sub> N		
	HC <sub>3</sub> NH <sup>+</sup>	HC <sub>3</sub> O	CH <sub>3</sub> CHNH	CH <sub>3</sub> NHCHO			<b>1-C<sub>19</sub>H<sub>7</sub>CN</b>		
	HC <sub>2</sub> CHO	HOCH <sub>2</sub> CN	CH <sub>3</sub> SiH <sub>3</sub>	HC <sub>7</sub> O			<b>2-C<sub>10</sub>H<sub>7</sub>CN</b>		
	<i>c</i> -H <sub>2</sub> C <sub>3</sub> O	HCCCHNH	H <sub>2</sub> NCONH <sub>2</sub>	<b>HC<sub>3</sub>HCHCN</b>			<b><i>c</i>-C<sub>9</sub>H<sub>8</sub></b>		
	C <sub>5</sub> H	HC <sub>4</sub> NC	HCCCH <sub>2</sub> CN	<b>H<sub>2</sub>CCHC<sub>3</sub>N</b>					
	C <sub>5</sub> N	<b><i>c</i>-C<sub>3</sub>HCCH</b>	HC <sub>5</sub> NH <sup>+</sup>	<b>H<sub>2</sub>CCHCCH</b>					
	C <sub>5</sub> N <sup>-</sup>	<b><i>l</i>-H<sub>2</sub>C<sub>5</sub></b>	<b>CH<sub>2</sub>CHCCH</b>						
	SiH <sub>3</sub> CN	<b>MgC<sub>5</sub>N</b>	<b>MgC<sub>6</sub>H</b>						
	CH <sub>3</sub> CO <sup>+</sup>	<b>CH<sub>2</sub>C<sub>3</sub>N</b>	<b>C<sub>2</sub>H<sub>3</sub>NH<sub>2</sub></b>						
	<b>C<sub>3</sub>H<sub>3</sub></b>								
	<b>H<sub>2</sub>C<sub>3</sub>S</b>								
	<b>HCCCHS</b>								

**Note.** All the chemical species that were detected in 2021 are highlighted in bold.

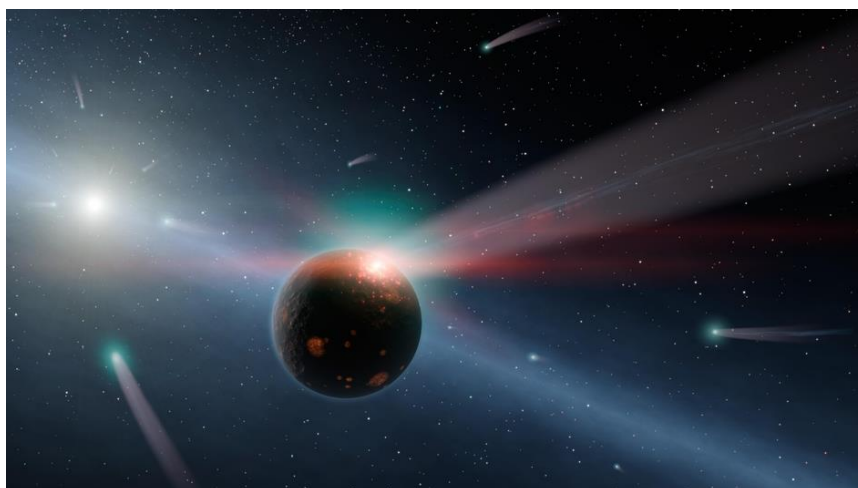
### 5.1. *Origin of life: “glycine, the missing molecule”*

Following on the previous point, the distribution of organic molecules of interest, in particular amino acids, in the universe could help to achieve a better understanding of life’s origin (Woon 2002). The chemistry of amino acids in the ISM, as well as in comets and meteorites is one of the most intriguing topics for Extraterrestrial Organic Chemistry and Astrobiology. Amino acids are usually considered the “building blocks of life”, since they are essential components of proteins and, therefore, indispensable constituents of living systems. So far, one of the indisputable evidences about the existence of amino acids outside our planet is their detection in carbonaceous chondrites (Cronin y et al. 1976, 1983, 1989). It is believed that amino acids have been shipped by meteorites and other small celestial bodies (comets, asteroids and interplanetary dust particles). Their origin is key to understanding how life emerged on the early Earth 3.5-3.7 billion years ago (see Figure 3).

The bombardment of the early Earth by comets and asteroids entailed the appearance of a large amount of organic material (Aponte et al. 2019). This reservoir of organic compounds, such as amino acids, could be of great relevance to astrobiology, as Earth-like planets located in extrasolar nebulae are likely to experience similar impacts. Therefore, the products of Extraterrestrial Organic Chemistry could, in principle, allow prebiotic chemistry to occur throughout the whole galaxy (Chyba et al. 1990).

---

<sup>1</sup>The complete list of detected molecules in the interstellar and circumstellar medium, together with the list of extragalactic molecules is available via <http://www.astro.uni-koeln.de/cdms/molecules>



**Figure 3.** Illustration of a storm of comets around Eta Corvi, a star located near our Sun, where a comet is depicted crashing into a rocky planet [Image credit: NASA/JPL-Caltech]

Despite the extensive search for **glycine** ( $\text{NH}_2\text{CH}_2\text{COOH}$ ), to date, the simplest amino acid has not been conclusively identified in the ISM. In 1994, a team of astronomers from the University of Illinois, under the direction of Lewis Snyder, claimed the detection of glycine molecules in space (Mehring et al. 1995). However, this detection was not confirmed by any other astrophysicists. Afterward, Kuan (2003) reported once more its detection based on the assignment of 27 rotational lines. This work was later questioned and finally discarded due to the wrong identification of the frequency lines. The viability to detect glycine in the ISM is strongly hampered because of the large size of the molecule. Therefore, weak rotational lines are expected due to their large molecular partition function. Thus, some of the targeted lines may be hindered by the emission of other species. In this context, the determination of accurate spectroscopic parameters of their structural isomers and the subsequent identification in the ISM will ultimately mean a deep cleaning of the corresponding Line Survey, making the detection of new chemical species accessible.

Our knowledge about the different formation pathways of glycine in space remains unexplained (Ruiz-Miraz et al. 2014). Two widely discussed proposals are found in the literature to simulate its synthesis in the laboratory. The first one is the well-known Strecker amino acid synthesis (Strecker 1850), and the second one is the Miller experiment (Miller 1953). In both cases, the key intermediate is **formaldehyde** ( $\text{HCHO}$ ).

Another suitable reactant, similar to formaldehyde, is **methanimine** ( $\text{CH}_2=\text{NH}$ ), an isoelectronic species compared to HCHO that is also suggested as a plausible prebiotic precursor of glycine (Godfrey et al. 1973; Salter et al. 2008; Koch et al. 2008).

In this context, ion-molecule reactions are a significant pathway for the formation of organic molecules in the gas-phase (Ferguson 1975, Herman & Futrell 2015), acting in multiple cases as a possible source to explain the formation of such species in the ISM (Petrie et al. 2007, Larsson et al. 2012, Geppert & Larsson 2013). Several years ago, a group under the supervision of prof. Bohme studied the gas-phase synthesis of glycine in detail. It was concluded that the following reactions allowed the formation of protonated glycine and water (Snow et al. 2007, Blagojevic et al. 2003).

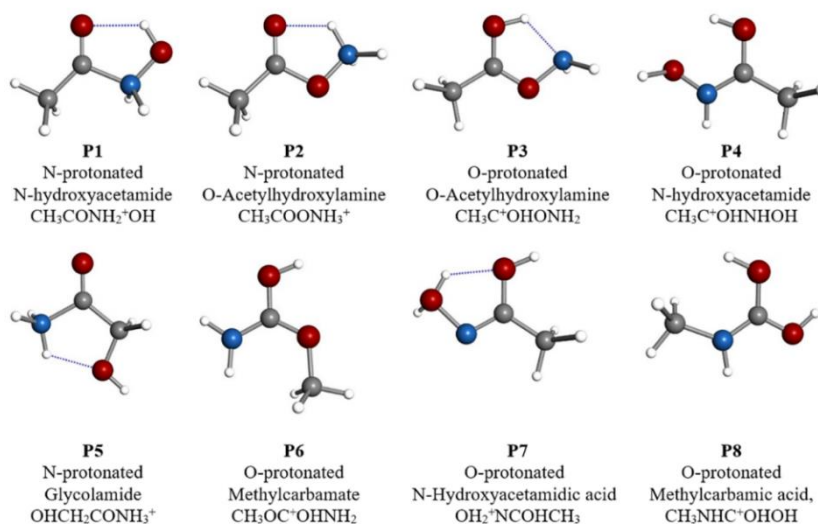


As mentioned above, the detection of glycine in the ISM was discarded after a heated debate within the astrophysical community (Snyder et al. 2005, Cunningham et al. 2007, Jones et al. 2007). However, the recent Rosetta mission to comet 67P/Churyumov-Gerasimenko reported glycine's presence along with other prebiotic molecules (Altwegg et al. 2016). Moreover, its generation via ion-molecule reactions is still a topic of significant astrophysical interest and should function as a model for the formation of other COMs.

On the other hand, several theoretical and experimental studies have suggested the formation of glycine in interstellar ices (Holtom et al. 2005) or via radical molecules (Rimola et al. 2012), varying the different reactions mechanisms depending on the composition of the ice (Elsila et al. 2007). It has usually been suggested that UV radiation is needed to initiate these reactions (Woon et al. 2002). It must be taken into account that condensed-phase reactions could include more than two reagents, as Maeda and Ohno (2004) pointed out for the formation of glycine via reaction of  $\text{NH}_3 + \text{CH}_2 + \text{CO}_2$ .

At the end of 2017, Largo et al. (2018) investigated the ion-molecule reactions 1 and 2 by molecular dynamics simulations, which were experimentally explored by Bohme and coworkers. Bohme characterized all the reaction products of mass-charge ratio  $m/z$  76 by mass spectrometry. Then, the fragmentation spectra at different energies

were compared to a glycine standard sample, where certain similarities were found. Largo et al. (2018) demonstrated that both tautomers of **protonated hydroxylamine** ( $\text{NH}_3\text{OH}^+$  y  $\text{NH}_2\text{OH}_2^+$ ), whose neutral analog has been recently detected in the ISM (Rivilla et al. 2020), enabling the formation of eight different molecular species (**P1-P8**) with the same mass as protonated glycine  $[\text{H}_6\text{C}_2\text{O}_2\text{N}]^+$  but different structures (see Figure 4). The product **P1** was found to be the most abundant species for both reaction pathways.



**Figure 4.** Distinct ions of  $m/z$  76 generated by collisions of  $\text{NH}_3\text{OH}^+ + \text{CH}_3\text{COOH}$ .

Among all the related neutral isomers studied to date, methylcarbamate (the unprotonated neutral counterpart of P6) has been the only one searched for in the hot molecular cloud W51e2, associated with high-mass protostar, and the intermediate-mass protostar IRAS21391+58502 (Demyk et al. 2004). Groner et al. (2007) studied the millimeter- and submillimeter-wave spectrum of methyl carbamate, providing rotational data in wide frequency ranges. Finally, glycolamide (the unprotonated neutral counterpart of P5) has been characterized by Stark modulation spectroscopy between 60 and 78.3 GHz (Maris 2004). To the best of our knowledge, neither structural nor spectroscopic information is available for the rest of the neutral species.

In the present Doctoral Thesis, a computational study on the potential energy surface (PES) corresponding to the reactions between **acetic acid** and the two **protonated hydroxylamine tautomers** has been carried out and is presented in the first section of Chapter IV. Eight different products have been studied, all of them obeying the molecular



formula  $[\text{H}_6\text{C}_2\text{O}_2\text{N}]^+$ , and, surprisingly, only one exothermic process with no net activation barrier was discovered, leading to the formation of protonated **N-hydroxyacetamide** or **acetohydroxamic acid**. Hence, its formation under interstellar conditions is expected to be feasible, and, therefore, the corresponding neutral counterpart, **acetohydroxamic acid** ( $\text{CH}_3\text{C}(\text{O})\text{NHOH}$ ), was suggested as a candidate molecule to be detected in the ISM. However, there is scarce experimental information for this molecular system, remaining unexplored by rotational spectroscopy. Thus, we proposed a first rotational study in the microwave region to enable its eventual astronomical search in the ISM. Finally, a comprehensive rotational study of **glycolamide** ( $\text{CH}_2(\text{OH})\text{C}(\text{O})\text{NH}_2$ ) up to 460 GHz, another important isomer of glycine, is presented together with its interstellar search in Sgr B2 using the *EMoCA* Alma Survey and the Effelsberg-100m radio telescope.

## 5.2. Cyanides in the interstellar medium

Among all COMs, cyano-bearing compounds, also referred to as nitriles, are organic molecules containing a cyano functional group (-CN), being among the most widely distributed and best-known species in the ISM. Since the detection of the diatomic cyano radical, CN, in the early 40's (McKellar 1940), which was just the second molecule to be identified in the ISM, more than 30 complex cyanide species -from an astronomical point of view- have been found in diverse types of astronomical environments. Of great relevance was the detection of the branched **i-propyl cyanide** or **isobutyronitrile** ( $\text{i-C}_3\text{H}_7\text{CN}$ ) (Belloche et al. 2014), **glycolonitrile** ( $\text{HO-CH}_2\text{-CN}$ ) (Zeng et al. 2019), **benzonitrile** ( $\text{c-C}_6\text{H}_5\text{CN}$ ), the first benzene-derived aromatic molecule in the ISM (McGuire et al. 2018), **cyanocyclopentadiene**, its five-membered ring analog (McCarthy et al. 2021) as well as **isocyanogen** ( $\text{CNCN}$ ), the simplest member of the **dicyanopolyynes** family (Agúndez et al. 2018). Also very recently, several -CN bearing molecules, highlighting  $\text{HC}_4\text{NC}$ ,  $\text{HCCH}_2\text{CN}$ ,  $\text{HC}_5\text{NH}^+$ ,  $\text{H}_2\text{CCHC}_3\text{N}$  and  $\text{HCCCHCHCN}$ , have been detected towards the prototypical cold dark molecular cloud TMC-1 (Cernicharo et al. 2020, McGuire et al. 2020, Marcelino et al. 2020 and Lee et al. 2021), together with the 1- and 2- isomers of **cyanonaftalene** (McGuire et al. 2021). However, many simple cyanides have yet to be discovered in the ISM. Consequently, new experimental studies exploring related species should represent an essential source of reference data in constraining models on the formation of -CN and -C(O)OH bearing

molecules in the ISM and predicting which other COMs might be lurking in the expanses of space, waiting to be detected.

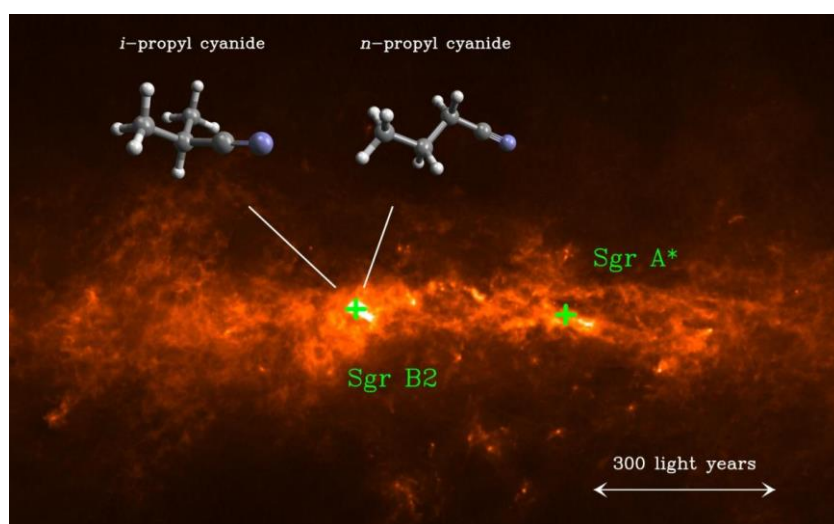
In this regard, **Chapter V** of the present Dissertation gathers the first spectroscopic study in the microwave region of two complex cyanides, which are candidates to be detected in the ISM: **cyanoacetamide** ( $\text{CH}_2(\text{CN})\text{C}(\text{O})\text{NH}_2$ ) and **cyanoacetic acid** ( $\text{CH}_2(\text{CN})\text{C}(\text{O})\text{OH}$ ). A set of accurate spectroscopic rotational parameters have been determined for each molecular system, including the nuclear quadrupole coupling constants attributed to the presence of one (or more)  $^{14}\text{N}$  nuclei. It is worth noticing that a precise resolution of the hyperfine structure will be fundamental to enable the unequivocal identification of a molecule in the ISM in case that low-frequency surveys (in the centimeter wavelengths) are employed. Moreover, a theoretical study on the formation of **protonated cyanoacetamide** has been proposed, and a feasible gas-phase ion-molecule reaction (under the conditions of the ISM) has been found.

### *5.3. Aldehydes and branched-chain species in the interstellar medium*

The aldehyde family embraces some of the most widely distributed species in nature, being precursors of a vast number of prebiotic and biomolecules. Fifty years ago, Snyder et al. (1969) reported the detection of formaldehyde ( $\text{H}_2\text{CO}$ ), the first aldehyde ever detected in space, in diverse galactic and extragalactic environments based on NRAO 140-ft observations. Thereafter, several aldehydes of increasing complexity levels were discovered in the ISM, remarking those detections toward the giant molecular cloud complex Sagittarius B2 (Sgr B2), located approximately 120 parsec (390 light years) away from the center of our Milky Way. It must be remembered the astronomical identification of **acetaldehyde** ( $\text{CH}_3\text{CHO}$ , Gottlieb 1973), **glycolaldehyde** ( $\text{CH}_2(\text{OH})\text{CHO}$ , Hollis et al. 2000), the simplest sugar-related molecule, **propionaldehyde** ( $\text{CH}_3\text{CH}_2\text{CHO}$ , Hollis et al. 2004) and **cyanoformaldehyde** ( $\text{CNCHO}$ , Remijan et al. 2008). Nevertheless, many complex aldehydes, such as **lactaldehyde** ( $\text{CH}_3\text{CH}(\text{OH})\text{CHO}$ , Alonso et al. 2019), are still unidentified in the ISM.

In current astronomical research, interstellar observations of novel molecular systems have usually driven the selection of new targets to study. Of paramount relevance was the radio astronomical identification of the abovementioned ***i*-propyl cyanide** (see Figure 5), the first interstellar aliphatic branched-chain molecule (Belloche et al. 2014).

Previously, Belloche et al. (2009) reported the detection of its straight-chain isomer, ***n*-propyl cyanide** ( $n\text{-C}_3\text{H}_7\text{CN}$ ,  $\text{CH}_3\text{CH}_2\text{CH}_2\text{CN}$ ), in Sgr B2 using the IRAM-30m radio telescope. The identification of new branched-chain species in the ISM will eventually decipher the link between the molecular inventory of the Milky Way and the chemical composition of small celestial objects such as comets (Altwegg et al. 2016), asteroids, as well as meteorites, their rocky remnants that land on Earth (Pizzarello 2006; Pizzarello et al. 2010; Burton et al. 2012), where several branched-chain amino acids have been found (Koga & Naraoka 2017).



**Figure 5.** Dust emission in a combination of data obtained with the APEX telescope and the Planck space observatory. The branched iso-propyl cyanide (left) as well as its linear isomer normal-propyl cyanide (right) were both detected in the star-forming region Sgr B2 using ALMA. [Image Credit: MPIfR/A. Weiss (background image), University of Cologne/M. Koerber (molecular models), MPIfR/A. Belloche (montage)]

Moreover, given the critical development of the new observational facilities, the detection of molecular systems of unparalleled size and complexity is expected to grow even beyond ***i*-propyl cyanide** (Belloche et al. 2014). This fact is mightily reinforced by some of the latest astronomical detections, highlighting the identification of the already mentioned **cyanonaftalene** isomers (McGuire et al. 2021), together with the observation of **indene** [Cernicharo et al. 2021 (b); Burkhardt et al. 2021] and **ethanolamine** ( $\text{HOCH}_2\text{CH}_2\text{NH}_2$ , Rivilla y et al. 2021).

In this context, in [Chapter VI](#) of the present Dissertation we compile a complete spectroscopic rotational study of two aliphatic aldehydes of great astrophysical interest and extraordinary complexity: **normal-butyraldehyde** ( $n\text{-C}_4\text{H}_8\text{O}$ ;  $\text{CH}_3\text{CH}_2\text{CH}_2\text{CHO}$ ) and its branched **iso-butyraldehyde** ( $i\text{-C}_4\text{H}_8\text{O}$ ;  $(\text{CH}_3)_2\text{CHCHO}$ ). Both species are suggested as candidates to be detected in the ISM and have already been identified in several chondritic meteorites (Aponte et al. 2019). In this Doctoral Thesis, we have carried out a comprehensive rotational characterization in the millimeter and submillimeter-wave region for both species (up to 325 GHz). Hence, the analysis of their rotational spectra allowed us to perform a confident search for both species in different regions of the ISM. In addition, we present in [Appendix I](#) a purely theoretical study of **amino acetaldehyde**, another appealing aldehyde for interstellar search, whose experimental rotational study remains, to date, unfeasible.

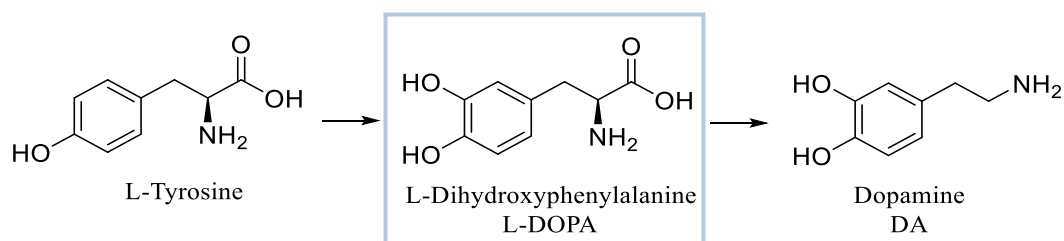
#### *6. The shape of biomolecules, fluxional systems and organic synthons*

The natural or biological medium in which the chemistry of biomolecules unfolds corresponds to the condensed phases. However, both its conformational panorama and its physiological behavior are ruled by a balance of intra- and intermolecular forces, which strongly depends on the nature of the environment itself. There are many spectroscopic studies of biomolecules in condensed phases, highlighting those based on nuclear magnetic resonance (NMR), infrared (IR) or Raman, among others. However, these works hardly provide information on the system under isolation conditions, that is, in the absence of intermolecular interactions with the environment surrounding it.

To undertake this problem, the synergy between gas-phase experiments based on rotational spectroscopy, highlighting those that combine laser ablation (LA) with Fourier transform microwave spectroscopy (FTMW), and theoretical calculations have proven to be one of the most robust methodologies to investigate the structural properties of biomolecules (Alonso & López 2015). This fact is mainly attributed to the gas-phase isolation conditions achieved through a supersonic expansion (Levy 1980; Scoles 1988). The experimental methodology regarding rotational spectroscopy is explained in detail in [Chapter III](#).

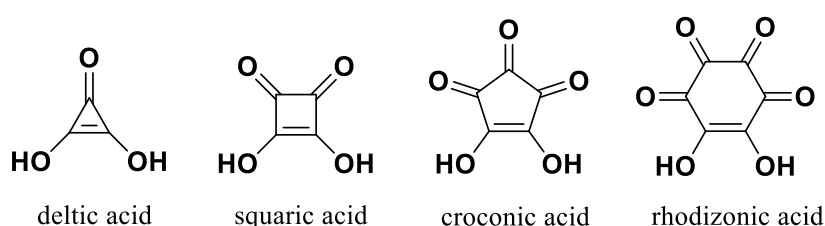
Moreover, the combination of laser ablation vaporization techniques with rotational spectroscopy has led to the study of numerous solid compounds with high melting points. This fact has allowed the structural characterization of most of the natural amino acids (Alonso & López 2015, Alonso et al. 2021), neurotransmitters [López et al. 2007; Ihiuchi et al. 2011, Cabezas et al. 2012, 2013(b)], along with a large list of biologically relevant species such as vitamin C and testosterone (Peña et al. 2013; León et al. 2021). Besides the experimental structural information, exploring the conformational panorama, that is, the different three-dimensional arrangements achieved by rotations around single bonds (without breaking or forming new bonds) will help to clarify the nature of the structure-property relationships. For instance, in the case of neurotransmitters, it is expected that if a particular conformation is much more stable than the rest, it will be favored in the “key and lock” recognition process at the receptor site (Ben-Naim 2001).

In the last decade, the study of the shape of neurotransmitters and their precursors has been a challenging endeavour for the chemical community. Although this topic is of major interest for biophysical chemists, the conformational space of these molecular systems, together with the interplay between inter- and intramolecular forces, remain in some cases very incomplete or even unknown. In recent years, the field of rotational spectroscopy has experienced unprecedented growth thanks to the development of new laboratory techniques combined with state-of-the-art theoretical methodologies. Hence, following this line of research, the conformational and structural characterization of **L-DOPA** (see Figure 6), a relevant supramolecular synthon and an amino acid precursor of neurotransmitters dopamine, [Cabezas et al. 2013 (b)] noradrenaline and adrenaline has been performed, whose rotational investigation has remained as a very challenging task for high-resolution spectroscopy until now. In addition, two other organic molecules of great importance for biochemistry and general chemistry have been investigated using the abovementioned approach: **squaric acid**, which is considered the archetypical cyclic oxocarbon (shown in Figure 7), and **phenyl-barbaralone**. The studies of these three molecular systems are presented in Chapter VII.



**Figure 6.** Schematic and classical biosynthesis pathway of dopamine (Meiser et al. 2013).

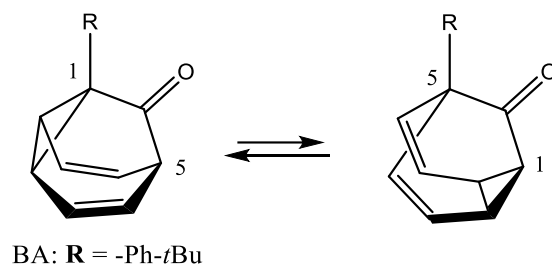
In the last decade, **squaric acid** and its derivatives have been extensively used as supramolecular synthons to design new crystal structures and different metal-organic frameworks (MOFs) (Das et al. 2005; Wang et al. 2007). Also, their direct derivatives, squaramides, have been widely employed in Organic Chemistry as organocatalysts (Chen et al. 2014; Andrés et al. 2017; Li et al. 2018), in Biomedical Chemistry (Liu et al. 2013; Hill et al. 2015) and also in Material Science (Skopinska-Wisnewska et al. 2016). However, its tridimensional structure, as well as the structure of its water-clusters, remains unexplored in the isolated conditions of the gas phase, emerging as a very attractive topic for high-resolution rotational spectroscopy.



**Figure 7.** Cyclic oxocarbon acids obeying to the general formula ( $C_nO_nH_2$ ).

Finally, the characterization and interpretation of constitutional dynamic systems remains as a puzzling and challenging task for physical organic chemistry. In particular, the phenomenon of valence tautomerism consists in an intramolecular reorganization involving just one reactive species, where a particular intramolecular bond migrates giving rise to a usually different molecular structure (shown in Figure 8). Within this

framework, the fluxional family of molecules called **barbaralones** are deeply ingrained in the understanding of valence tautomerism phenomena. These species can undergo [3, 3]-sigmatropic processes, where the number of constitutional isomers (valence tautomers in this case) depend on the presence and chemical nature of any possible substituent, which may induce isomerization restrictions.



**Figure 8.** Sigmatropic equilibrium for 1-substituted barbaralones (i.e. phenyl-barbaralone).

To the best of our knowledge no gas-phase rotational study has been conducted to investigate the [3.3]-sigmatropic equilibria so far. Thus, it is of great interest to employ high-resolution spectroscopic techniques to decipher the valence tautomerism phenomena. As a proof of concept, we chose a **phenyl-barbaralone** to fully understand these fluxional molecular systems and characterize their three-dimensional structure, and to ultimately study the valence tautomerism equilibrium in the absence of intermolecular interactions. In **Chapter VII**, the main results regarding the spectroscopic study for **phenyl-barbaralone** are shown.

### 7. Thesis outline

The subject of this Doctoral Thesis must be therefore understood as two different thematic streams. First, the backbone of the thesis is focused on the spectroscopic characterization of molecules of astrochemical relevance and their possible identification in the ISM, as well as the theoretical study and analysis of some of the plausible formation processes (ion-molecule reactions) in the gas-phase starting from molecules already detected in the ISM. The second one proposes the structural and conformational study of several relevant organic and biomolecules. Table 2 resumes the layout of the molecular

systems studied - either in theoretically or in the laboratory- in the main chapters of the Doctoral Thesis.

**Table 2.** List of the molecular systems studied in the present thesis.

Chapter	Molecular systems	Chemical formula	Search in the ISM
IV: Experimental and theoretical study of interstellar glycine isomers	Protonated glycine isomers	$[\text{H}_6\text{C}_2\text{O}_2\text{N}]^+$	No
	Glycolamide	$\text{CH}_2(\text{OH})\text{C}(\text{O})\text{NH}_2$	Yes
	Acetohydroxamic acid	$\text{CH}_3\text{C}(\text{O})\text{NHOH}$	No
V: Hunting complex cyanides in the interstellar medium	Cyanoacetamide	$\text{CH}_2(\text{CN})\text{C}(\text{O})\text{NH}_2$	Yes
	Cyanoacetic acid	$\text{CH}_2(\text{CN})\text{C}(\text{O})\text{OH}$	Yes
VI: Toward de limits of complexity of interstellar chemistry: rotational spectroscopy and astronomical search for n- and i-butyraldehyde	<i>n</i> -butyraldehyde	$\text{CH}_3\text{CH}_2\text{CH}_2\text{CHO}$	Yes
	<i>i</i> -butyraldehyde	$(\text{CH}_3)_2\text{CHCHO}$	Yes
VII: The shape of biomolecules, fluxional systems and organic synthons	L-DOPA <sup>[a]</sup>	$\text{C}_9\text{H}_{11}\text{NO}_4$	-
	Squaric acid	$c\text{-C}_4\text{H}_2\text{O}_4$	-
	Phenyl-barbaralone	$\text{C}_{15}\text{H}_{12}\text{O}$	-
Appendix I	Amino acetaldehyde	$\text{NH}_2\text{CH}_2\text{CHO}$	Yes <sup>[b]</sup>

**Nota:** [a] Squaric acid, L-DOPA and phenyl-barabaralone are systems of organic and biochemical relevance which are not candidates for interstellar search; [b] To date, the predicted spectroscopic parameters, even at the highest level of theory available, do not allow an unequivocal identification of the molecule in the ISM (a precision of about 0.2% is estimated for the rest frequencies). Nevertheless, several astrophysicists have tried, without success, to search for this molecule in the ISM (Belloche 2021, priv. com.).

In the present Dissertation it will be shown that, in order to adequately understand the different pathways for the chemical evolution of more complex molecular systems, it is mandatory to carry out a coordinated and thorough multidisciplinary effort (observational, theoretical and in the laboratory). This work demonstrates the great importance of the data obtained in the laboratory using rotation spectroscopy to answer fundamental questions for Astronomy and Astrochemistry, such as the level of chemical complexity reached in the ISM, ultimately seeking to shed some light on those species related to the origin of life on Earth.



## 8. References

- Agúndez, M., Marcelino, N., & Cernicharo, J. **2018**, *ApJ*, 861, L22.
- Andrés, J.M., Losada, J., Maestro, A., Rodríguez-Ferrer, P., Pedrosa, R., **2017**, *J. Org. Chem*, 82, 8444,8454.
- Alonso, E. R., Kolesníková, L., Peña, I., Shipman, S. T., Tercero, B., Cernicharo, J., Alonso J. L., **2015**, *J. Mol. Spectrosc.*, 316, 84–89.
- Alonso, E. R., Kolesníková, L., Alonso, J. L., **2017**, *J. Ch.em. Phys.* 147, 124312..
- Alonso, E. R., Kolesníková, L., Białkowska-Jaworska, E., Kisiel, Z., León, I. Guillemin J.-C., & Alonso, J. L., **2018**, *APJ*, 861, 70-77.
- Alonso, E. R., McGuire, B. A., Kolesníková, L., et al., **2019**, *ApJ*, 883, 18.
- Alonso, E.R., Kolesníková, L., Belloche A., Mata. S. et al. **2021**, *A&A* 647, A55 (a)
- Alonso, E. R., León, I., Alonso, J. L., *Intra- and Intermolecular Interactions between Non-covalently Bonded Species*. Elsevier, **2021**, 93-141 (b)
- Alonso, J. L., López, J. C., Springer Berlin Heidelberg, **2015**, pp. 1-67.
- Altwegg, K., Balsiger, H., Bar-Nun, A., Berthelier, J.-J., et al. **2016**, *Sci. Adv.* 2, e1600285.
- Aponte, J. C., Whitaker, D., Powner, M. W., et al. **2019**, *ACS Earth and Space Chemistry*, 3, 463.
- Aprile, E. & Profumo, S., **2009**, *New J. Phys.* 11 105002.
- Barone V, Biczysko M, Puzzarini C., **2015**, *Acc. Chem. Res.*, 48:1413–1422.
- Barrientos, C., Redondo, P., Largo, L., et al. **2012**, *ApJ*, 748, 99
- Barrientos, C., Rayón, V.M., Largo, A., Sordo, J.A., Redondo, P. **2013**, *Journal of Physical Chemistry*, 117, 7742.
- Belloche, A., Garrod, R. T., Müller, H. S. P., et al., **2009**, *A&A*, 499, 215.
- Belloche, A., Garrod, R. T., Müller, H. S. P., et al., **2014**, *Science*, 345, 1584.
- Ben-Naim, A. *Cooperativity and Regulation in Biochemical Processes*, Kluwer Academic/Plenum Publishers, New York, **2001**.
- Bergner, J. B., Guzmán, V. G., Öberg, K. I., Loomis, R. A., & Pegues, J. **2018**, *ApJ*, 857, 69.
- Biczysko, M., Bloino, J., Puzzarini, C. **2018**, *Comput. Mol. Sci.*, 8:e1349.
- Blagojevic, V.; Petrie, S.; Bohme, D. K., **2003**, *Mon. Not. R. Astron. Soc.* 339, L7–L11.
- Burkhardt, A. M., Long Kelvin Lee, K., Bryan Changala, P., et al. **2021**, *ApJ*, 913, L18.
- Burton, A. S., Stern, J. C., Elsila, J. E. et al. **2012**, *Chem. Soc. Rev.*, 41, 5459-5472.
- Cabezas, C.; Varela, M.; Peña, I.; Lopez, J. C.; Alonso, J. L., **2012**, *Phys. Chem. Chem. Phys.*, 14, 13618–13623.
- Cabezas, C., Cernicharo, J., Alonso, J.L., Agúndez, M., Mata, S., Guelin, M., Peña, I., **2013**, *ApJ*, 775, 133 (a).
- Cabezas, C.; Peña, I. López, J. C.; Alonso, J.L., *J. Phys. Chem. Lett.* **2013**, 4, 486-490 (b).
- Cernicharo, J., Kisiel, Z., Tercero, B., Kolesníková, L., et al., **2016**, *A&A*, 587, L4.
- Cernicharo, J., Marcelino, N., Angúdez, M., et al., **2020**, *APJL*, 900, L9.
- Cernicharo, J., Cabezas, C., Endo, Y. et al., **2021**, *A&A*, 650, L14. (a)

- Cernicharo, J., Agúndez, M., Cabezas, C., et al. **2021**, *A&A*, 649, L15. (b)
- Chen, S., Pan, J., Wang, Y., Zhou, Z., **2014**, *Eur. J. Org. Chem.*, 7940.
- Chyba, C. F., Thomas, P. J., Brookshaw, L., & Sagan, C. **1990**, *Science*, 249, 366.
- Chyba, C., & Sagan, C. **1992**, *Nature*, 355, 125.
- Chyba, C. F., & Hand, K. P. **2005**, *ARA&A*, 43, 31.
- Cronin, J. R., **1989**, *Advances in Space Research*, 9, 59.
- Cronin, J. R., & Moore, C. B., **1976**, *Geochim. Cosmochim. Acta*, 40, 853.
- Cronin, J. R., & Pizzarello, S., **1983**, *Advances in Space Research*, 3, 5.
- Cunningham, M. R.; Jones, P. A.; Godfrey, P. D.; Cragg, D. M.; et al., **2007**, *Mon. Not. R. Astron. Soc.* 376, 1201–1210.
- Dae, N., Ghosh, A., Arif, M. & Stan, P.J., **2005**, *Inorg. Chem.*, 44, 7130-7137.
- Daly, A.M., Kolesniková, L., Mata, S. & Alonso J.L., **2014**, *J. Mol. Spectrosc.*, 306, 11–18.
- Daly, A.M., Bermúdez, C., Kolesniková, L., Alonso, J.L. **2015**, *ApJS*, 218:30 (7pp) .
- Dartois, E. **2005**, *Space Sci. Rev.*, 119, 293.
- Demyk, K.; Wlodarczak, G.; Dartois, E. in *Semaine de l’Astrophysique Francaise*, ed. Combes, F.; Barret, D.; Contini, T.; Meynadier, F.; Pagani, L. SF2A-2004 (Les Ulis: EDP-Sciences), **2004**, 493.
- Draine, B. T., **2003**, *Annu. Rev. Astron. Astrophys.* 41 241–89.
- Draine, B. T. *Physics of the Interstellar and Intergalactic Medium*, Princeton University Press, **2011**. ISBN: 978-0-691-12214-4.
- Elsila, J. E.; Dworkin, J. P.; Bernstein, M. P.; Martin, M. P.; Sandford, S. A. **2007**, *APJ.*, 660, 911–918.
- Ferguson, E. E. **1975**, *Annu. Rev. Phys. Chem.*, 26, 17–38.
- Fortman, S. M., Medvedev, I. R., Neese, C. F., & De Lucia, F. C. **2010**, *ApJ*, 725, 1682.
- Garrod, R. T. & Herbst, E. **2006**, *A&A*, 457, 927.
- Garrod, R. T. **2013**, *ApJ*, 765, 60.
- Geppert, W. D.; Larsson, M., **2013**, *Chem. Rev.* 113, 8872– 565 8905.
- Godfrey, P. D., Brown, R. D., Robinson, B. J. & Sinclair, M.W. **1973**, *APJL*, 13, 119–121.
- Goldsmith, P., Bergin, T., De Lucia, F. C., et al. **2006**, Report from the Workshop on Laboratory Spectroscopy in Support of Herschel, Sofia, and Alma, Tech. Rep. (Pasadena, CA: California Institute of Technology)
- Gottlieb, C. A., **1973**, *Molecules in the Galactic Environment*, 181.
- Groner, P.; Winnewisser, M.; Medvedev, I. R., et al.. **2007**, *Astrophys. J. Suppl. Series.*, 169, 28-36.
- Herman, Z.; Futrell, J. H. **2015**, *Int. J. Mass Spectrom.* 377, 84–92.
- Jortner, J., **2006**, *Phil. Trans. R. Soc. B*, 361, 1877.
- Jørgensen, J. K., Belloche, A., & Garrod, R. T., **2020**, *ARA&A*, 58, 727.
- Koga, T. & Naraoka, H., **2017**, *Scientific Reports*, 7, 636.
- Herbst, E., & van Dishoeck, E. F. **2009**, *ARA&A*, 47, 427.
- Hill, N.D., Bunata, K., Hebert, A., **2013**, *Clinics in Dermatology*, 33, 300-304.
- Hollis, J. M., Lovas, F. J., & Jewell, P. R. **2000**, *ApJ*, 540, L107.

- Hollis, J. M., Jewell, P. R., Lovas, F. J., et al. **2004**, *ApJ*, 610, L21
- Holtom, P. D.; Bennett, C. J.; Osamura, Y.; Mason, N. J.; Kaiser, R. I., **2005**, *APJ*, 626, 940 – 952.
- Ishiuuchi, S.; Asakawa, T.; Mitsuda, H.; Miyazaki, M.; Chakraborty, S.; Fuvii, M., **2011**, *J. Phys Chem. A*, 115, 10363–10369.
- Jeanvoine, Y.; Largo, A.; Hase, W.L.; Spezia R.; **2018**, *J. Phys. Chem. A*. 122: 869-877.
- Jones, P. A.; Cunningham, M. R.; Godfrey, P. D.; Cragg, D. M. **2007**, *Mon. Not. R. Astron. Soc.*, 374, 579–589.
- Kallman, T. R., and Palmeri, P., **2007**, *Rev. Mod. Phys.* 79 79–133
- Kappeler, F., Gallino, R., Bisterzo, S. & Aoki, W. **2011**, *Rev. Mod. Phys.* 83 157–93.
- Koch, D. M., Toubin, C., Peslherbe, G. H., & Hynes, J. T. J. **2008**, *Phys. Chem. C*, 112, 2972–2980.
- Kolesníková, L., Tercero, B. Cernicharo, J., Alonso, J. L., et al., **2014**, *ApJ*, 784 L7.
- Kolesníková, L., Tercero, B., Alonso, E. R., et al. **2017**, *A&A*, 609, A24.
- Kolesníková, L. Alonso; E.R., Mata, S., Alonso, J.L. **2017**, *A&A Supplement Series*, 229:26 (8pp).
- Kolesníková, L., León, I., Alonso, E. R., Mata, S., Alonso, J. L., **2019**, *J. Phys. Chem. Lett.*, 10, 1325-1330.
- Kuan Y.-J. *Astrophys. J.*, **2003**, 848 – 867.
- Larsson, M.; Geppert, W. D.; Nyman, G. **2012**, *Rep. Prog. Phys.* 75, 066901.
- Lee, K. L. L., Changala, P. B., Loomis, R. A., et al., **2021**, *ApJ*, 910, L2.
- León, I., Alonso, E. R., Mata, S., Alonso, J.L., **2021**, *J. Phys. Chem. Lett.*, 29; 12 (29): 6983-6987
- Levy D.H., **1980**, *Annu Rev Phys Chem.*, 31:197, 2.
- Li, S.C., Zhang, T., Deng, X. P., et al., **2018**, *Inor. Chem. Comm*, 92, 59-73.
- Liu, Z. Wang U, Han W et al., **2013**, *European Journal of Medical Chemistry*, 65, 187-194.
- Lopez, J. C.; Cortijo, V.; Blanco, S.; Alonso, J. L. **2007**, *Phys. Chem. Chem. Phys.*, 9, 4521– 4527.
- Scoles G, *Atomic and molecular beam methods*, vol 1. Oxford University Press, New York/Oxford, **1988**.
- Loomis, R. A., Burkhardt, A. M., Shingledecker C. N. et al., **2021**, *N. Nat. Astron.* 5, 188–196.
- Maeda, S.; Ohno, K. **2004**, *Chem. Phys. Lett.*, 398, 240–244.
- Marcelino, N., Angúdez, M., Tercero, B., et al. **2020**, *A&A*, 643, L6.
- Maris, A. **2004**, *Phys. Chem. Chem. Phys.*, 6, 2611-2616.
- Maury, A. J., Belloche, A., André, P., et al., **2014**, *A&A*, 563, L2.
- McCarthy, M. C., Kelvin Lee, K. L., Loomis, R. A. et al., **2021**, *Nature Astronomy*, 5, 176-180.
- McGuire, B. A. **2018**, *ApJS*, 2. (a)
- McGuire, B. A., Burkhardt, A. M., Kalenskii, S., et al., **2018**, *Science*, 359, 202. (b)
- McGuire, B., Burkhardt, A., Loomis, R., et al., **2020**, *APJL*, 900, L110.
- McGuire, B. A., Loomis, R. A., Burkhardt, A. M., et al., **2021**, *Science*, 371, 1265.
- McKellar, A. **1940**, *PASP*, 52, 187.
- Mehring, D. M., Miao, Y., Kuan, Y.-J., & Snyder, L. E., **1995**, *American Astronomical Society Meeting Abstracts*, 184, 26, 906.

- Meiser, J., Weindl, D. & Hiller, K., **2013**, *Cell Commun Signal*, 11, 34.
- Miller, S. L., **1953**, *Science*, 117, 528–529.
- I. Peña, A. M. Daly, C. Cabezas, S. Mata, C. Bermúdez, A. Niño, J.C. López, J-U. Grabow, and J. L. Alonso, **2013**, *J. Phys. Chem. Lett.*, 4(1), pp 65–69.
- Petrie, S.; Bohme, D. K. Ions in space. **2007**, *Mass Spectrom. Rev.*, 26, 258–280.
- Pizzarello, S. **2006**, *Acc. Chem. Res.*, 39, 321.
- Pizzarello, S., Shock, E. **2010**, *Cold Spinrg Harb. Perspect. Biol.*, 2, 1-20.
- Puzzarini C, Biczysko M, Bloino J, Barone V. **2014**, *APJ*, 785:107.
- Redondo, P.; Martínez, H.; Cimas, A.; Barrientos, C.; Largo, A. **2013**, *Phys. Chem. Chem. Phys.* 15, 13005.
- Redondo, P., Barrientos, C., & Largo, A. **2014**, *ApJ*, 793, 32
- Redondo, P., Largo, A., & Barrientos, C. **2015**, *A&A*, 579, A125
- Redondo, P., Martínez, H., Largo, A., et al., **2017**, *A&A*, 603, A13
- Redondo, P.; Largo, A.; Barrientos, C., **2017**, *ApJ*, 836, 240.
- Remijan, A. J., Hollis, J. M., Lovas, F. J., et al. **2008**, *ApJ*, 675, L85.
- Remington, B. A., Drake, R. P., & Ryutov, D. D., **2006**, *Rev. Mod. Phys.* 78 755–807.
- Rimola, A.; Sodupe, M.; Ugliengo, P., **2012**, *APJ*, 754, 24.
- Rivilla, V. M., Martín-Pintado, J., Jiménez-Serra, I., et al., **2020**, *ApJ*, 899, L28.
- Rivilla, V. M., Jiménez-Serra, I., Martín-Pintado, J., et al., **2021**, *Proceedings of the National Academy of Science*, 118, 2101314118.
- Ruiz-Mirazo, K., Briones, C., & de la Escosura, **2014**, *A. Chem. Rev.*, 114, 285–366.
- Sakai, N. & Yamamoto, S. **2011**, 280, 43.
- Salter, C. J., Ghosh, T., Catinella, B., Lebron, M., Lerner, M. S., Minchin, R. and Momjian, E., **2008**, *Astron. J.*, 136, 389–399.
- Sanz-Novo, M., Belloche, A., Alonso, J. L., et al., **2020**, *A&A*, 639, A135.
- Savin, D. W., Brickhouse, N. N., Cowan, J.J., et al., **2012**, *Rep. Prog. Phys.* 75, 036902.
- Skopinska-Wisniewska, J. Kuderko, J., Bajek, A. et al., **2016**, *Materials Science and Engineering C*, 60, 100-109.
- Snow, J. L.; Orlova, G.; Blagojevic, V.; Bohme, D. K., **2007**, *J. Am. Chem. Soc.* 129, 9910–9917.
- Snyder, L. E., Buhl, D., Zuckerman, B., et al. **1969**, *Phys. Rev. Lett.*, 22, 679.
- Snyder, L. E.; Lovas, F. J.; Hollis, J. M.; Friedel, D. N.; Jewell, P. R.; Remijan, A. **2005**, *APJ*, 619, 914–930.
- Spitzer, L. *Physical Processes in the Interstellar Medium*, New York Wiley-Interscience, **1978**, 333.
- Strecker, A., **1850**, *Annalen der Chemie und Pharmazie*, 75, 27–45.
- Tercero, B., Cernicharo, J. López, A., Brouillet, N., et al., **2015**, *A&A* 582, L1.
- Thaddeus, P. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **2006** 361, 1681.
- Tielens, A. G. G. M., **2013**, *Rev. Mod. Phys.* 85. 3.
- Vazart F, Calderini D, Puzzarini C, Skouteris D, Barone V.. **2016**, *J. Chem. Theory Comput*, 12:5385–5397.

Walsh, C., Loomis, R. A., Öberg, K. I., et al. **2016**, *ApJ*, 823, L10.

Wiescher M, Gorres J, Uberseder E, Imbriani G and Pignatari M **2010**, *Annu. Rev. Nucl. Part. Sci.*, 60 381–404.

Willis, E. R. & Garrod, R. T. **2017**, *ApJ*, 840, 61.

Wilson, T. Rohlfs, K., Huettemeister, S. *Tools of Radio Astronomy*, Springer-Verlag Berlin Heidelberg, 6<sup>a</sup> edición, **2013**.

Woon, D. E. **2002**, *ApJ*, 571, L177.

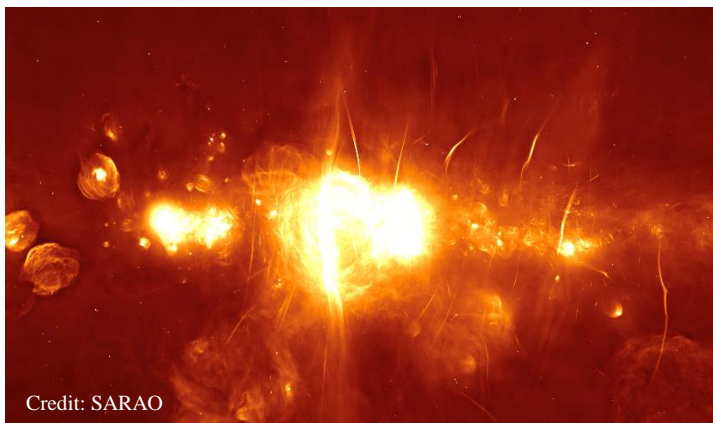
Yamada, M., Kulsrud, R., and Ji H. **2010**, *Rev. Mod. Phys.* 82 603–64

Zeng, S., Quénard, D., Jiménez-Serra, I., et al., **2019**, *MNRAS*, 484, L43-L48.



## CHAPTER II.b INTRODUCCIÓN

---



Dentro de la Química-Física actual, uno de los campos de mayor interés es la Astroquímica o Química Interestelar. Por un lado, la Astroquímica de Laboratorio juega un papel fundamental en la comprensión de nuestro universo molecular, complementando las observaciones astronómicas y acompañando el modelado teórico astroquímico (Tielens, 2013). Dicho campo engloba varias áreas de la física relevantes para la Astronomía y la Astrofísica (Savin y col., 2012). Entre ellas, destacan los experimentos de laboratorio basados en la espectroscopía de rotación, cuya simbiosis con las búsquedas astronómicas ha permitido la detección de gran cantidad de moléculas interestelares (Herbst & van Dishoeck 2009, Jorgensen 2020). No obstante, las condiciones generales del medio interestelar (ISM) hacen que la química que allí se desarrolla tenga unas peculiaridades diferentes a la terrestre y compuestos allí presentes son difícilmente aislables en el laboratorio. Por tanto, la Química Computacional puede desempeñar también un papel fundamental en este campo. En este capítulo se desarrollarán una serie de temas fundamentales dentro de la Astrofísica de Laboratorio y la Astroquímica Computacional tratando de lograr un primer acercamiento a este campo. También se realizará una revisión de los antecedentes bibliográficos más interesantes y recientes del mismo, centrándonos las últimas detecciones de moléculas interestelares y en el estudio y búsqueda de la glicina, el aminoácido más sencillo, en el ISM.

### 1. Antecedentes

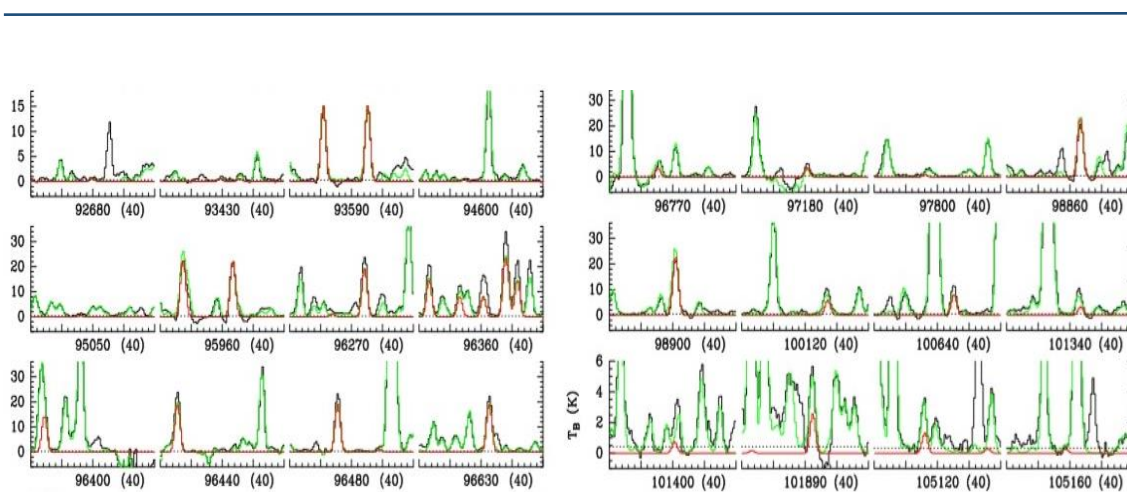
La presente Tesis Doctoral se encuentra en el marco de una beca predoctoral FPU (FPU17/02987) otorgada por el Ministerio de Ciencia, Innovación y Universidades, asociada a los proyectos del Plan Nacional CTQ2016-76393-P y PID2019-111396GB-I00, que permitió financiar los estudios teórico-experimentales de moléculas de relevancia astrofísica, así como la investigación acerca del comportamiento conformacional y estructura tridimensional de biomoléculas realizados durante el periodo 2018-2021. El trabajo recogido en la presente Memoria se ha realizado en los Laboratorios de Espectroscopía y Bioespectroscopía del Grupo de Espectroscopía Molecular (GEM) de la Universidad de Valladolid, Unidad Asociada del CSIC, ubicados en el Edificio Quifima del Parque Científico UVa, así como en las instalaciones del Grupo de Química Computacional (CCG) de la Universidad de Valladolid.

En la última década se ha desarrollado una instrumentación puntera que combina técnicas de microondas en el dominio del tiempo con expansiones supersónicas y sistemas de ablación láser para la vaporización de muestras sólidas. Como se explicará en el Capítulo III, dichas técnicas se conocen por los acrónimos LA-MB-FTMW (Laser Ablation Molecular Beam Fourier Transform Microwave) y LA-CP-FTMW (Laser Ablation Chirped Pulse Fourier Transform Microwave) y, junto con las técnicas convencionales de calentamiento, han permitido la realización de los primeros estudios espectroscópicos, en concreto caracterizaciones en rotación en la región de microondas, de moléculas candidatas a ser detectadas en el ISM, así como de diversas biomoléculas y moléculas de índole orgánico bajo las condiciones de aislamiento en fase gas que ofrece la expansión supersónica.

La identificación de nuevas moléculas en el ISM requiere indispensablemente de una caracterización espectroscópica previa. En este contexto, las medidas de laboratorio resultantes de los estudios espectroscópicos a baja frecuencia permitirán posibles búsquedas interestelares a longitudes de onda centimétricas empleando los radio telescopios de Yebes 40-m (Yebes, Guadalajara, Spain), Effelsberg 100-m (Bad Münstereifel, Germany) y GBT 100-m (Green Bank Telescope, West Virginia, US). Posteriormente, los datos se extenderán en la medida de lo posible a las regiones milimétrica y submilimétrica (mm-submm) del espectro electromagnético, permitiendo su comparación directa con los registros o “*Line Surveys*” de diferentes radio



observatorios, destacando el interferómetro ALMA (Atacama Large Millimeter Array, Chile) y el radio telescopio IRAM 30-m (Institute for Radio Astronomy in the Millimeter Range, Pico Veleta, España). En este contexto, destaca la participación de nuestra Unidad Asociada CSIC en el proyecto europeo ERC-SyG-2013 Grant Agreement n. 610256 Nanocosmos, así como la estrecha colaboración con el Instituto Max Planck de Radioastronomía, una de las principales instituciones a nivel mundial en el ámbito de la astrofísica, remarcando su papel en la detección e identificación de numerosas moléculas en el ISM. En la Figura 1 se muestra a modo de ejemplo una porción del registro de ALMA denominado EMoCA (*Exploring Molecular Complexity with ALMA*) de Sgr B2(N2) incluyendo las transiciones detectadas del **acetaldehído** ( $\text{CH}_3\text{CHO}$ ).



**Figura 1.** a) Colección de transiciones del estado de vibración fundamental,  $v=0$  del acetaldehído ( $\text{CH}_3\text{CHO}$ ) cubiertas por el registro de ALMA EMoCA de Sgr B2(N2). El espectro sintético bajo condiciones de equilibrio termodinámico local (LTE) para dicha especie (en rojo), empleado para derivar el límite superior para la columna de densidad del acetaldehído, se muestra superpuesto sobre el espectro de Sgr B2(N2) (en negro). El espectro sintético indicado en color verde muestra las contribuciones de todas las moléculas identificadas en dicho registro hasta la fecha. Para cada panel la frecuencia central y la anchura están indicadas en MHz en el eje x, mientras el eje y se corresponde con la temperatura de brillo (*brightness temperature*, en K) [Figura adaptada de Sanz-Novo y col. 2019].

Además, en paralelo a los estudios experimentales, para algunos de los sistemas de estudio se ha realizado un análisis de los posibles procesos de formación en fase gas a partir de moléculas ya detectadas en el ISM, así como diversos cálculos computacionales a alto nivel, cuya interpretación astroquímica es fundamental para elucidar el modelado del universo molecular (Tielens 2013), las vías de formación en el espacio de las distintas familias de compuestos químicos, así como en las investigaciones sobre el origen de la vida.

## 2. Medio interestelar (ISM): composición e importancia de los granos

Brevemente, la materia que se encuentra entre las estrellas, o medio interestelar (ISM), consiste en una mezcla muy diluida de iones, átomos, moléculas, granos de polvo (del orden de un 99 % gas y un 1 % de partículas polvo, en porcentaje en masa, Spitzer, 1978), bañados en un campo de radiación electromagnético.

Por un lado, el gas interestelar está compuesto mayoritariamente por hidrógeno, el elemento más abundante en el Universo (~70% de la masa total del ISM), que se puede encontrar en diferentes estados. Aproximadamente, el 60% del hidrógeno interestelar se encuentra en forma atómica (H), un 20% genera hidrógeno diatómico o molecular (H<sub>2</sub>), y finalmente el resto se halla en su estado ionizado (H<sup>+</sup>). En función de las propiedades del gas interestelar y dependiendo del estado del hidrógeno, se puede dividir el ISM en distintas fases (Draine, 2011): el medio ionizado caliente (*hot ionized médium*, HIM), el medio ionizado cálido (*warm ionized médium*, WIM), el medio neutro cálido (*warm neutral médium*, WNM), el medio neutro frío (*cold neutral médium*, CNM) y, por último, las nubes moleculares (*molecular clouds*, MC). Además, el ISM también contiene un contenido minoritario de Helio y otros elementos más pesados. Conjuntamente con el gaseoso, otro de los estados de la materia que se encuentra en el espacio interestelar es el polvo cósmico (Thaddeus, 2006). Este juega un papel fundamental en la Química Interestelar actuando en muchos casos como superficie catalizadora para diversos procesos químicos, y la regulación de los procesos de adsorción-desorción controla la química presente en las distintas regiones del espacio. La composición de los granos se basa fundamentalmente en silicio, carbono y oxígeno. En función de las proporciones se forman silicatos o granos carbonáceos. Además de la estructura principal de los granos, se encuentran otras especies moleculares en forma de hielo como CO, CO<sub>2</sub>, CH<sub>3</sub>OH, NH<sub>2</sub>, CH<sub>2</sub>. Durante la evolución del grano, este material sufrirá numerosas modificaciones químicas causadas por irradiaciones iónicas y UV, así como efectos térmicos (Dartois, 2005). Las partículas de polvo interestelar han podido ser un posible origen de átomos prebióticos como el C, H y N (Chyba & Sagan 1992; Chyba & Hand 2005), apuntando hacia la interconexión entre la Química Interestelar y el origen terrestre de la vida (Jorten, 2006).

### 3. *Astroquímica o Astrofísica de laboratorio*

Nuestro entendimiento acerca del Universo molecular y la Astroquímica que subyace detrás depende enormemente de los avances de la Astroquímica o Astrofísica de laboratorio. Dicho campo, como indica su propio nombre, tiene una motivación puramente astrofísica, y nos permite complementar las observaciones astronómicas y acompañar el modelado teórico astroquímico (Tielens, 2013). En línea con la revisión de Savin y col. (2012), se puede hacer un desglose en cinco grandes temáticas o categorías diferentes: a) Los sistemas planetarios y la formación de estrellas, b) La evolución estelar, c) La estructura de la Vía Láctea y otras galaxias cercanas d) La evolución temporal de las galaxias y e) Cosmología y física fundamental. El estudio de todas ellas conlleva el desarrollo de varias áreas de la física, destacando la químico-física atómica y molecular (Kallman & Palmeri 2007; Herbst & van Dishoeck 2009), la física de la materia condensada (Draine 2003), la física de plasmas (Remington y col. 2006; Yamada y col. 2010), la física nuclear (Käppeler y col. 2011; Wiescher y col. 2010) y la física de partículas (Aprile & Profumo 2009).

Uno de los objetivos fundamentales consiste en la detección de moléculas prebióticas y de interés biológico en diversos ambientes astroquímicos, aspirando en última instancia a obtener un censo completo para los registros espectrales (idealmente de sensibilidad uniforme) de un objeto astronómico concreto (Jørgensen 2020). En la presente tesis, el trabajo ha centrado en la químico-física y la física molecular, siendo principalmente los datos de laboratorio (experimentales y teóricos) basados en la espectroscopía de rotación los que han permitido la detección una gran cantidad de moléculas interestelares (Herbst & van Dishoeck 2009, Jørgensen 2020).

En este contexto, en nuestros laboratorios de la Universidad de Valladolid se han estudiado los espectros de rotación del **etil mercaptano**  $\text{CH}_3\text{CH}_2\text{SH}$  (Kolesníková y col. 2014) y **isocianuro de hidromagnesio**  $\text{HMgNC}$  [Cabezas y col. 2013(a)], **etil metil éter**  $\text{CH}_3\text{CH}_2\text{OCH}_3$  (Tercero y col. 2015) y **metilisocianato**  $\text{CH}_3\text{NCO}$  (Cernicharo y col. 2016), que han permitido su identificación inequívoca en el ISM. Igualmente se han analizado los espectros de la **acroleína**  $\text{CH}_2\text{CHCHO}$  (Daly y col. 2015), **metil vinil éter**  $\text{CH}_2\text{CH}_2\text{OCH}_3$  (Daly y col. 2014) **acetato de vinilo**  $\text{CH}_2\text{CHCO}_3\text{H}$  (Alonso y col. 2015), la **metoxiamina** (Kolesníková y col. 2017), el **lactaldehído**  $\text{CH}_3\text{CH}(\text{OH})\text{CHO}$  (Alonso y col. 2019) y la **propiolamida**  $\text{HCCC}(\text{O})\text{NH}_2$  (Alonso y col. 2021) entre otras muchas

especies, efectuando en algunos casos detecciones tentativas. Como se explicará en detalle en el **Capítulo III** correspondiente a la Metodología, a parte de las técnicas de calentamiento convencionales, recientemente se han acoplado fuentes de ablación láser con técnicas espectroscópicas de microondas con transformada de Fourier de alta resolución. Muchos compuestos sólidos con elevados puntos de fusión suelen presentar una presión de vapor muy reducida, escapando de los estudios convencionales en fase gas en la región tanto centimétrica como milimétrica. Empleando técnicas de ablación láser, en los últimos años se han investigado diversas especies de relevancia astrofísica como prerequisite para su posterior identificación astronómica, destacando la **hidantoína** (Alonso y col. 2017), el **ácido hidantoico** (Kolesniková y col. 2019) y la **glicinamida**  $\text{CH}_2(\text{NH}_2)\text{C}(\text{O})\text{NH}_2$  (Alonso y col. 2018), posible precursor interestelar de la glicina.

Cabe comentar para la realización de estos trabajos se requiere una colaboración estrecha entre una comunidad muy diversa de astrofísicos y astroquímicos de laboratorio junto con importantes grupos de radio astrónomos y químicos teóricos. En este sentido, la Astroquímica se encuentra en auge y paulatinamente está transicionando a una nueva era de medidas de altísima precisión (tanto a nivel experimental como referido a observaciones astronómicas) y de modelados químicos de alta fidelidad (Garrod 2013).

#### *4. Astroquímica computacional*

Otro de los objetivos fundamentales dentro de la Astroquímica es la elucidación de los mecanismos de formación y transformación de las especies químicas presentes en el ISM (Vazart y col. 2016), que hoy en día siguen siendo objeto de debate, y cuyo estudio experimental no siempre es accesible. En consecuencia, surge un nuevo campo de investigación denominado Astroquímica Computacional, que aplicará los métodos de la química cuántica al estudio de sistemas de relevancia astroquímica (Biczysko y col. 2018). Esta disciplina se ha convertido en una herramienta muy útil para afrontar de manera global y racional los diferentes problemas químicos (Barone y col. 2015). Además, el campo de la Química Interestelar constituye un medio ideal para la aplicación de los métodos teóricos, y resulta también muy útil cuando la experimentación es mucho más difícil de llevar a cabo.

La diversidad y complejidad química de las diferentes fuentes astronómicas se puede explicar como resultado de interacciones de gas, granos, y entre gas y granos. Consideradas en conjunto, permitirán la construcción de intrincadas redes químicas que contienen miles de procesos y reacciones, generando complejos modelos astroquímicos, como por ejemplo el modelo MAGICKAL (*Model for Astrophysical Gas and Ice Chemical Kinetics and Layering*, Garrod 2013), al cual se hará alusión en varios capítulos de la presente tesis.

Por un lado, el estudio de los procesos superficiales ha atraído recientemente la atención de multitud de investigadores teórico-experimentales (Willis & Garrod 2017), tratando de comprender la composición de los mantos de hielo, generados por la acreción y acumulación de especies en fase gas en los granos de polvo, y suponiendo en ciertos casos una alternativa a las reacciones en fase gas. Algunas de las moléculas interestelares se forman probablemente en mantos de hielo, como por ejemplo el **formato de metilo** ( $\text{CH}_3\text{OCHO}$ ), mientras que el **dimetil éter** ( $\text{CH}_3\text{OCH}_3$ ) y el **ácido fórmico** ( $\text{HCOOH}$ ) probablemente tengan su origen tanto en una química en fase gas como en una química de superficies en los granos (Garrod & Herbst 2006). Sin embargo, las reacciones ión-molécula en fase gas sigue teniendo un papel fundamental para la síntesis de muchas moléculas orgánicas interestelares (Petrie y col. 2007, Larsson y col. 2012, Geppert & Larsson 2013; Barrientos y col. 2012; Redondo y col. 2014, 2015, 2017). El grupo de investigación de Química Teórica y Computacional (Computational Chemistry Group, CCG) de la Universidad de Valladolid dirigido por el profesor Antonio Largo lleva estudiando la formación de compuestos de interés astrofísico durante más de una década, destacando los estudios enfocados a la formación de **glicina** ( $\text{NH}_2\text{CH}_2\text{COOH}$ ) en el marco de la Química Computacional (Barrientos y col. 2013, Redondo y col. 2014, 2015, 2017).

Gracias a los últimos avances teóricos, es posible realizar hoy en día estudios mecanísticos, estructurales, termodinámicos y también predecir propiedades espectroscópicas de moléculas de tamaño mediano con una altísima precisión (Barone y col. 2015), empleando métodos de agregados acoplados (CC) y modelos híbridos (ver **Capítulo III**, Metodología). Finalmente, cabe destacar que los datos teóricos han permitido en algunos casos guiar observaciones astronómicas [Cernicharo y col. 2021 (a)], así como proponer nuevos sistemas de estudio de especies aún no detectadas (Puzzarini y col. 2014), esenciales para explicar el origen de moléculas interestelares de

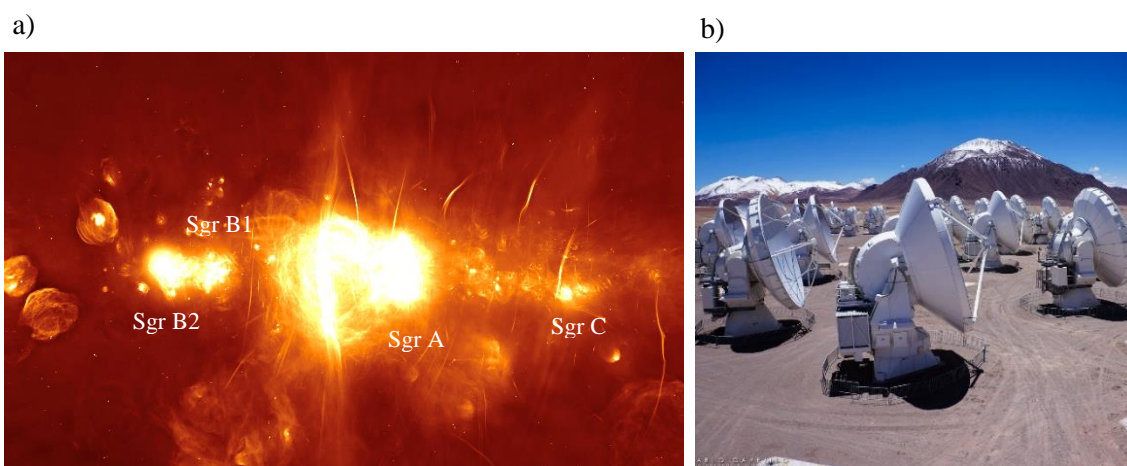
gran tamaño, en algunos casos ingredientes fundamentales de la vida tal y como la conocemos en la Tierra.

### 5. *Moléculas descubiertas en el ISM: moléculas orgánicas complejas (COMs)*

En los últimos años, gracias al gran esfuerzo de la comunidad científica, el número de moléculas detectadas en el medio interestelar y circunestelar, alrededor de 250, ha crecido casi exponencialmente [ver McGuire 2018 (a) para un censo global]. Dentro de las familias de moléculas de mayor relevancia astroquímica encontramos las denominadas moléculas orgánicas complejas (*complex organic molecules*, COMs), definidas como especies de seis o más átomos incluyendo al menos un átomo de carbono (Jørgensen 2020). La mayoría de estos compuestos fueron detectados por primera vez en la región milimétrica y submilimétrica en núcleos moleculares calientes (*Hot molecular cores*, HMCs,  $T \approx 100\text{-}300$  K) tales como Sagitario B2(N) [ver Figura 2(a)] y Orion-K (Herbst & van Dishoeck 2009). Posteriormente, los estudios astronómicos se han extendido a diversos entornos tales como nubes de gas frío ( $T \approx 10$  K) de núcleos preestelares (Vastel y col. 2014) coronas calientes, es decir, los análogos de menor masa de los núcleos calientes (Sakai & Yamamoto 2001, Maury y col. 2014), discos protoplanetarios (Walsh y col. 2016, Bergner y col. 2018), y cometas (Altwegg y col. 2016), entre otros.

Actualmente, la detección de nuevos sistemas complejos en el ISM se ha visto enormemente facilitada por el desarrollo de los radiotelescopios modernos, junto con el uso de métodos interferométricos (Wilson y col. 2013). Los radiotelescopios permiten detectar y analizar la radiación emitida en la zona de microondas y ondas de radio, donde se producen las transiciones entre los distintos estados rotacionales de las moléculas. Encontramos numerosas instalaciones radio astronómicas de interés, destacando el Proyecto ALMA (Atacama Large Millimetre Array) en el desierto de Atacama (Chile) como uno de los más prometedores [ver Figura 2(b)]. La identificación de una especie molecular en el espacio implica una comparación directa de las líneas de frecuencias registradas de emisión con medidas espectroscópicas obtenidas a partir de información teórica (cálculos computacionales) y experimental (registros en un laboratorio). Normalmente, con una decena de identificaciones positivas en rotación (huella dactilar) de líneas no contaminadas se considera una molécula identificada en el ISM, aunque

recientemente se están empleando técnicas más novedosas, destacando el denominado “*Line Stacking*” (Loomis y col. 2021, Lee y col. 2021). Dichas técnicas permiten un incremento de la relación señal/ruido al obtener un espectro promedio que encapsula la formación total de todas las líneas observadas en vez de examinar cada una de las líneas individualmente (McGuire y col. 2021), siempre y cuando el registro no se encuentre cerca del límite de confusión (técnica muy útil para ciertos registros de nubes moleculares oscuras y frías como TMC-1). Además, los descubrimientos astronómicos de nuevos sistemas moleculares suelen ser uno de los principales motivos a la hora de seleccionar nuevos sistemas de estudio.



**Figura 2.** a) Centro de la Galaxia y Sagitario B2 (a aproximadamente 390 años luz de distancia de Sgr A), visto por el interferómetro MeerKat [Crédito de la imagen: MeerKat, SARAO]; b) Proyecto ALMA, desierto de Atacama (Chile) [Crédito de la imagen: NRAO].

El desarrollo tecnológico que han experimentado las instalaciones radio astronómicas en las últimas décadas, destacando los nuevos receptores de banda ancha como los construidos dentro del proyecto Nanocosmos e instalados en el radio telescopio de Yebes-40m (+18 GHz de ancho de banda), ha supuesto la obtención de espectros de una resolución y sensibilidad tan alta - entre 0.25 y 1 mK para la banda *Q* del registro QUIJOTE [*Q-band Ultrasensitive Inspection Journey to the Obscure TMC-1 Environment*, Cernicharo 2021(a)] - donde una gran cantidad de líneas espectrales no han podido ser aún asignadas (“*Unidentified lines*” o *U-lines*). En muchos casos esta cuestión se debe a la existencia de *Line Surveys* que prácticamente llegan al denominado “límite de confusión”. Dichas transiciones de rotación podrán pertenecer a nuevas moléculas aún sin identificar en el ISM, pero también a estados vibracionales excitados de unas pocas

moléculas denominadas “astrophysical weeds” (Goldsmith y col. 2006; Fortman y col. 2010, Kolesniková y col. 2017), o incluso a isotopólogos o conformeros de moléculas ya detectadas para los cuales se desconoce el espectro de rotación. Dicha cuestión es crucial cuando las observaciones astronómicas implican regiones calientes del ISM, como son las fuentes relacionadas con la formación de estrellas conocidas como núcleos calientes o “*hot cores*”. Se espera que los nuevos observatorios europeos y americanos aumenten notablemente el repertorio conocido del universo molecular.

A continuación, se presenta en la Tabla 1 un listado de las moléculas detectadas, hasta la fecha, en el medio interestelar y circunestelar, obtenido de la base de datos del Centro de Colonia para Espectroscopía Molecular (Cologne Database for Molecular Spectroscopy, CDMS<sup>1</sup>).

---

<sup>1</sup>La lista completa de moléculas descubiertas en el medio interestelar y circunestelar, así como las moléculas extragalácticas, está disponible vía <http://www.astro.uni-koeln.de/cdms/molecules>



**Tabla 1.** Moléculas descubiertas en el ISM y en envoltorios circunestelares.

2 átomos		3 átomos		4 átomos		5 átomos		10 átomos	
H <sub>2</sub>	SH <sup>+</sup>	H <sub>3</sub> <sup>(*)</sup>	OCS	CH <sub>3</sub>	C <sub>3</sub> S	CH <sub>4</sub> <sup>*</sup>	CH <sub>3</sub> COCH <sub>3</sub>		
AlF	SH	CH <sub>2</sub>	SO <sub>2</sub>	NH <sub>3</sub>	HMgNC	CH <sub>2</sub> NH	(CH <sub>2</sub> OH) <sub>2</sub>		
AlCl	HCl <sup>+</sup>	NH <sub>2</sub>	c-SiC <sub>2</sub>	H <sub>3</sub> O <sup>+</sup>	HCCO	SiH <sub>4</sub>	CH <sub>3</sub> CH <sub>2</sub> CHO		
C <sub>2</sub> <sup>**</sup>	TiO	H <sub>2</sub> O	SiCN	C <sub>2</sub> N <sub>2</sub>	CNCN	H <sub>2</sub> COH <sup>+</sup>	CH <sub>3</sub> C <sub>5</sub> N		
CH	ArH <sup>+</sup>	C <sub>2</sub> H	C <sub>2</sub> P	H <sub>2</sub> CN	HONO	c-C <sub>3</sub> H <sub>2</sub>	<b>HC<sub>3</sub>S<sup>+</sup></b>		
CH <sup>+</sup>	N <sub>2</sub>	HCN	AlNC	HCNH <sup>+</sup>	MgC <sub>2</sub> H	<i>l</i> -C <sub>3</sub> H <sub>2</sub>	CH <sub>3</sub> OCH <sub>2</sub> OH		
CN	NO <sup>+</sup> ?	HNC	AlOH	H <sub>2</sub> CO	<b>HCCS</b>	CH <sub>2</sub> CN	<b>H<sub>2</sub>CCCHC<sub>3</sub>N</b>		
CO	NS <sup>+</sup>	HCO	H <sub>2</sub> O <sup>+</sup>	PH <sub>3</sub> ?	<b>HNCN</b>	H <sub>2</sub> C <sub>2</sub> O	<b>C<sub>2</sub>H<sub>5</sub>NCO</b>		
CO <sup>+</sup>	HeH <sup>+</sup>	HCO <sup>+</sup>	CaNC	C <sub>3</sub> H <sup>*</sup>	<b>H<sub>2</sub>NC</b>	H <sub>2</sub> CCN	<b>c-C<sub>6</sub>H<sub>4</sub></b>		
CP	CS	HOC <sup>+</sup>	<b>NCS</b>	c-C <sub>3</sub> H	<b>HCCS<sup>+</sup></b>	NH <sub>2</sub> CN			
SiC	HF	N <sub>2</sub> H <sup>+</sup>	NaCN	HC <sub>2</sub> N		HCOOH		<b>11 átomos</b>	
HCl	HD	HNO	C <sub>2</sub> S	HNCO		C <sub>4</sub> H	C <sub>2</sub> H <sub>5</sub> OCHO		
KCl	FeO?	H <sub>2</sub> S	H <sub>2</sub> Cl <sup>+</sup>	HCNO		C <sub>4</sub> H	CH <sub>3</sub> C <sub>6</sub> H		
NH	O <sub>2</sub>	HCP	KCN	HNCN?		HC <sub>3</sub> N	HC <sub>9</sub> N		
NO	CF <sup>+</sup>	CO <sub>2</sub>	HSC	HCO <sub>2</sub> <sup>+</sup>		HC <sub>2</sub> NC	CH <sub>3</sub> OC(O)CH <sub>3</sub>		
NS	SiH?	HCS <sup>+</sup>	NCO	H <sub>2</sub> CS		HNC <sub>3</sub>	CH <sub>3</sub> C(O)CH <sub>2</sub> OH		
NaCl	PO	C <sub>2</sub> O	HO <sub>2</sub>	C <sub>3</sub> H <sup>+</sup>		CNCHO	<b>c-C<sub>5</sub>H<sub>6</sub></b>		
OH	AlO	C <sub>3</sub> <sup>*</sup>	C <sub>2</sub> N	C <sub>3</sub> N		C <sub>5</sub> <sup>*</sup>	<b>HOCH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub></b>		
PN	OH <sup>+</sup>	MgNC	Si <sub>2</sub> C	C <sub>3</sub> N <sup>-</sup>		SiC <sub>4</sub>			
SO	CN <sup>-</sup>	MgCN	HS <sub>2</sub>	C <sub>3</sub> O		CH <sub>3</sub> Cl		<b>12 átomos</b>	
SO <sup>+</sup>	SiO	FeCN	NCO	HNCS		NCCNH <sup>+</sup>	c-C <sub>6</sub> H <sub>6</sub> <sup>*</sup>		
SiN	SiS	N <sub>2</sub> O	HNO	c-SiC <sub>3</sub>		C <sub>4</sub> S	<i>n</i> -C <sub>5</sub> H <sub>7</sub> CN		
	<b>6 átomos</b>	<b>7 átomos</b>	<b>8 átomos</b>	<b>9 átomos</b>		<b>HC<sub>3</sub>S<sup>+</sup></b>	C <sub>2</sub> H <sub>5</sub> OCH <sub>3</sub>		
	C <sub>2</sub> H <sub>4</sub>	CH <sub>3</sub> NH <sub>2</sub>	CH <sub>2</sub> CHCHO	CH <sub>2</sub> CHCH <sub>3</sub>		<b>H<sub>2</sub>C<sub>2</sub>S</b>	<i>i</i> -C <sub>3</sub> H <sub>7</sub> CN		
	CH <sub>3</sub> OH	CH <sub>3</sub> C <sub>2</sub> H	NH <sub>2</sub> CH <sub>2</sub> CN	CH <sub>3</sub> OCH <sub>2</sub>		<b>HC(O)SH</b>	<b>1-c-C<sub>2</sub>H<sub>5</sub>CN</b>		
	CH <sub>3</sub> CN	CH <sub>3</sub> CHO	CH <sub>3</sub> COOH	CH <sub>3</sub> CH <sub>2</sub> CN		<b>HC(O)SH</b>	<b>2-c-C<sub>2</sub>H<sub>5</sub>CN</b>		
	CH <sub>3</sub> NC?	c-C <sub>2</sub> H <sub>4</sub> O	CH <sub>2</sub> OHCHO	CH <sub>3</sub> CH <sub>2</sub> OH					
	CH <sub>2</sub> CNH	CH <sub>2</sub> CHOH	HC(O)OCH <sub>3</sub>	<b>H<sub>2</sub>C<sub>2</sub>S</b>				<b>&gt; 12 átomos</b>	
	NH <sub>2</sub> CHO	CH <sub>2</sub> CHCN	CH <sub>3</sub> C <sub>3</sub> N	<b>HC(O)SH</b>				C <sub>60</sub> <sup>*</sup>	
	CH <sub>3</sub> SH	C <sub>6</sub> H	CH <sub>2</sub> CCHCN	<b>HC(S)CN</b>				C <sub>70</sub> <sup>*</sup>	
	CH <sub>3</sub> SH	C <sub>6</sub> H <sup>-</sup>	C <sub>7</sub> H	C <sub>8</sub> H				C <sub>60</sub> <sup>+</sup> *	
	C <sub>4</sub> H <sub>2</sub>	HC <sub>3</sub> N	C <sub>6</sub> H <sub>2</sub>	C <sub>8</sub> H <sup>-</sup>				c-C <sub>6</sub> H <sub>5</sub> CN	
	HC <sub>4</sub> N	CH <sub>3</sub> NCO	H <sub>2</sub> C <sub>6</sub>	HC <sub>7</sub> N				HC <sub>11</sub> N	
	HC <sub>3</sub> NH <sup>+</sup>	HC <sub>5</sub> O	CH <sub>3</sub> CHNH	CH <sub>3</sub> NHCHO				<b>1-C<sub>19</sub>H<sub>7</sub>CN</b>	
	HC <sub>2</sub> CHO	HOCH <sub>2</sub> CN	CH <sub>3</sub> SiH <sub>3</sub>	HC <sub>7</sub> O				<b>2-C<sub>10</sub>H<sub>7</sub>CN</b>	
	c-H <sub>2</sub> C <sub>3</sub> O	HCCCHNH	H <sub>2</sub> NCONH <sub>2</sub>	<b>HC<sub>3</sub>HCHCN</b>				<b>c-C<sub>9</sub>H<sub>8</sub></b>	
	C <sub>5</sub> H	HC <sub>4</sub> NC	HCCCH <sub>2</sub> CN	<b>H<sub>2</sub>CCHC<sub>3</sub>N</b>					
	C <sub>5</sub> N	<b>c-C<sub>3</sub>HCCH</b>	HC <sub>5</sub> NH <sup>+</sup>	<b>H<sub>2</sub>CCHCCH</b>					
	C <sub>5</sub> N <sup>-</sup>	<b><i>l</i>-H<sub>2</sub>C<sub>5</sub></b>	<b>CH<sub>2</sub>CHCCH</b>						
	SiH <sub>3</sub> CN	<b>MgC<sub>5</sub>N</b>	<b>MgC<sub>6</sub>H</b>						
	<b>CH<sub>3</sub>CO<sup>+</sup></b>	<b>CH<sub>2</sub>C<sub>3</sub>N</b>	<b>C<sub>2</sub>H<sub>3</sub>NH<sub>2</sub></b>						
	<b>C<sub>3</sub>H<sub>3</sub></b>								
	<b>H<sub>2</sub>C<sub>3</sub>S</b>								
	<b>HCCCHS</b>								

**Nota.** Se señala en negrita aquellas especies químicas detectadas en el año 2021.

### 5.1. Origen de la vida: “glicina, la molécula perdida”

Enlazando con el punto anterior, la distribución de moléculas orgánicas de interés, y en concreto de aminoácidos, a través del universo podría contribuir a una mejor comprensión del origen mismo de la vida (Woon 2002). La química de los aminoácidos en el ISM, los cometas y los meteoritos es uno de los temas más intrigantes de la Química Orgánica Extraterrestre y la Astrobiología, debido a su condición de *building blocks of life*, ladrillos o bloques de construcción para la síntesis de proteínas, constituyentes esenciales de los seres vivos. Hasta el momento una de las evidencias indiscutibles a cerca de la existencia de aminoácidos fuera de nuestro planeta es su detección en condritas carbonosas (Cronin y col. 1976, 1983, 1989). Se cree que los aminoácidos han sido transportados por meteoritos y otros cuerpos pequeños (cometas, asteroides, partículas de polvo interplanetarias, etc.). Su origen es un punto clave para comprender cómo surgió la vida en la Tierra primitiva hace 3.500-3.700 millones de años (ver Figura 3).

---



**Figure 3.** Illustration of a storm of comets around Eta Corvi, a star located near our Sun, where a comet is depicted crashing into a rocky planet [Image credit: NASA/JPL-Caltech]

---

El bombardeo de la Tierra primitiva por cometas y asteroides implicó la aparición de una gran cantidad de material orgánico (Aponte y col. 2019). Este suministro de compuestos orgánicos, como los aminoácidos, podría tener una gran importancia para la astrobiología, pues es probable que planetas similares a la Tierra en las nebulosas extrasolares experimenten impactos similares. Por lo tanto, los productos de la Química

Orgánica Extraterrestre podrían, en un principio, permitir que la química prebiótica se produzca en toda la galaxia (Chyba y col. 1990).

Cabe destacar que, a pesar de su búsqueda extensiva, a día de hoy no se ha identificado el aminoácido más sencillo de todos, la **glicina** ( $\text{NH}_2\text{CH}_2\text{COOH}$ ). En 1994, un equipo de astrónomos de la Universidad de Illinois, dirigido por Lewis Snyder, afirmó que habían encontrado moléculas de glicina en el espacio (Mehringer y col. 1995), pero esta detección no fue confirmada por otros grupos de astrofísicos. Posteriormente, en 2003 Kuan (2003) reportó nuevamente su detección realizando una asignación de hasta 27 transiciones rotacionales, trabajo que fue rebatido posteriormente debido a errores del método y a identificaciones incorrectas de las líneas. La viabilidad de su detección está dificultada debida al gran tamaño de la molécula, presentando líneas débiles en su espectro de rotación (debido al tamaño de la función de partición). Por ello, cabría la posibilidad de que las transiciones buscadas se encuentren ocultas por la emisión de otras especies químicas. La determinación de parámetros espectroscópicos precisos de sus isómeros estructurales y su subsecuente identificación en el ISM supondría en última instancia una limpieza del *Line Survey*, haciendo accesible la detección de nuevas especies químicas.

Nuestro conocimiento acerca de la formación de la glicina en el espacio mediante diversos caminos químicos parece aún inconcluso (Ruiz-Miraz y col. 2014). Para emular su síntesis en el laboratorio encontramos dos grandes propuestas ampliamente discutidas en la bibliografía. La primera es la síntesis de Strecker (Strecker 1850), y la segunda consiste en el experimento de Miller (Miller 1953). En ambos casos el intermediato clave es el **formaldehído** ( $\text{HCHO}$ ). Otro reactivo de gran importancia, similar al formaldehído, es la **metanoimina** ( $\text{CH}_2=\text{NH}$ ), especie isoelectrónica con el  $\text{HCHO}$  y también se propone como precursor prebiótico para la glicina (Godfrey y col. 1973; Salter y col. 2008; Koch y col 2008).

En este contexto, las reacciones ión-molécula constituyen una vía de gran importancia para la formación de moléculas orgánicas en fase gas (Ferguson 1975, Herman & Futrell 2015) actuando en múltiples ocasiones como una posible fuente para la formación de dichas especies en el medio interestelar (Petrie y col. 2007, Larsson y col. 2012, Geppert & Larsson 2013). Hace algunos años, el grupo liderado por Bohme estudió en detalle la síntesis en fase gas de la glicina, concluyendo que las siguientes

reacciones permiten la formación de glicina protonada y agua (Snow y col. 2007, Blagojevic y col. 2003).



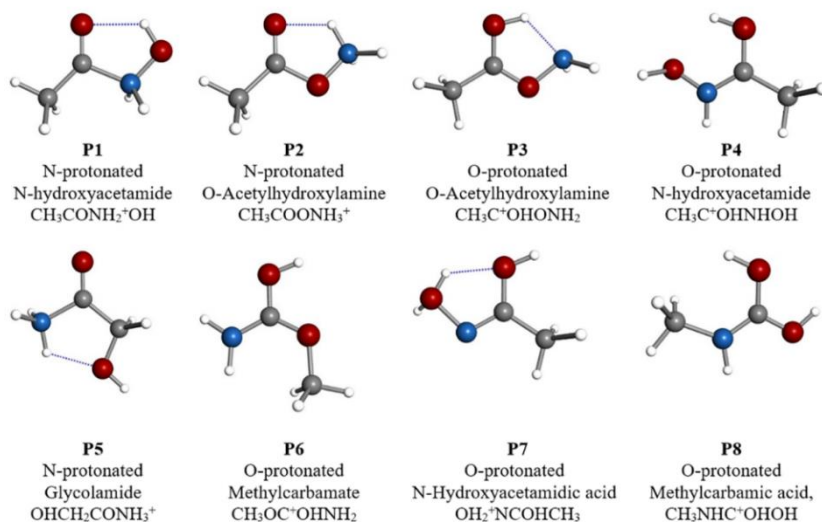
Como comentamos previamente, la detección de la glicina en el ISM fue desechada después de un gran debate dentro de la comunidad astrofísica (Snyder y col 2005, Cunningham y col. 2007, Jones y col. 2007). Recientemente, la misión espacial Rosseta dirigida al estudio del cometa 67P/Churyumov-Gerasimenko reportó la presencia de glicina junto con otras moléculas prebióticas (Altwegg y col. 2016). Además, la posibilidad de su generación vía reacción ión-molécula sigue siendo una temática de gran interés astrofísico, además de servir como reacción modelo para síntesis de COMs.

Diversos estudios teóricos y experimentales han propuesto la formación de glicina en hielos interestelares (Holtom y col. 2005) o por vía radicalaria (Rimola y col. 2012) variando los distintos caminos de reacción dependiendo de la composición del hielo (Elsila y col. 2007). Habitualmente se ha sugerido que la radiación UV es necesaria para iniciar dichas reacciones (Woon y col. 2002), pues hay que tener en cuenta que las reacciones en fase condensada pueden incluir más de dos reactivos, como apuntaron Maeda y Ohno para la formación de glicina vía reacción de  $\text{NH}_3 + \text{CH}_2 + \text{CO}_2$  (Maeda & Ohno 2004).

A finales de 2017, Largo y colaboradores investigaron mediante simulaciones de dinámica molecular las reacciones ion-molécula 1 y 2, realizadas experimentalmente por Bohme y colaboradores (Jeanvoine & Largo y col. 2018). Bohme caracterizó por espectrometría de masas en primer lugar todos los productos de relación masa-carga  $m/z$  76. A continuación, el espectro de fragmentación a diferentes energías se comparó con una muestra patrón de la glicina protonada encontrando ciertas similitudes. Largo y colaboradores demostraron que ambos tautómeros de la **hidroxilamina protonada** ( $\text{NH}_3\text{OH}^+$  y  $\text{NH}_2\text{OH}_2^+$ ), cuyo análogo neutro ha sido recientemente detectado en el ISM (Rivilla y col. 2020), permiten la formación de ocho especies moleculares (**P1-P8**) con la misma masa que la glicina protonada  $[\text{H}_6\text{C}_2\text{O}_2\text{N}]^+$  pero diferentes estructuras, siempre

que se aporte cierta energía translacional al sistema (ver Figura 4), siendo el producto **P1** la especie más abundante para ambas reacciones.

Entre todos los isómeros estudiados hasta la fecha, únicamente se ha buscado el **carbamato de metilo** en la nube molecular caliente W51e2 y en la protoestrella de masa intermedia IRAS21391+58502 (Demyk y col. 2004). Previamente, Groner y col. (2007) estudiaron su espectro de rotación en la región milimétrica y submilimétrica, generando datos en un intervalo muy amplio de frecuencia. Finalmente, la **glicolamida** se ha caracterizado utilizando espectroscopía con modulación de Stark entre 60 y 78.3 GHz (Maris 2004). Que se sepa, no hay disponible información estructural ni espectroscópica para el resto de las especies neutras.



**Figura 4.** Distintos iones de  $m/z$  76 obtenidos por colisiones de  $\text{NH}_3\text{OH}^+ + \text{CH}_3\text{COOH}$ .

Durante la presente tesis, se ha realizado un estudio computacional de la superficie de energía potencial (PES) correspondiente a las reacciones entre **ácido acético** y los dos tautómeros de la **hidroxilamina protonada**, generando diferentes productos, todos ellos obedeciendo a la fórmula molecular  $[\text{H}_6\text{C}_2\text{O}_2\text{N}]^+$ , recogido en la primera sección del Capítulo IV. Sorprendentemente, el único proceso exotérmico sin ninguna barrera de activación consiste en la generación del derivado protonado de la N-hidroxiacetamida o **ácido acetohidroxiámico**. Por tanto, se espera que su formación sea viable bajo condiciones interestelares, de forma que el correspondiente isómero neutro, el **ácido acetohidroxiámico** ( $\text{CH}_3\text{C}(\text{O})\text{NHOH}$ ) se propuso como molécula candidata a ser

detectada en el ISM. Sin embargo, al no existir ninguna información experimental previa en rotación para este sistema molecular, se decidió realizar un primer estudio experimental en la región de microondas, para posibilitar eventualmente su búsqueda en el ISM. Finalmente, se recoge un estudio exhaustivo de la **glicolamida** ( $\text{CH}_2(\text{OH})\text{C}(\text{O})\text{NH}_2$ ) hasta 460 GHz, otro isómero de gran relevancia de la glicina, junto con su búsqueda en Sgr B2 empleando el registro *EMoCA* y el radio telescopio Effelsberg-100m.

### 5.2. Nitrilos en el medio interestelar

Entre todos los COMs, los nitrilos, es decir, moléculas orgánicas que contienen un grupo funcional cianuro ( $-\text{C}\equiv\text{N}$ ), son algunas de las especies químicas mejor conocidas y más ampliamente distribuidas en el ISM. Desde la detección del radical diatómico CN en el ISM a comienzos de los años 40 (McKellar 1940), más de 30 nitrilos complejos – hablando desde un punto de vista astroquímico – han sido identificados en diversos tipos de entornos astronómicos. Cabe destacar la detección del ***i*-butironitrilo** ( $i\text{-C}_3\text{H}_7\text{CN}$ ) (Belloche y col. 2014), el **glicolonitrilo** ( $\text{CH}_2(\text{OH})\text{CN}$ ) (Zeng y col. 2019), el **benzonitrilo** ( $c\text{-C}_6\text{H}_5\text{CN}$ ), la primera especie aromática derivada del benceno detectada en el ISM [McGuire y col. 2018 (b)] y el **cianociclopentadieno**, su análogo con un anillo de cinco miembros (McCarthy y col. 2021), así como el **isocianógeno** ( $\text{CNCN}$ ), el miembro más sencillo de la familia de dicianopolinos (Agúndez y col. 2018). Muy recientemente, numerosos nitrilos han sido detectados en la prototípica nube molecular oscura TMC-1, destacando  $\text{HC}_4\text{NC}$ ,  $\text{HCCH}_2\text{CN}$ ,  $\text{HC}_5\text{NH}^+$ ,  $\text{H}_2\text{CCHC}_3\text{N}$  and  $\text{HCCCHCHCN}$  (Cernicharo y col. 2020; McGuire y col. 2020; Marcelino y col. 2020, y Lee y col. 2021), así como los isómeros 1- y 2- del cianonaftaleno (McGuire et al. 2021). No obstante, quedan aún muchos nitrilos por descubrir en el espacio interestelar. Por tanto, la realización de nuevos estudios teórico-experimentales que exploren especies relacionadas serán de vital importancia para la construcción de modelos sobre la formación de los nitrilos en el ISM, así como para predecir qué otros COMs pueden estar vagando en el espacio, a la espera de ser detectados.

El Capítulo V de la presente tesis recoge el primer estudio espectroscópico en la región de microondas de dos nitrilos complejos, candidatos a ser detectados en el ISM: la **ciano acetamida** ( $\text{CH}_2(\text{CN})\text{C}(\text{O})\text{NH}_2$ ) y el **ácido cianoacético** ( $\text{CH}_2(\text{CN})\text{C}(\text{O})\text{OH}$ ). Para ambos sistemas moleculares se ha determinado un set preciso de parámetros

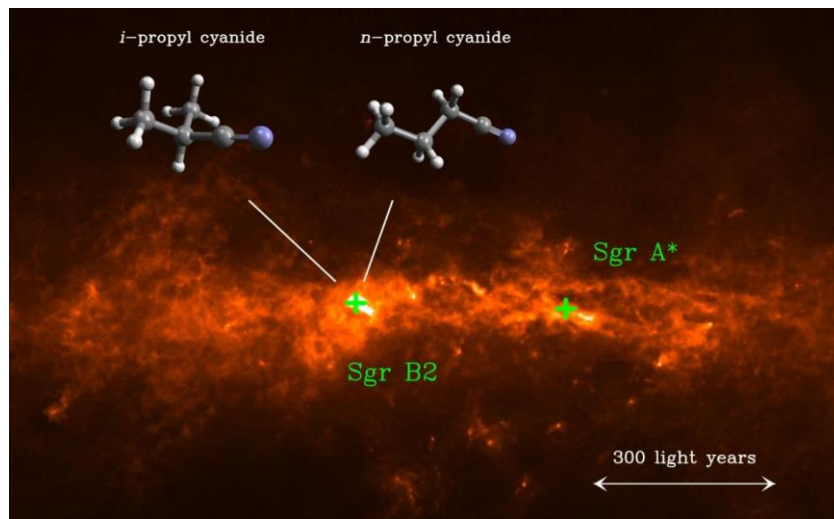
espectroscópicos de rotación, incluyendo las constantes de acoplamiento de cuadrupolo nuclear debida a la presencia de uno o más núcleos de  $^{14}\text{N}$ . La resolución detallada de la estructura hiperfina supondrá una cuestión fundamental para permitir la búsqueda e identificación de nuevas moléculas en el ISM, siempre y cuando se empleen registros de baja frecuencia (longitud de onda centimétrica). También se ha realizado un estudio teórico proponiendo un posible mecanismo de formación de la **ciano acetamida protonada** mediante reacciones ión-molécula en fase gas, factible bajo las condiciones del ISM.

### 5.3. Aldehídos y especies de cadena ramificada en el medio interestelar

La familia de los aldehídos comprende algunas de las especies químicas más extendidas en la naturaleza, siendo precursores de multitud de moléculas de gran relevancia biológica. Hace 50 años, Snyder y col. (1969) reportaron la detección del **formaldehído** ( $\text{H}_2\text{CO}$ ), el primer aldehído interestelar, en diversas fuentes (galácticas y extragalácticas) empleando el telescopio NRAO-140 ft. Desde ese momento, varios aldehídos de niveles de complejidad química cada vez mayores han sido descubiertos en el ISM, principalmente en la nube molecular gigante Sagitario B2 (Sgr B2), localizada a unos 120 parsecs (390 años luz) del centro de la Vía Láctea. Destacan la detección del **acetaldehído** ( $\text{CH}_3\text{CHO}$ , Gottlieb 1973), el **glicolaldehído** ( $\text{CH}_2(\text{OH})\text{CHO}$ , Hollis y col. 2000), que es la molécula más sencilla con relación con los azúcares, el **propionaldehído** ( $\text{CH}_3\text{CH}_2\text{CHO}$ , Hollis y col. 2004) y el **cianoformaldehído** ( $\text{CNCHO}$ , Remijan y col. 2008). Sin embargo, muchos aldehídos complejos, tales como el **lactaldehído** ( $\text{CH}_3\text{CH}(\text{OH})\text{CHO}$ , Alonso y col. 2019), están aún por descubrir.

Por otro lado, a la hora de seleccionar nuevos sistemas de estudio, la comunidad astrofísica suele guiarse por los últimos descubrimientos astronómicos. La detección del ***i*-butironitrilo**, previamente mencionado, fue de una importancia crucial al ser la primera molécula alifática de cadena ramificada detectada en el ISM (Belloche et al. 2014). Previamente, Belloche y col. (2009) reportaron la detección de su isómero lineal, ***n*-butironitrilo** ( $n\text{-C}_3\text{H}_7\text{CN}$ ,  $\text{CH}_3\text{CH}_2\text{CH}_2\text{CN}$ ), basada en observaciones del IRAM-30m de Sgr B2. La identificación de nuevas especies de cadena ramificada en el ISM nos permitirá descifrar cuál es la conexión entre el inventario molecular de nuestra Vía Láctea y la composición química de pequeños objetos celestiales, como cometas (Altwegg y col. 2016), asteroides, así como también de meteoritos, sus restos rocosos que llegan a la

Tierra (Pizzarello 2006; Pizzarello y col. 2010; Burton y col. 2012), donde han sido encontrados varios aminoácidos de cadena ramificada (Koga & Naraoka 2017).



**Figura 5.** Emisiones de polvo combinadas con datos obtenidos con el telescopio APEX y el observatorio espacial Planck. El *i*-butironitrilo (izquierda) junto con su isómero lineal, el *n*-butironitrilo fueron detectados en la región de formación de estrellas Sgr B2 con ALMA. [Crédito de la imagen: MPIfR/A. Weiss (fondo), University of Cologne/M. Koerber (modelos moleculares), MPIfR/A. Belloche (montaje)]

Asimismo, con el desarrollo y mejora de las nuevas instalaciones observacionales se espera que la detección de sistemas moleculares de gran tamaño crezca por encima incluso del *i*-butironitrilo (Belloche y col. 2014). Esta premisa se ve reforzada con alguna de las últimas detecciones astronómicas, destacando la identificación de los isómeros ya mencionados del **cianonaftaleno** (McGuire y col. 2021), junto con la observación del **indeno** [Cernicharo y col. 2021 (b); Burkhardt y col. 2021] y la **etanolamina** ( $\text{HOCH}_2\text{CH}_2\text{NH}_2$ , Rivilla y col. 2021).

En este contexto, el **Capítulo VI** de la presente Memoria recoge un estudio espectroscópico empleando espectroscopía de rotación de dos aldehídos alifáticos de crucial interés y extraordinaria complejidad: el ***n*-butiraldehído** ( $n\text{-C}_4\text{H}_8\text{O}$ ;  $\text{CH}_3\text{CH}_2\text{CH}_2\text{CHO}$ ) y la el isómero ramificado ***i*-butiraldehído** ( $i\text{-C}_4\text{H}_8\text{O}$ ;  $(\text{CH}_3)_2\text{CHCHO}$ ). Ambas especies se presentan como candidatas a ser detectadas en el ISM y han sido previamente identificadas en varios meteoritos condriticos (Aponte y col. 2019). Dicha investigación consiste en la realización de una caracterización exhaustiva en rotación de ambas especies en la región milimétrica y submilimétrica (hasta 325 GHz).



El análisis de sus espectros de rotación nos ha permitido también realizar su búsqueda interestelar en diferentes regiones del ISM. Además, se presenta en el Apéndice I un estudio puramente teórico del **amino acetaldehído**, dado que su experimentación sigue siendo a día de hoy inaccesible.

#### 6. *Estructura tridimensional de biomoléculas, sistemas fluxionales y sintones orgánicos*

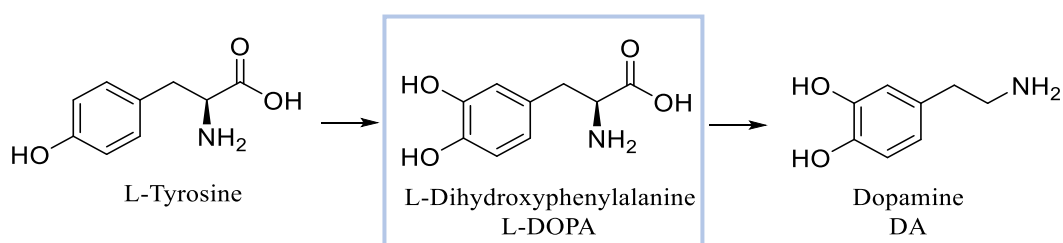
El medio natural o biológico en el que se despliega la química de las biomoléculas se corresponde con las fases condensadas. No obstante, tanto su panorama conformacional como su comportamiento fisiológico está condicionado por un balance de fuerzas intra- e intermoleculares, que dependerá enormemente de la naturaleza del propio medio. Son muchos los estudios espectroscópicos existentes en fases condensadas de biomoléculas, destacando los basados en resonancia magnética nuclear (NMR), infrarrojo (IR) o Raman, entre otros. No obstante, dichos estudios difícilmente proporcionan información sobre el sistema de estudio aislado, es decir, en ausencia de interacciones intermoleculares con el medio que lo envuelve.

Para abordar este problema, la sinergia entre los experimentos en fase gas basados en espectroscopía de rotación, destacando aquellos que combinan ablación laser (LA) con espectroscopía de microondas con transformada de Fourier (FTMW), y los cálculos teóricos ha demostrado ser una de las metodologías más robustas para investigar las propiedades estructurales de biomoléculas (Alonso & López 2015), debido fundamentalmente a las condiciones de aislamiento en fase gas lograda mediante la expansión supersónica (Levy 1980; Scoles 1988). Dicha metodología experimental se explica en detalle en el capítulo III.

Por otro lado, el acoplamiento con técnicas de vaporización laser ha supuesto el acceso al estudio de numerosas sustancias sólidas de elevados puntos de fusión, permitiendo la caracterización estructural de la mayoría de los aminoácidos naturales (Alonso & López 2015, Alonso y col. 2021), neurotransmisores [López y col. 2007; Ihiuchi y col. 2011, Cabezas y col. 2012, 2013(b)] junto con una larga lista de especies de gran relevancia biológica tales la vitamina C o la testosterona (Peña y col. 2013; León y col. 2021). Además de la información estructural obtenida, la exploración del panorama conformacional, es decir, las distintas disposiciones tridimensionales logradas mediante

giros entorno a enlaces sencillos (sin implicar ruptura o formación de nuevos enlaces), ayudará a clarificar la naturaleza de las relaciones estructura-propiedad. Por ejemplo, en el caso de los neurotransmisores, se espera que si una conformación específica es mucho más estable que el resto, esta será la favorita en el proceso de reconocimiento tipo “llave-cerradura” en el punto de anclaje del receptor (Ben-Naim 2001).

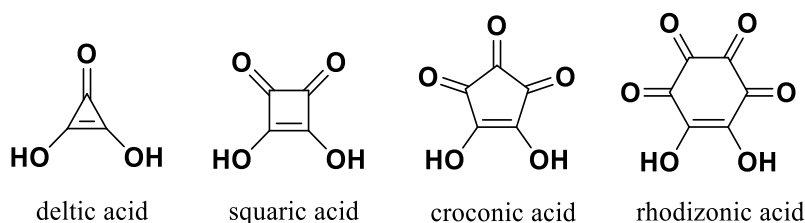
Siguiendo esta línea de investigación, durante la presente tesis se ha abordado el estudio conformacional y estructural de la **L-DOPA** (ver Figura 5), precursor directo de la dopamina [Cabezas y col. 2013 (b)], cuyo estudio en rotación ha permanecido como una tarea muy desafiante para la espectroscopía de alta resolución durante la última década, debido a su inherente inestabilidad y la difícil vaporización de la muestra, También se han estudiado otras dos moléculas orgánicas de gran importancia para la bioquímica y la química general, respectivamente: el **ácido escuárico**, considerado como uno de los óxidos de carbon, o oxocarbonos, cíclicos arquetípicos (mostrados en la Figura 6), y la **fenil-barbaralona**. Los estudios de estas tres especies químicas se presentan conjuntamente en el **Capítulo VII**.



**Figura 5.** Esquema del proceso clásico de biosíntesis de la dopamina (Meiser y col. 2013).

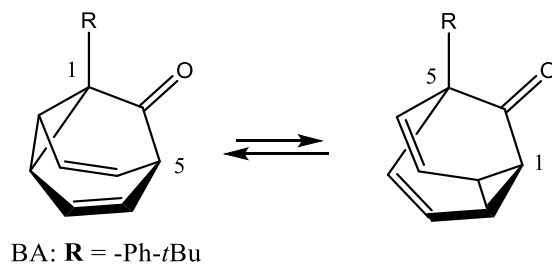
En la última década, el **ácido escuárico** y sus derivados han sido muy empleados como sintones fundamentales para el diseño de nuevas estructuras cristalinas y de algunos *metal-organic frameworks* (MOFs) dentro la química supramolecular (Das y col 2005; Wang y col. 2007). Sus derivados presentan aplicaciones muy relevantes dentro de la Química Orgánica (Chen y col. 2014; Andrés y col. 2017; Li y col. 2018), Química Biomédica (Liu y col. 2013; Hill y col. 2015) y Ciencia de Materiales (Skopinska-Wisnewska y col. 2016). No obstante, la estructura tridimensional de la molécula aislada,

así como la de sus agregados con agua ha permanecido desconocida hasta la fecha, siendo un un sistema de estudio muy atractivo para la espectroscopía de rotación de alta resolución.



**Figura 6.** Ácidos oxocarbónicos cíclicos de fórmula  $(C_nO_nH_2)$ .

Finalmente, la caracterización e interpretación de sistemas constitucionales dinámicos permanece como un desafío intrigante para la química física orgánica. En concreto el fenómeno de tautomerismo de valencia consiste en una reorganización intramolecular que implica una única especie reactiva, donde un enlace intramolecular migra dando lugar a una estructura normalmente diferente. En este contexto, se ha estudiado la familia de las **barbaralonas**, moléculas fluxionales profundamente arraigadas a la comprensión de los fenómenos de tautomerismo de valencia. Dichas especies pueden experimentar procesos [3, 3]-sigmatrópicos, es decir reorganizaciones estructurales donde el número de isómeros constitucionales (en este caso tautómeros de valencia) depende de la presencia y naturaleza química de cualquier posible sustituyente, que induzca alguna restricción sobre el proceso de isomerización.



**Figura 7.** Equilibrio sigmatrópico para barbaralonas 1-sustituidas (por ejemplo la fenil barbaralona).

Hasta la fecha no tenemos conocimiento de ningún estudio en fase gas empleando espectroscopía de rotación enfocado en la investigación de los procesos [3.3]-sigmatrópicos. Por tanto, resulta de gran interés el empleo de técnicas espectroscópicas de alta resolución para descifrar dicho equilibrio. Como una prueba de concepto, escogimos la **fenil-barbaralona** para obtener información acerca de estos sistemas y del fenómeno de tautomerismo de valencia en ausencia de interacciones intermoleculares, así como también caracterizar la estructura tridimensional de estas moléculas fluxionales.

### *7. Contenido de la Memoria*

El contenido de la presente tesis debe entenderse, por tanto, en dos vertientes diferentes. Una primera, y columna vertebral de la tesis, está centrada en la caracterización espectroscópica de moléculas de relevancia astroquímica y su posible identificación o búsqueda en el ISM, así como el estudio teórico y análisis de algunos de los posibles procesos de formación (reacciones ión-molécula) en fase gas a partir de moléculas ya detectadas en el ISM. La segunda propone el estudio estructural y conformacional de diversas moléculas orgánicas y de relevancia biológica. En la Tabla 2 se presentan los nombres y fórmulas químicas de los principales sistemas estudiados experimental y/o teóricamente durante esta tesis.

A lo largo de esta tesis doctoral se mostrará que, para comprender adecuadamente los diferentes caminos para la evolución de sistemas químicos de mayor complejidad, es necesario realizar un esfuerzo coordinado y meticuloso multidisciplinar (observacional, teórico y de laboratorio). En concreto, este trabajo demuestra la gran importancia de los datos obtenidos en el laboratorio empleando espectroscopía de rotación para responder cuestiones fundamentales para la Astronomía y la Astroquímica, tales como el nivel de complejidad química alcanzado en el ISM, buscando en última instancia aportar información sobre aquellas especies relacionadas con el origen de la vida en la Tierra.

**Tabla 2.** Listado de los sistemas moleculares estudiados en la presente tesis.

Capítulo	Sistemas moleculares	Fórmula química	Búsqueda ISM
IV: Experimental and theoretical study of interstellar glycine isomers	Isómeros de la glicina protonada	$[H_6C_2O_2N]^+$	No
	Glicolamida	$CH_2(OH)C(O)NH_2$	Si
	Ácido acetohidroxiámico	$CH_3C(O)NHOH$	No
V: Hunting complex cyanides in the interstellar medium	Cianoacetamida	$CH_2(CN)C(O)NH_2$	Si
	Ácido cianoacético	$CH_2(CN)C(O)OH$	Si
VI: Toward de limits of complexity of interstellar chemistry: rotational spectroscopy and astronomical search for n- and i-butiraldehyde	<i>n</i> -butiraldehído	$CH_3CH_2CH_2CHO$	Si
	<i>i</i> -butiraldehído	$(CH_3)_2CHCHO$	Si
VII: The shape of biomolecules, fluxional systems and organic synthons	L-DOPA <sup>[a]</sup>	$C_9H_{11}NO_4$	-
	Ácido escuárico	<i>c</i> - $C_4H_2O_4$	-
	Fenil-barbaralona	$C_{15}H_{12}O$	-
Apéndice I	Amino acetaldehído	$NH_2CH_2CHO$	Si <sup>[b]</sup>

**Nota:** [a] Tanto la L-DOPA, el ácido escuárico y la fenil-barabralona son sistemas de relevancia orgánica y bioquímica, que no son candidatos para la búsqueda en el ISM; [b] Hasta la fecha, los parámetros espectroscópicos predichos, incluso a los niveles de teoría más altos posibles, no permiten una identificación inequívoca de la molécula en el ISM (se estima una precisión de aproximadamente 0.2% para los valores de las frecuencias). No obstante, varios astrofísicos han tratado, sin éxito, de buscar esta molécula en el ISM (Belloche 2021, com. priv.).

## 8. Referencias

- Agúndez, M., Marcelino, N., & Cernicharo, J. **2018**, *ApJ*, 861, L22.
- Andrés, J.M., Losada, J., Maestro, A., Rodríguez-Ferrer, P., Pedrosa, R., **2017**, *J. Org. Chem*, 82, 8444,8454.
- Alonso, E. R., Kolesniková, L., Peña, I., Shipman, S. T., Tercero, B., Cernicharo, J., Alonso J. L., **2015**, *J. Mol. Spectrosc.*, 316, 84–89.
- Alonso, E. R., Kolesniková, L., Alonso, J. L., **2017**, *J. Ch.em. Phys.* 147, 124312.
- Alonso, E. R., Kolesniková, L., Białkowska-Jaworska, E., Kisiel, Z., León, I. Guillemin J.-C., & Alonso, J. L., **2018**, *ApJ*, 861, 70-77.
- Alonso, E. R., McGuire, B. A., Kolesniková, L., y col., **2019**, *ApJ*, 883, 18.
- Alonso, E.R., Kolesniková, L., Belloche A., Mata. S. y col. **2021**, *A&A* 647, A55 (a)
- Alonso, E. R., León, I., Alonso, J. L., *Intra- and Intermolecular Interactions between Non-covalently Bonded Species*. Elsevier, **2021**, 93-141 (b)
- Alonso, J. L., López, J. C., Springer Berlin Heidelberg, **2015**, pp. 1-67.

- Altwegg, K., Balsiger, H., Bar-Nun, A., Berthelier, J.-J., y col. **2016**, *Sci. Adv.* 2, e1600285.
- Aponte, J. C., Whitaker, D., Powner, M. W., y col. **2019**, *ACS Earth and Space Chemistry*, 3, 463.
- Aprile, E. & Profumo, S., **2009**, *New J. Phys.* 11 105002.
- Barone V, Biczysko M, Puzzarini C., **2015**, *Acc. Chem. Res.*, 48:1413–1422.
- Barrientos, C., Redondo, P., Largo, L., y col. **2012**, *ApJ*, 748, 99
- Barrientos, C., Rayón, V.M., Largo, A., Sordo, J.A., Redondo, P. **2013**, *Journal of Physical Chemistry*, 117, 7742.
- Belloche, A., Garrod, R. T., Müller, H. S. P., y col. **2009**, *A&A*, 499, 215.
- Belloche, A., Garrod, R. T., Müller, H. S. P., y col. **2014**, *Science*, 345, 1584.
- Ben-Naim, A. *Cooperativity and Regulation in Biochemical Processes*, Kluwer Academic/Plenum Publishers, New York, **2001**.
- Bergner, J. B., Guzmán, V. G., Öberg, K. I., Loomis, R. A., & Pegues, J. **2018**, *ApJ*, 857, 69.
- Biczysko, M., Bloino, J., Puzzarini, C. **2018**, *Comput. Mol. Sci.*, 8:e1349.
- Blagojevic, V.; Petrie, S.; Bohme, D. K., **2003**, *Mon. Not. R. Astron. Soc.* 339, L7–L11.
- Burkhardt, A. M., Long Kelvin Lee, K., Bryan Changala, P., y col., **2021**, *ApJ*, 913, L18.
- Burton, A. S., Stern, J. C., Elsila, J. E. y col., **2012**, *Chem. Soc. Rev.*, 41, 5459-5472.
- Cabezas, C.; Varela, M.; Peña, I.; Lopez, J. C.; Alonso, J. L., **2012**, *Phys. Chem. Chem. Phys.*, 14, 13618–13623.
- Cabezas, C., Cernicharo, J., Alonso, J.L., Agúndez, M., Mata, S., Guelin, M., Peña, I., **2013**, *ApJ*, 775, 133 (a).
- Cabezas, C.; Peña, I. López, J. C.; Alonso, J.L., *J. Phys. Chem. Lett.* **2013**, 4, 486-490 (b).
- Cernicharo, J., Kisiel, Z., Tercero, B., Kolesniková, L., y col. , **2016**, *A&A*, 587, L4.
- Cernicharo, J., Marcelino, N., Angúdez, M., y col., **2020**, *APJL*, 900, L9.
- Cernicharo, J., Cabezas, C., Endo, Y. y col. **2021**, *A&A*, 650, L14. (a)
- Cernicharo, J., Agúndez, M., Cabezas, C., y col., **2021**, *A&A*, 649, L15. (b)
- Chen, S., Pan, J., Wang, Y., Zhou, Z., **2014**, *Eur. J. Org. Chem.*, 7940.
- Chyba, C. F., Thomas, P. J., Brookshaw, L., & Sagan, C. **1990**, *Science*, 249, 366.
- Chyba, C., & Sagan, C. **1992**, *Nature*, 355, 125.
- Chyba, C. F., & Hand, K. P. **2005**, *ARA&A*, 43, 31.
- Cronin, J. R., **1989**, *Advances in Space Research*, 9, 59.
- Cronin, J. R., & Moore, C. B., **1976**, *Geochim. Cosmochim. Acta*, 40, 853.
- Cronin, J. R., & Pizzarello, S., **1983**, *Advances in Space Research*, 3, 5.
- Cunningham, M. R.; Jones, P. A.; Godfrey, P. D.; Cragg, D. M.; y col. **2007**, *Mon. Not. R. Astron. Soc.* 376, 1201–1210.
- Dae, N., Ghosh, A., Arif, M. & Stan, P.J., **2005**, *Inorg. Chem.*, 44, 7130-7137.
- Daly, A.M., Kolesniková, L., Mata, S. & Alonso J.L., **2014**, *J. Mol. Spectrosc.*, 306, 11–18.
- Daly, A.M., Bermúdez, C., Kolesniková, L., Alonso, J.L. **2015**, *ApJS*, 218:30 (7pp) .
- Dartois, E. **2005**, *Space Sci. Rev.*, 119, 293.

- Demyk, K.; Wlodarczak, G.; Dartois, E. in *Semaine de l'Astrophysique Francaise*, ed. Combes, F.; Barret, D.; Contini, T.; Meynadier, F.; Pagani, L. SF2A-2004 (Les Ulis: EDP-Sciences), **2004**, 493.
- Draine, B. T., **2003**, *Annu. Rev. Astron. Astrophys.* 41 241–89.
- Draine, B. T. *Physics of the Interstellar and Intergalactic Medium*, Princeton University Press, **2011**. ISBN: 978-0-691-12214-4.
- Elsila, J. E.; Dworkin, J. P.; Bernstein, M. P.; Martin, M. P.; Sandford, S. A. **2007**, *APJ.*, 660, 911–918.
- Ferguson, E. E. **1975**, *Annu. Rev. Phys. Chem.*, 26, 17–38.
- Fortman, S. M., Medvedev, I. R., Neese, C. F., & De Lucia, F. C. **2010**, *ApJ*, 725, 1682.
- Garrod, R. T. & Herbst, E. **2006**, *A&A*, 457, 927.
- Garrod, R. T. **2013**, *ApJ*, 765, 60.
- Geppert, W. D.; Larsson, M., **2013**, *Chem. Rev.* 113, 8872– 565 8905.
- Godfrey, P. D., Brown, R. D., Robinson, B. J. & Sinclair, M.W. **1973**, *APJL*, 13, 119–121.
- Goldsmith, P., Bergin, T., De Lucia, F. C., et al. **2006**, Report from the Workshop on Laboratory Spectroscopy in Support of Herschel, Sofia, and Alma, Tech. Rep. (Pasadena, CA: California Institute of Technology)
- Gottlieb, C. A. **1973**, *Molecules in the Galactic Environment*, 181
- Groner, P.; Winnewisser, M.; Medvedev, I. R, et al.. **2007**, *Astrophys. J. Suppl. Series.*, 169, 28-36.
- Herman, Z.; Futrell, J. H. **2015**, *Int. J. Mass Spectrom.* 377, 84–92.
- Jortner, J. **2006**, *Phil. Trans. R. Soc. B*, 361, 1877.
- Jørgensen, J. K., Belloche, A., & Garrod, R. T. **2020**, *ARA&A*, 58, 727.
- Koga, T. & Naraoka, H. **2017**, *Scientific Reports*, 7, 636.
- Herbst, E., & van Dishoeck, E. F. **2009**, *ARA&A*, 47, 427.
- Hill, N.D., Bunata, K., Hebert, A., **2013**, *Clinics in Dermatology*, 33, 300-304.
- Hollis, J. M., Lovas, F. J., & Jewell, P. R. **2000**, *ApJ*, 540, L107.
- Hollis, J. M., Jewell, P. R., Lovas, F. J., y col., **2004**, *ApJ*, 610, L21
- Holtom, P. D.; Bennett, C. J.; Osamura, Y.; Mason, N. J.; Kaiser, R. I., **2005**, *APJ*, 626, 940 – 952.
- Ishiuchi, S.; Asakawa, T.; Mitsuda, H.; Miyazaki, M.; Chakraborty, S.; Fuvii, M., **2011**, *J. Phys Chem. A*, 115, 10363–10369.
- Jeanvoine, Y.; Largo, A.; Hase, W.L.; Spezia R.; **2018**, *J. Phys. Chem. A*. 122: 869-877.
- Jones, P. A.; Cunningham, M. R.; Godfrey, P. D.; Cragg, D. M. **2007**, *Mon. Not. R. Astron. Soc.*, 374, 579–589.
- Kallman, T. R., and Palmeri, P., **2007**, *Rev. Mod. Phys.* 79 79–133
- Kappeler, F., Gallino, R., Bisterzo, S. & Aoki, W. **2011** *Rev. Mod. Phys.* 83 157–93.
- Koch, D. M., Toubin, C., Peslherbe, G. H., & Hynes, J. T. J. **2008**, *Phys. Chem. C*, 112, 2972–2980.
- Kolesníková, L., Tercero, B. Cernicharo, J., Alonso, J. L., y col. **2014**, *ApJ*, 784 L7.
- Kolesníková, L., Tercero, B., Alonso, E. R., y col. **2017**, *A&A*, 609, A24.
- Kolesníková, L. Alonso; E.R., Mata, S., Alonso, J.L. **2017**, *A&A Supplement Series*, 229:26 (8pp).
- Kolesníková, L., León, I., Alonso, E. R., Mata, S., Alonso, J. L., **2019**, *J. Phys. Chem. Lett.*, 10, 1325-

1330.

- Kuan Y.-J. *Astrophys. J.*, **2003**, 848 – 867.
- Larsson, M.; Geppert, W. D.; Nyman, G. **2012**, *Rep. Prog. Phys.* 75, 066901.
- Lee, K. L. L., Changala, P. B., Loomis, R. A., y col. **2021**, *ApJ*, 910, L2.
- León, I., Alonso, E. R., Mata, S., Alonso, J.L., **2021**, *J. Phys. Chem. Lett.*, 29; 12 (29): 6983-6987
- Levy D.H., **1980**, *Annu Rev Phys Chem.*, 31:197, 2.
- Li, S.C., Zhang, T., Deng, X. P., y col. **2018**, *Inor. Chem. Comm*, 92, 59-73.
- Liu, Z. Wang U, Han W y col. **2013**, *European Journal of Medical Chemistry*, 65, 187-194.
- Lopez, J. C.; Cortijo, V.; Blanco, S.; Alonso, J. L. **2007**, *Phys. Chem. Chem. Phys.*, 9, 4521– 4527.
- Scoles G, *Atomic and molecular beam methods*, vol 1. Oxford University Press, New York/Oxford, **1988**.
- Loomis, R. A., Burkhardt, A. M., Shingledecker C. N. y col. **2021**, *N. Nat. Astron.* 5, 188–196.
- Maeda, S.; Ohno, K. **2004**, *Chem. Phys. Lett.*, 398, 240–244.
- Marcelino, N., Angúdez, M., Tercero, B., y col., **2020**, *A&A*, 643, L6.
- Maury, A. J., Belloche, A., André, P., y col. **2014**, *A&A*, 563, L2.
- Maris, A. **2004**, *Phys. Chem. Chem. Phys.*, 6, 2611-2616.
- McCarthy, M. C., Kelvin Lee, K. L., Loomis, R. A. y col., **2021**, *Nature Astronomy*, 5, 176-180.
- McGuire, B. A. **2018**, *ApJS*, 2. (a)
- McGuire, B. A., Burkhardt, A. M., Kalenskii, S., y col. **2018**, *Science*, 359, 202. (b)
- McGuire, B., Burkhardt, A., Loomis, R., y col., **2020**, *APJL*, 900, L110.
- McGuire, B. A., Loomis, R. A., Burkhardt, A. M., y col. **2021**, *Science*, 371,1265.
- McKellar, A. **1940**, *PASP*, 52, 187.
- Mehringer, D. M., Miao, Y., Kuan, Y.-J., & Snyder, L. E., **1995**, *American Astronomical Society Meeting Abstracts*, 184, 26, 906.
- Meiser, J., Weindl, D. & Hiller, K., **2013**, *Cell Commun. Signal*, 11, 34.
- Miller, S. L., **1953**, *Science*, 117, 528–529. Z
- I. Peña, A. M. Daly, C. Cabezas, S. Mata, C. Bermúdez, A. Niño, J.C. López, J-U. Grabow, and J. L. Alonso, **2013**, *J. Phys. Chem. Lett.*, 4(1), pp 65–69.
- Petrie, S.; Bohme, D. K. Ions in space. **2007**, *Mass Spectrom. Rev.*, 26, 258–280.
- Pizzarello, S. **2006**, *Acc. Chem. Res*, 39, 321.
- Pizzarello, S., Shock, E. **2010**, *Cold Spinrg Harb. Perspect. Biol.*, 2, 1-20.
- Puzzarini C, Biczysko M, Bloino J, Barone V. **2014**, *APJ*, 785:107.
- Redondo, P.; Martínez, H.; Cimas, A.; Barrientos, C.; Largo, A. **2013**, *Phys. Chem. Chem. Phys.* 15, 13005.
- Redondo, P., Barrientos, C., & Largo, A. **2014**, *ApJ*, 793, 32
- Redondo, P., Largo, A., & Barrientos, C. **2015**, *A&A*, 579, A125
- Redondo, P., Martínez, H., Largo, A., y col. **2017**, *A&A*, 603, A13
- Redondo, P.; Largo, A.; Barrientos, C., *ApJ*, **2017**, 836, 240.

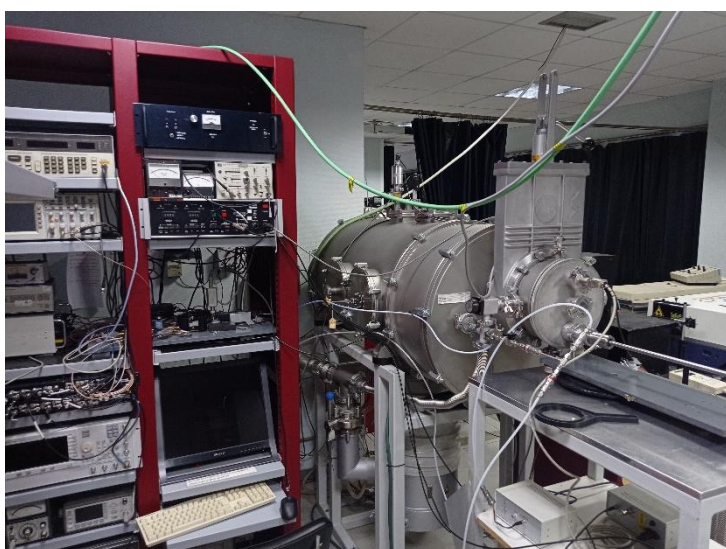


- Remijan, A. J., Hollis, J. M., Lovas, F. J., y col., **2008**, *ApJ*, 675, L85.
- Remington, B. A., Drake, R. P., & Ryutov, D. D., **2006**, *Rev. Mod. Phys.* 78 755–807.
- Rimola, A.; Sodupe, M.; Ugliengo, P., **2012**, *APJ*, 754, 24.
- Rivilla, V. M., Martín-Pintado, J., Jiménez-Serra, I., y col. **2020**, *ApJ*, 899, L28.
- Rivilla, V. M., Jiménez-Serra, I., Martín-Pintado, J., y col. **2021**, *Proceedings of the National Academy of Science*, 118, 2101314118.
- Ruiz-Mirazo, K., Briones, C., & de la Escosura, **2014**, *A. Chem. Rev.*, 114, 285–366.
- Sakai, N. & Yamamoto, S. 2011, 280, 43.
- Salter, C. J., Ghosh, T., Catinella, B., Lebron, M., Lerner, M. S., Minchin, R. and Momjian, E., **2008**, *Astron. J.*, 136, 389–399.
- Sanz-Novo, M., Belloche, A., Alonso, J. L., y col. **2020**, *A&A*, 639, A135.
- Savin, D. W., Brickhouse, N. N., Cowan, J.J. y col. **2012**, *Rep. Prog. Phys.* 75, 036902.
- Skopinska-Wisniewska, J. Kuderko, J., Bajek, A. y col., **2016**, *Materials Science and Engineering C*, 60, 100-109.
- Snow, J. L.; Orlova, G.; Blagojevic, V.; Bohme, D. K., **2007**, *J. Am. Chem. Soc.* 129, 9910–9917.
- Snyder, L. E., Buhl, D., Zuckerman, B., et al. **1969**, *Phys. Rev. Lett.*, 22, 679.
- Snyder, L. E.; Lovas, F. J.; Hollis, J. M.; Friedel, D. N.; Jewell, P. R.; Remijan, A. **2005**, *APJ*, 619, 914–930.
- Spitzer, L. *Physical Processes in the Interstellar Medium*, New York Wiley-Interscience, **1978**, 333.
- Strecker, A., **1850**, *Annalen der Chemie und Pharmazie*, 75, 27–45.
- Tercero, B., Cernicharo, J. López, A., Brouillet, N. y col., **2015**, *A&A* 582, L1.
- Thaddeus, P. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **2006** 361, 1681.
- Tielens, A. G. G. M., **2013**, *Rev. Mod. Phys.* 85. 3.
- Vazart F, Calderini D, Puzzarini C, Skouteris D, Barone V., **2016**, *J. Chem. Theory Comput*, 12:5385–5397.
- Walsh, C., Loomis, R. A., Öberg, K. I., y col., **2016**, *ApJ*, 823, L10.
- Wiescher M, Gorres J, Überseder E, Imbriani G and Pignatari M **2010**, *Annu. Rev. Nucl. Part. Sci.*, 60 381–404.
- Willis, E. R. & Garrod, R. T. **2017**, *ApJ*, 840, 61.
- Wilson, T. Rohlfs, K., Huettmeister, S. *Tools of Radio Astronomy*, Springer-Verlag Berlin Heidelberg, 6ª edición, **2013**.
- Woon, D. E. **2002**, *ApJ*, 571, L177.
- Yamada, M., Kulsrud, R., & Ji H. **2010**, *Rev. Mod. Phys.* 82 603–64
- Zeng, S., Quénard, D., Jiménez-Serra, I., y col. **2019**, *MNRAS*, 484, L43-L48.



## CHAPTER III. METHODOLOGY

---



This chapter is divided in four fundamental sections. First, after a brief historical introduction about rotational spectroscopy, a description of each of the instruments used during this Ph.D will be presented (see Section 1). In this context, two main different spectroscopic techniques emerge: time domain and frequency domain spectroscopies. Afterward, in Section 2, the general work strategy used to carry out the experiments in this thesis will be given. In the present Dissertation the synergy between theory and experiments is evident, thus, it is necessary to treat in detail the computational methods used during the investigation of the systems under study (Section 3). Finally, in the last section of the chapter, the main tools needed for the analysis of a rotation spectrum will be briefly introduced.

## 1. EXPERIMENTAL TECHNIQUES: ROTATIONAL SPECTROSCOPY

High-resolution rotational spectroscopy is acknowledged among the most robust techniques to distinguish between subtle changes in the three-dimensional structure of gas-phase molecules (Graner et al. 1998). It is a very suitable technique for providing experimental data that enables eventual astronomical searches for molecules in the ISM, as well as to unveil the shape of relevant organic and biomolecules in the gas phase. The only required criteria is that the molecule must present a permanent electric dipole moment (Gordy & Cook, 1984).

Briefly, experiments based on rotational spectroscopy involve the interaction of molecules in gas phase with electromagnetic radiation, ranging from the microwave (mw), with a wavelength between 1 cm – 1 m, to the millimeter (mmw), with wavelengths of 1 mm – 1 cm, and sub-millimeter (sub-mmw) regions. Regarding time domain spectroscopy, its foundation relies on the fact that if the radiation-matter interaction is resonant with a molecular transition, it will induce a macroscopic polarization of the gaseous sample. Then, its “decay”, consequence of the return of population to the low energy levels, will be recorded in the form of one or more "free induction decays" (FIDs), coherent with the molecular emission. Afterward, the time domain spectrum is Fourier transformed into the frequency domain. Finally, its analysis will entail an iterative procedure of fittings and predictions (described in detail in [Section 3.4](#) of this chapter) that will enable to obtain a set of spectroscopic parameters to fully describes each molecular system.

Microwave spectroscopy has evolved and progressed remarkably in recent decades through various technological milestones. The first major landmark consisted in the development of the radar technology, which advanced significantly during World War II. In this regard, the physicist Robert Watson-Watt was one of the main pioneers responsible for the construction of the modern radar model. In the 1980s and early 1990s, Balle and Flygare developed a Fourier transform microwave (FTMW) spectrometer, which coupled a pulsed microwave source with a Fabry-Pérot cavity, and also used supersonic expansion to introduce the sample (Balle et al. 1979, Balle & Flygare 1981). Through successive improvements, emphasizing the coaxial arrangement of microwave radiation with respect to the molecular beam (MB-FTMW) (Grabow et al. 1996), this

technique quickly became one of the fundamental pillars of rotational spectroscopy. At the beginning of the 21st century, Brooks Pate and co-workers presented the design of a Chirped-pulse (CP) spectrometer. They introduced the use of broadband horn antennas, thus replacing the resonant cavity of the previous Balle-Flygare design (Brown et al. 2008). Although this new setup is based on the same principles as the MB-FTMW instrument, allowed the use of excitation pulses that cover large frequency ranges. The term “Chirped-pulse” consists of a linear frequency sweep performed in a short interval of time. Hence, it enables to polarize a wide range of frequencies with a single radiation pulse, remarkably reducing the spectral acquisition time, and also allowing to simultaneously analyze all the species populating the jet (tautomers, conformers... etc.).

To sum everything up, microwave spectroscopy coupled with a supersonic expansion have led to the development of two families of spectroscopic techniques: narrowband or cavity based FTMW techniques (MB-FTMW) (Andresen et al. 1990, Grabow et al. 1996) and broadband Chirped-pulse FTMW techniques (CP-FTMW) (Brown et al. 2008). These techniques have been largely employed to studying easily vaporized samples or gaseous compounds, along with short-lived species generated *in situ* via electric discharge procedures (Endo et al. 1994, Sutter & Dreizler, 2001; McCarthy et al. 2001, 2021). Nevertheless, many organic and biologically relevant molecular systems are solids with relatively high melting points which exhibit very low vapor pressures at room temperature. Some of them even decompose when exposed to conventional heating procedures, further preventing their easy transfer to the gas phase and therefore being out of scope of conventional rotational studies (Suenram et al. 1988; Welzel et al. 1999; Blanco et al. 2003). To address such molecular systems, a laser ablation vaporization technique was developed in our group.

### *1.1 Laser ablation time domain Fourier transform microwave spectroscopy*

In the nineties, Suenram et al. (1990) and Walker & Gerry (1997) developed separately two distinct experimental setups combining a narrowband microwave spectrometer with a laser ablation vaporization system (LA-MB-FTMW) to study several metal-containing species. However, both teams did not succeed to study organic solid samples using the abovementioned devices (Lovas et al. 1995, Kretschmer et al. 1996).

Laser ablation has been proved to be a very efficient tool for the ultrafast vaporization of solid samples and has been applied in a wide range of experiments (Russo et al. 2013; Stafe et al 2013; Kerse et al. 2016; Bergevin et al. 2018). In this context, over the last years, we have combined laser ablation with the state-of-the-art Fourier transform microwave instrumentation (both narrowband and broadband setups) (Lesarri et al. 2003, Bermúdez et al. 2014; Mata et al. 2012) to study the three-dimensional shape of the neutral forms of natural amino acids in a supersonic jet (Alonso et al 2015, León et al. 2019). These experimental approaches have also been extended to study species of astrochemical relevance (Cabezas et al. 2013; Alonso et al. 2018; Kolesníková et al. 2019), along with many other building blocks of life (Peña et al 2015; Cabezas et al. 2017; Alonso et al. 2021) and weakly bounded molecular complexes (Alonso et al. 2006).

Very recently, new applications of the laser ablation have been reported in our laboratory, highlighting an innovative experimental approach devoted to the generation of interstellar species in a terrestrial laboratory (Kolesníková et al. 2021). Thus, with this new developed “micro-laboratory”, molecules of astrophysical relevance such as aminocynoacetylene have been generated for the first time together with a vast inventory of already detected interstellar compounds by laser ablation of diaminomaleonitrile as solid organic precursor.

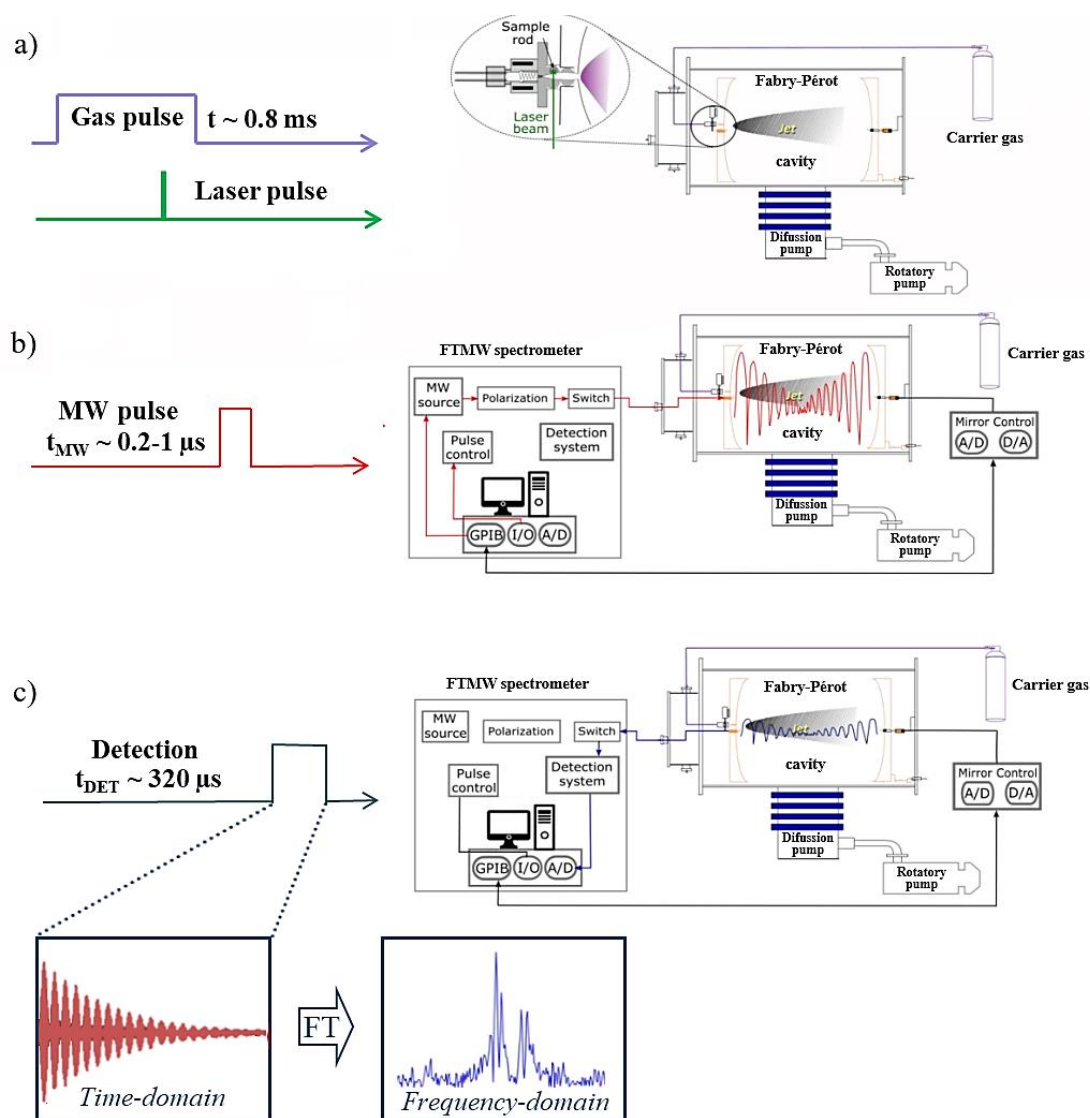
In brief, both techniques have been extensively exploited in our laboratory at the University of Valladolid, emerging as very suitable approaches for the study of the molecular systems proposed in [Chapter II](#). In the following sections, a description of the LA-FTMW spectrometers employed during the course of the present thesis will be presented. A complete description of each experimental setup is given elsewhere (Lesarri et al. 2003; Mata et al. 2012, Bermúdez et al. 2014).

### *1.1.a Laser ablation narrowband (cavity based) Fourier transform microwave spectrometers (LA-MB-FTMW)*

As it is shown in Figure 1, the LA-MB-FTMW instrument consists of three main blocks or sections: a laser ablation source, a Fabry-Pérot resonator, and a FTMW spectrometer.

Regarding the laser ablation process, it will be necessary to couple the ablation system with the Fabry-Pérot cavity. Thus, the backside of one of the mirrors of the resonator will be modified so that it can adapt a nozzle with a pulsed valve specifically designed to support ablation (see Figure 2). The samples will consist of solid rods, the preparation of which will be explained in detail in Section 2. The solid rod is placed vertically in the homemade nozzle, and subsequently a focused laser beam will produce the vaporization of the sample. What is more, a homemade motored device will be employed to constantly rotate and translate the rod, looking forward to enhancing the reproducibility of the ablation process and optimize the use of sample. To carry out the experiments reported in this Doctoral Thesis, the third (355 nm) harmonic of a Nd:YAG picosecond laser (*Q*-switched, ca. 200 ps, ~ 20 mJ / pulse, which changes depending on the chemical system under study) was employed.

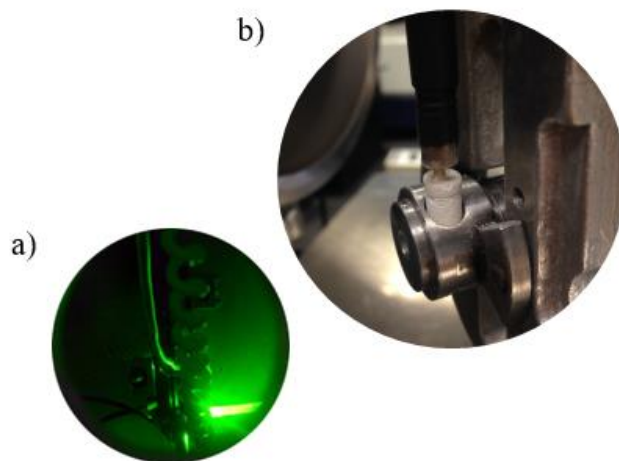
Generally, an experimental cycle starts with a pulse of a carrier gas (backing pressures between 1-15 bars of a noble gas, such as Ne, Ar or He) of typically approximately 0.8 ms. After a short delay, which is a crucial parameter to optimize the signal, a short laser pulse removes matter from the solid sample rod and drag it into the plasma plume (see Figure 1 and 2). In addition, a meticulous control of the pulse energy will be required to manage plausible photofragmentation issues. Afterward, the laser ablated molecules will be seeded in the carrier gas and the gaseous mixture is adiabatically expanded (under an isentropic approximation) through the nozzle orifice into the vacuum chamber of the spectrometer ( $\sim 10^{-7}$  mbar) (Levy et al. 1977; Levy 1981; Hayers & Small, 1983) [see Figure 2(a)].



**Figure 1.** Schematic drawing of a complete LA-MB-FTMW experimental cycle [Figure adapted from Blanco et al. 2007 and Alonso 2018].

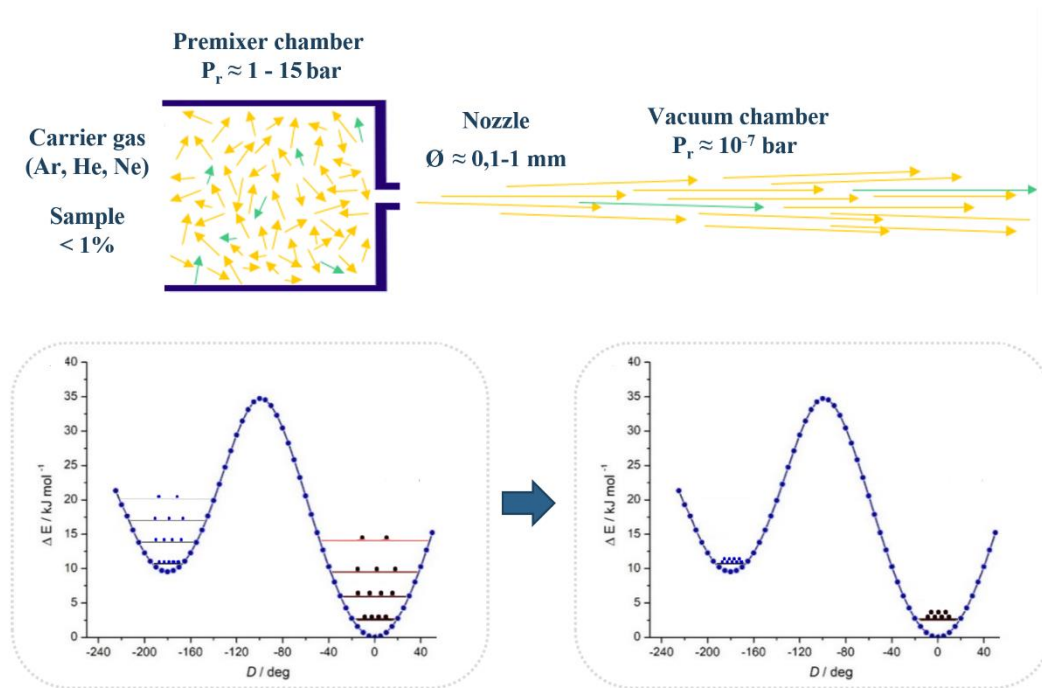
During this process, multiple molecular collisions will occur in the reservoir prior expansion (near the exit channel of the nozzle) that will induce a significant acceleration of the molecules as well as a highly directional gas flux near the exit channel of the nozzle. Thus, the generated jet will present a very narrow velocity distribution. The seeded molecules will suffer an effective cooling of the rotational and vibrational degrees of freedom and, therefore, the distinct conformations will freeze into the ground vibrational state of the corresponding potential energy surface (PES) (see Figure 3). In case the conformational interconversion barriers are high enough, the conformer distribution before the molecular beam occurs will be preserved.





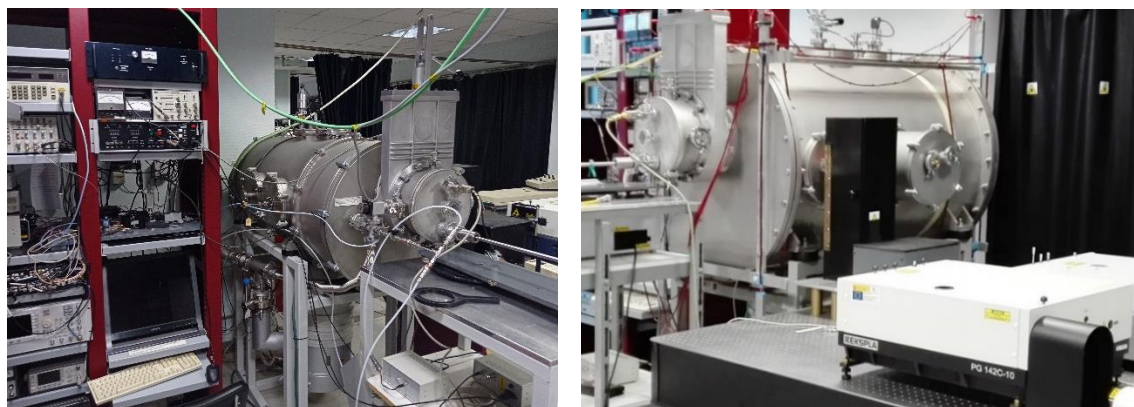
**Figure 2.** a) Laser ablation plume; b) Solid rod placed vertically in a homemade nozzle.

Subsequently, a short microwave pulse (between 0.2 and 1  $\mu\text{s}$ ) is applied to induce a macroscopic polarization of the molecular species populating the jet (see Figure 3b). As seen in Figure 3(c), once the excitation pulse ends and after an adequate protection width, the molecular emission signal or free induction decay (FID) will be recorded in the time domain. Very recently, León et al. (2017) have implemented a new pulse sequence that allow us to collect multiple FIDs to improve the sensitivity of the spectrometer. Afterward, the spectrum in the time domain will be Fourier transformed to the frequency domain, where each of the transitions appear as Doppler doublets due to the coaxial disposition of the supersonic jet and the resonator axis (microwave radiation). The final rest frequency will be computed as the arithmetic mean of the Doppler doublets, reaching an experimental precision of about 10 kHz (theoretical full width at the half maximum, FWHM) and an accuracy or experimental uncertainty better than 5 kHz. Finally, once the vacuum chamber has been evacuated, a new experimental cycle may start (usually a 2 Hz repetition rate is used). What is more, the number of experimental cycles needed to record a rotational signal will depend on the corresponding intensity of the line. Then, hundreds or even thousands of experimental cycles may be added coherently for very weak signals.



**Figure 3.** Schematic representation of the supersonic expansion (or molecular beam) and the vibrational and rotational cooling implied in the process for a fictitious conformational space.

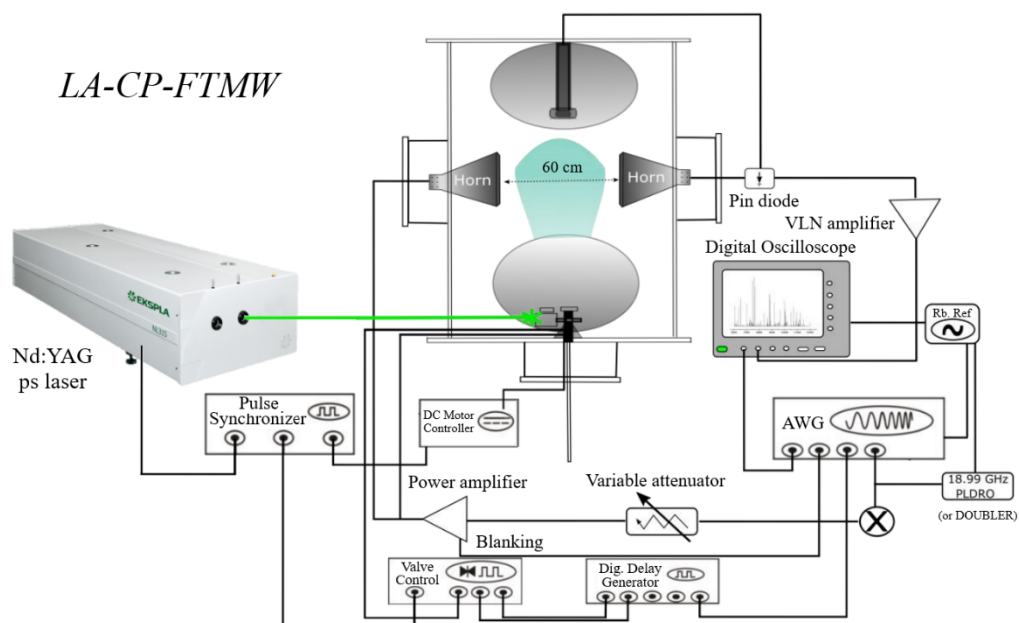
In this thesis, two LA-MB-FTMW spectrometers that cover different frequency ranges have been employed, together with an old MB-FTMW with no laser ablation system which works in the 6-18 GHz frequency range (Alonso et al. 1997). For middle-size molecular systems a 6-18 GHz system has been employed. In addition, a recently build instrument operating in the 2-8 GHz frequency range (Bermúdez et al. 2014), shown in Figure 4, has been developed to the study of large molecules (see Chapter VII).



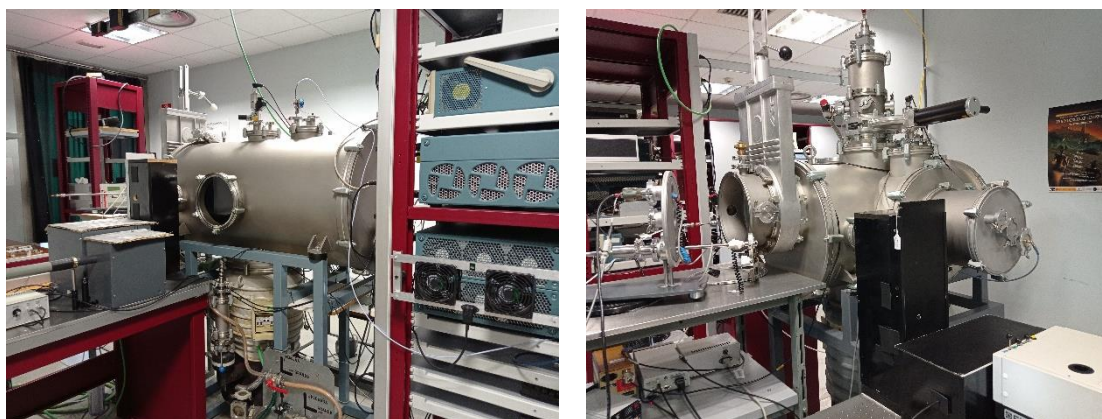
**Figure 4.** Current status of the 6-18 GHz (left) and 2-8 GHz (right) LA-MB-FTMW instruments.

1.1.b Laser ablation broadband chirped pulse Fourier transform (LA-CP-FTMW) spectrometers

Besides the use of our LA-MB-FTMW instruments, in recent years several CP-FTMW spectrometers, with and without a laser ablation source, have been built in our laboratory. Of great relevance is the tank devoted to the 6-18 GHz bandwidth, which allows two different configurations: we can work with a parabolic set (2-32 GHz), located in a coaxial arrangement with respect to the molecular beam (described in detail in Peña et al. 2013), or we can operate with horns (6-18 GHz), situated perpendicular to the jet (a scheme of the experimental layout is depicted in Figure 5). The second one covers the 2-8 GHz frequency range and operates exclusively with a set of horns (see Figure 6). For both instruments a Nd:YAG picosecond laser is employed, and the injection system is referable to that of the narrowband spectrometers.

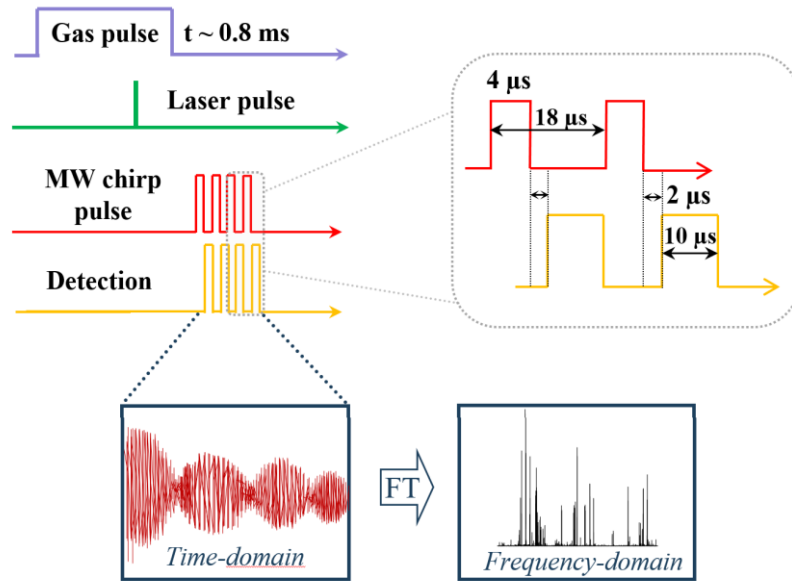


**Figure 5.** Scheme of the LA-CP-FTMW spectrometer operating in the 6-18 GHz (horns) or 2-36 GHz (parabolic set) frequency ranges, respectively. [Figure adapted from Alonso 2018]



**Figure 6.** Picture of the 6-18 GHz (left) 2-8 GHz (right) LA-CP-FTMW spectrometers.

A schematic diagram of a complete LA-CP-FTMW experimental cycle can be described as follows (see Figure 7). First, a digital delay generator sends a pulse to the valve driver, opening the next channel of the nozzle for approximately 0.8 ms (molecular pulse). Hence, the carrier gas (Ar, He or Ne) adiabatically expands into the vacuum chamber generating the supersonic jet, already described. After an adequate delay, a picosecond laser pulse of a Nd:YAG laser (3<sup>rd</sup> harmonic) incides in the rod and vaporizes the solid sample at a repetition rate of 2 Hz. In this case, for each injection of gas, an arbitrary waveform generator (AWG) is used to generate four individual fast chirped microwave pulses (of 4  $\mu$ s width) that macroscopically polarize the molecules present in the jet. Depending on the frequency range of the broadband excitation pulse, a 300 W traveling wave tube amplifier (TWT) or a 200 W solid state amplifier is used to amplify each chirped pulse. Then, 2  $\mu$ s after the excitation pulse ceases, the corresponding FID is acquired for 10  $\mu$ s by a digital oscilloscope and Fourier transformed to the frequency domain. In this case, line widths (FWHM) of about 100 kHz are obtained. Also, a 10 MHz rubidium frequency standard oscillator is employed to achieve the phase reproducibility of the experiment. Hence, over 100.000 individual FIDs are usually averaged in the time domain and Fourier transformed to obtain the final broadband spectra.



**Figure 7.** A complete LA-CP-FTMW experimental cycle (with multifid).

### 1.1.c Double resonance (DR) configuration

One of the biggest drawbacks in the abovementioned techniques is the limitation in frequency of the conventional setups (up to 18 GHz for the MB-FTMW spectrometers). During this Ph.D. different double resonance (DR) configurations up to 40 GHz have been developed seeking for an extension of our narrowband instruments. (DR-MB-FTMW y DR-LA-MB-FTMW). This fact is of paramount importance since it serves as a straightforward connection between the microwave and the millimeter-wave data, guiding the analysis of the high-frequency spectra. Furthermore, it allows us to measure in the same frequency region covered by astronomical surveys conducted at centimeter wavelengths, such as those performed with the GBT and the Yebes 40-m radiotelescope.

In this DR technique (Mori et al. 2009), we monitor the FID or multi-FID signal of a known microwave transition by the FTMW system while centimeter-wave radiation is sent into the Fabry-Pérot chamber (using an additional synthesizer) perpendicular to the directions of the “low frequency” microwave radiation and the supersonic jet. Then, we scan the frequency of the centimeter-wave radiation; when it is resonant with a particular transition, i.e., connecting the lower or upper level of the monitored microwave transition, the monitored microwave transition's coherence is broken, and the FID signal decreases. As a result, transitions induced by the centimeter-wave radiation, even with

the hyperfine structure (HPS) resolved (see Section 3.3), can be observed as a change in the intensity of the monitored transition using an active background subtraction method.

In Figure 8, we present a simplified diagram of our double resonance configurations, which operate in two different frequency intervals, 18-26 GHz and 26-40 GHz, respectively, which overlap with the range covered by several observational facilities.

Using this technique, several molecular systems have been studied by DR-MB-FTMW or LA-DR-MB-FTMW during this thesis: **aceto hydroxamic acid**, **glycolamide** (Chapter IV) and **cianoacetic acid** (Chapter V); providing very accurate laboratory rotational data that further enable their eventual search in the ISM using low frequency Surveys.

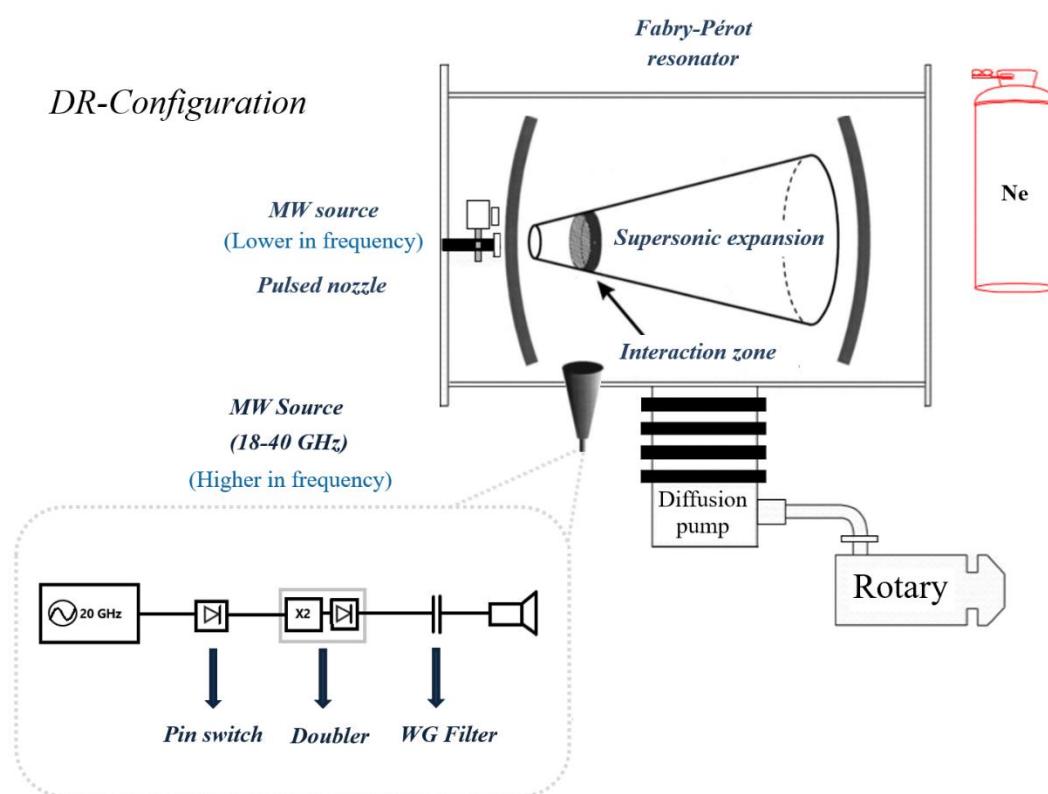


Figure 8. a) Schematic representation of the newly built double resonance setup.

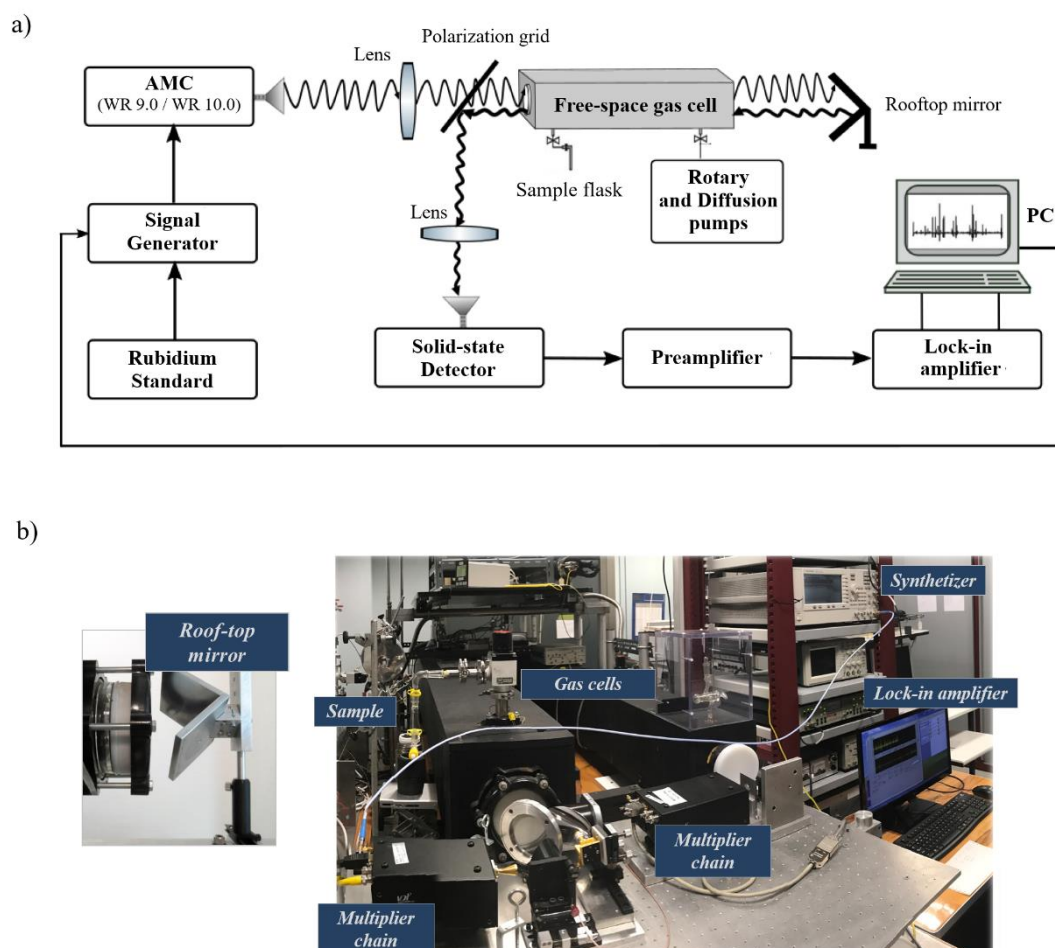
## 1.2 Frequency modulated (FM) millimeter and submillimeter-wave spectroscopy

Frequency domain spectroscopy is considered a very useful and suited technique to study astrophysically relevant molecules, that is, the investigation of molecular systems that are candidates to be detected in the ISM. The aim of this technique is to measure the laboratory rotational spectra of these species in the same operation range of some of the latest observational facilities (IRAM, ALMA... etc.), to enable a direct comparison between highly accurate experimental rotational data and millimeter or submillimeter-wave astronomical Surveys.

In contrast to the time domain spectroscopy, in this technique we usually perform the experiment at room temperature (i.e., without a supersonic jet). Thus, the rotational spectra of the excited vibrational states of the molecules can be measured. Also, we modulate or vary the frequency with a radiation source and simultaneously measure the absorption spectrum of the system under study. At the University of Valladolid, a millimeter and submillimeter-wave spectrometer that cover from 50 to 1080 GHz is available. Two different configurations have been constructed: a single pass and a double pass configuration. In this thesis, only the double pass configuration has been employed, which is described in detail as follows (for detailed information see Daly et al. 2014, Alonso et al 2016):

As shown in Figure 9, this double pass spectrometer is based on a cascade multiplication of the output of an Agilent E8257D microwave synthesizer, which operates up to 20 GHz, by a set of active and passive multipliers. Depending on the desired frequency range, different amplifier-multiplier chains from Virginia Diods (WR10.0 and WR9.0) will be employed. Also, to cover the whole spectral range, additional frequency doublers and triplers will be used (WR2.2, WR4.3, WR2.8, VDI, Inc.). The output of the synthesizer is frequency modulated ( $f = 10.2$  kHz), using a modulation depth between 20 and 40 kHz. Then, the optical path length of 3.6m long is doubled using a polarization grid (Millitech, Inc.) and a homemade rooftop mirror. After the second pass through the cell, we detect the signal using the corresponding solid-state zero-bias detector and is sent to a lock-in amplifier (SR510, Stanford Research Systems, Inc.), where a  $2f$  demodulation (time constant of typically 30 ms) is applied to enhance the sensitivity of the measurements. Both the synthesizer and the lock-in are connected through GPIB to a computer and controlled by a LabVIEW program, developed in our laboratory. Hence,

the rotational spectra are measured as the average of two scans: one recorded in increasing frequency and another in decreasing frequency. Finally, we use the Assignment and Analysis of Broadband Spectra (AABS) package (Kisiel et al. 2005) to adjust the rotational lines to a Gaussian profile. The experimental uncertainty of the isolated symmetric lines is estimated to be around 30 kHz.

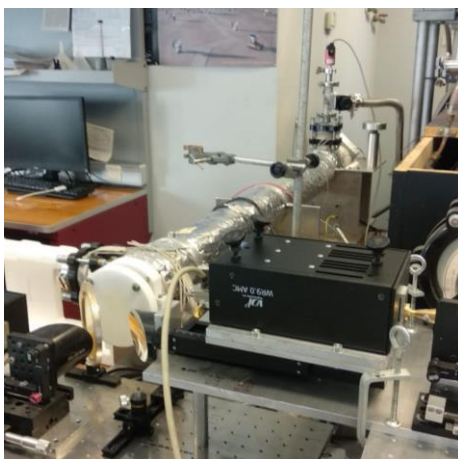


**Figure 9.** a) Scheme of the double pass configuration; b) Current state of the mm- and submm-wave spectrometer in our laboratory.

However, some of the suggested molecular systems (i.e., **glycolamide**) are solids with very low vapor pressures at room temperature, which can be associated with condensation problems in the walls of the cell, therefore preventing their transfer to the gas phase (if it is even possible). To overcome this issue, a new free-space cell has been recently build in our laboratory (see Figure 10). It consists of a 140 cm-long Pyrex cell with an inner diameter of 10 cm and Teflon Windows covered by an aluminium heating



blanket. To monitor and adjust the temperature, we use a heating tape controlled by a thermocouple. It is worth noticing that the sample is continuously flowed through the cell (optimum gas pressure between 10 y 30  $\mu$ bar) to minimize plausible condensation problems as well as to constantly provide fresh sample. In this case, the experimental uncertainty can be a bit higher than the usual 30 kHz (of about 50 kHz) due to the corresponding temperature broadening.



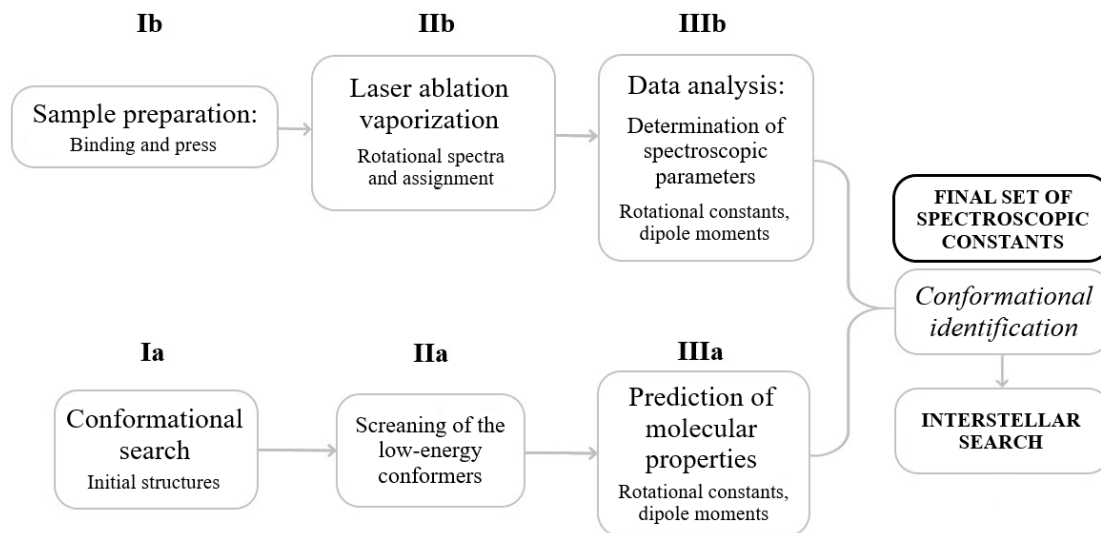
**Figure 10.** Picture of the newly built heating cell.

## 2. GENERAL WORK STRATEGY

The general methodology to study and identify the molecular systems presented in this thesis is schematically shown in Figure 11. It can be divided in two parallel blocks or sequence lines: a preliminar theoretical and computational study (block a) and the complete experimental procedure (block b)

In a first step (**Ia**) molecular mechanics methods are used to systematically explore the conformational space of each molecular system. All the conformations resulted from this search are submitted to density functional theory (DFT) computations. Then, during step **IIa**, we select the most stable conformations. Afterward, each conformation is submitted to higher level computations (see Section 3 for additional details) to predict the relevant spectroscopic parameters that allows us to interpret a rotational spectrum (step **IIIa**), such as the rotational constants, which strongly depend on the mass distribution of

the molecule, the electric dipole moment components and the nuclear quadrupole coupling constants, if needed.



**Figure 11.** Block diagram of the general strategy.

In a parallel way, we perform the experiment (block b). Most of the species presented in this thesis are solid compounds with very high melting points at room temperature, which need to be laser ablated in order to avoid decomposition problems during the vaporization process. In this case, samples are prepared grounding approximately 1g of the solid, which is subsequently mixed with few drops of a commercial copolymeric binder (step **Ib**). Then, the mixture is pressed at 200 MPa in a hydraulic press till compact solid rods are obtained. Finally, these rods are placed in a sealable vessel (vacuum desiccator), where a drying agent is used to evacuate the air and moisture and perfectly dry them.

Once the samples are placed in the corresponding nozzle (**IIb**), the rotational spectra is recorded using LA-CP-FTMW and LA-MB-FTMW (or CP-FTMW/MB-FTMW if the molecule can be conventionally heated, i.e., **glycolamide**). In this step, the instrumental conditions (stagnation pressure, power, timing, laser ablation...etc..) are optimized. Generally, a broadband spectrum is acquired at first, using once of the LA-CP-FTMW spectrometer. Then, we search for characteristic patterns, which depend on the shape of the molecule and rotational selection rules (see [Section 4](#)), to assign each of

the rotamers present in the jet (step **IIb**). On the other hand, long scans around the predicted frequency lines of the lowest energy conformers are carried out to perform the conformational search using the narrowband (MB-FTMW and LA-MB-FTMW) techniques. When a transition is found, an iterative line-by-line measurement and fitting procedure is executed until the corresponding conformer is conclusively identified.

In the case of molecules of astrochemical relevance, the aim of the investigation is not only to unveil the conformational panorama, but also to obtain rotational parameters accurate enough to perform a conclusive astronomical search. Thus, it is usually mandatory to perform an extension of the measurements up to higher frequencies.

Once the microwave spectra are registered and assigned, we use the preliminary first set of spectroscopic constants from the microwave study to predict and model the rotational spectra in the millimeter-wave region. Then, we will try to record large records of hundreds of GHz progressively increasing in frequency using different configurations of the FM modulated spectrometer, and ultimately acquiring, if possible, thousand of transitions. Afterwards, the analysis of the spectrum is carried out using the Assignment and Analysis of Broadband Spectra (AABS) package (Kisiel et al. 2005). Also, Loomis-Wood type graphic representations will be used to simplify the assignments. Finally, the final set of constants will be shared with the astrophysical community to enable an astronomical search for each system using the abovementioned Line Surveys.

### 3. COMPUTATIONAL METHODS

#### 3.1. *Introduction*

Computational simulation methods are a powerful and widely used tool in numerous fields of scientific research. This is due to its usefulness for analyzing data, making predictions on different models, and solving problems numerically. The field of Interstellar Chemistry is a clear example for the application of theoretical methods, since it allows the characterization and analysis of chemical-physical properties for different systems where, sometimes, a laboratory-based experimentation is very hampered.

Computational Chemistry bases its methods on the application of a set of algorithms and approximations from Classical Physics and, fundamentally, from

Quantum Mechanics, to specific chemical problems, which ultimately allows us to resolve the Quantum Chemistry equations to describe microscopic systems.

### 3.1.a. Schrödinger equation and wave function

The Schrödinger equation (1) is one of the fundamental pillars of Quantum Mechanics, whose resolution is usually the main target for any theoretical-computational method.

$$H\Psi(r) = E\Psi(r) \quad (1)$$

Where  $\Psi(r)$  represents the wave function,  $E$  is the energy of the system and  $H$  is the Hamiltonian operator, which consists of a first term corresponding to the kinetic energy of the system and a second term attributed to the potential energy:

The knowledge of the wave function is mandatory to describe the state of a system. These states are generally time-dependent, so the wave function is also a function of time  $\Psi(r, t)$ , although for most chemical applications, a time-independent treatment is used instead. In the latter case, the states are known as stationary states.

In the case of a real polyelectronic system, the Hamiltonian operator is described by the following expression:

$$\begin{aligned} H &= T_{elec} + T_{nuc} + V_{ne} + V_{ee} + V_{nn} = \\ &= - \sum_i^{elec} \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_\alpha^{nuc} \frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 \\ &+ \sum_i^{elec} \sum_\alpha^{nuc} \frac{z_\alpha e^2}{4\pi\epsilon_0 r_{i\alpha}} + \sum_i^{elec} \sum_{j>i}^{elec} \frac{e^2}{4\pi\epsilon_0 r_{ij}} + \sum_\alpha^{nuc} \sum_{\beta>\alpha}^{nuc} \frac{z_\alpha z_\beta e^2}{4\pi\epsilon_0 r_{\alpha\beta}} \end{aligned} \quad (2)$$

where  $T_{elec}$  and  $T_{nuc}$  are the kinetic energy operators for the electrons and nuclei, of the system, respectively, and  $V_{ne}$ ,  $V_{ee}$  and  $V_{nn}$  represent the potential energy operators for nucleus – electron attraction, electron –electron repulsion, and nucleus – nucleus repulsion, respectively. Through the resolution of the Schrödinger equation, the energy

of the system will be obtained, which is a crucial parameter in the study of any molecular system. However, only for the simplest cases, such as the hydrogen atom and hydrogen-like or hydrogenic atoms, this equation presents exact solutions. Nevertheless, the description of average-size systems will require physical-mathematical approximations, highlighting the Born-Oppenheimer approximation.

### 3.1.b. *Born-Oppenheimer approximation*

Among all plausible approaches, the Born-Oppenheimer approximation is crucial for the quantum mechanical treatment of any molecule. It is based on the separability of electronic and nuclear movements, given the great mass difference between both particles ( $m_{nucleus} > 1600 m_e$  for the hydrogen atom) and is related to a much faster movement of the electrons compared to the nuclei. Therefore, it can be considered that the position of the nuclei does not vary during the movement of the electrons, making it possible to subtract the nuclear terms of the Hamiltonian ( $T_n$  y  $V_{nn}$ ) and significantly simplify the Schrödinger equation.

$$\hat{H} = \hat{H}_e(q_e, q_n) + \hat{H}_n(q_n) \quad (3a)$$

$$\Psi(q_e, q_n) = \psi(q_e, q_n) + \Psi(q_n) \quad (3b)$$

where  $q_n$  and  $q_e$  are the electronic and nuclear coordinates, respectively.

Thus, the introduction of the Hamiltonian in the Schrödinger equation for the wave function of the system results in the obtention of the electronic energy, which will depend on the geometric parameters of the molecule. If a diatomic molecule is considered, its only variable geometric parameter is the bond distance, where a two-dimensional  $E(r)$  function is generated. By increasing the number of parameters to two, a surface will be obtained, giving rise to the concept or property of Potential Energy Surface (PES). Finally, in the case of a general molecule of  $N$  atoms, we will have  $3N-6$  degrees of freedom ( $3N-5$  for linear molecules) or needed parameters to achieve a complete description of a system, obtaining in most cases Potential Energy Hyper Surfaces. The mathematical analysis and understanding of the PES enable to identify local minima,

global minima (most stable geometries), reaction paths ... etc., being a fundamental tool for all computational processes.

On the other hand, the study of a reaction path implies a comprehensive and detailed analysis of the PES. A system will go through a series of energy minima as well as some energy maxima, where there is an exchange in the position of the nuclei in the three-dimensional space, thus varying the geometry between two structures of minimum energy. The PES maxima are first-order saddle points (maximum in one direction and minimum in the rest) and are known as transition states (TS). Also, they will be of great interest to understanding the mechanisms involved in the different chemical processes. Currently, there are many different methods to search for transition states. Usually, a potential energy gradient is used and, taking a starting structure, a second structure where the energy gradient is zero is reached. Although the energy minima can be verified by verifying that all their vibrational frequencies are real, the transition states have vibration frequencies with an imaginary component. In general, a saddle point of order  $N$  has  $N$  imaginary frequencies in the transition state. TS are usually located at first-order saddle points; thus, we can represent the reaction path in a plane, which is known as the potential energy curve.

Both Gaussian09 (Gaussian09 Revision B.01) and Gaussian16 (Gaussian 16 Revision A.03), widely well-known chemical software packages, have been used during this thesis. The enable the use of the Quadratic Synchronous Transit (QST) method, implemented by the group of H. B. Schlegel (Peng et al. 1993, 1996) to locate the transition structures. The QST2 and QST3 options stand out above the rest, where QST2 requires the specification of the structure of two molecules (reactant and product, or two intermediates) in the input, while QST3 implies also including the structure for the transition state. This method behaves in a similar way to the algorithm used for minimum optimizations and converges efficiently when the proposed starting structures and the Hessian estimation are adequate. Finally, we can verify the connections between suggested TS structures and adjacent minima (reactive species and products) using the intrinsic reaction coordinate (IRC) technique (Gonzalez & Schlegel 1989, 1990). In this approach, the evolution of the proposed transition state in both directions will be evaluated until the starting species (energy minima) are obtained.

### 3.2. *Theoretical methods*

Theoretical Chemistry develops theories and methods to study the different chemical systems. There are a multitude of theoretical and computational methods, which fundamentally differ in the computational cost, their applicability to different systems and the quality or precision of results. They can be classified into two broad categories:

- **Molecular Mechanics:** In this methodology, molecules are considered using a model of balls (atoms) and springs (bonds), trying to evaluate the mechanical energy involved in stretching bonds, twists, torsions, etc. The different elements will be parameterized by means of a force field from experimental or computational data. It is a very simplified approach, in which no quantum effect is considered, and only classical mechanics are applied. However, this allows the study of large molecular systems and serves as a first approach to common problems such as conformational searches and analysis, where it is necessary to perform a thorough screening for each of the possible structures (in some cases this implies hundreds of geometries).

In this Ph.D., it has been considered the Merck Molecular Force Field (MMFF) (Halgren 1996), which is a force field designed to study a broad range of organic molecules of biochemical and pharmacological interest. In general, a systematic search on the conformational space is carried out for each molecular system, where “Large scales Low Mode” and a Monte Carlo-based search algorithms in MacroModel ([www.schrodinger.com](http://www.schrodinger.com)) are used to locate all possible minima energy structures on the potential hypersurface (within an energy window of 1000-1500 cm<sup>-1</sup>), which are those most likely to be detected in the experiment. Also, the TINKER programme has been used for small and rigid systems (<http://dasher.wustl.edu/tinker/>).

- **Quantum Mechanics:** Contrary to Molecular Mechanics methods, the methods discussed below apply the equations from the Quantum Chemistry to the investigation of molecular systems. This opens the window to the study of its electronic properties and the analysis of chemical processes in which electronic changes occur, such as reactions, where bonds are constantly created or broken. They are considerably more powerful than Molecular Mechanics methods; however, they entail a substantial increase in the computational cost of calculations, further limiting the size of the

candidate systems to be studied. They can be divided into three main categories: semi-empirical methods, methods based on the Density Functional Theory and *ab initio* methods. The methods used in the calculations of the present thesis are presented below.

### 3.2.a. Density Functional Theory

The Density Functional Theory (DFT) focuses its attention on the electron density  $\rho$ . While the wave function  $\Psi$  of  $n$  electrons that depends on  $4N$  coordinates (the three spatial coordinates and one spin coordinate), the main strategy is to replace it by  $\rho$ , which depends only on three coordinates. It is based on the Hohenberg-Kohn Theorems (Hohenberg 1964), which state that the energy of the ground state of a system is an electron density functional,  $E = E(\rho)$ . However, although this correlation is known, the exact mathematical formulation of this functional is hidden. In general, these methods enable to reproduce some of the main structural properties, energies, and spectroscopic parameters (rotational constants, vibration frequencies, dipole moments, etc.) of different molecular systems with a significantly small computational cost.

The application of the Density Functional Theory requires the selection of an adequate functional. One of the most popular functionals is the B3LYP exchange-correlation functional, which includes the Becke's three-parameter exchange functional (Becke 1993) and the Lee-Yang-Parr (LYP) correlation functional (Lee et al. 1988):

$$B3LYP = (1 - A)E_X^{LSDA} + AE_X^0 + B\Delta E_X^{B88} + (1 - C)E_C^{VWN} + CE_C^{LYP} \quad (4)$$

where  $A = 0.2$ ,  $B = 0.72$  and  $C = 0.81$ .

To better understand this functional, we shall go back to the development of the so-called *half-and-half* functional (Becke 1993), considered as the "father" of hybrid functionals, which tries to model the electron-electron interaction energy,  $W$ , using a linear equation  $W_\lambda = a + b\lambda$ , where  $a$  and  $b$  are fitted to obtain  $W_0$  (Becke approximates  $W_0 = a = E_X^0$ , where  $E_X^0$  is the exchange energy of a noninteracting electron system) and  $W_1$  ( $W_1 = a + b = E_{XC}^{LSDA}$ ). The *half-and-half* functional can be defined as follows:

$$E_{XC}[\rho] = a + \frac{b}{2} = \frac{1}{2}E_X^0 + \frac{1}{2}E_{XC}^{LSDA} \quad (5)$$



In this equation, the electron density is treated under the *Local Spin Density Approximation* (LSDA), which applies the so-called LDA functional ( $E_{XC}^{LDA}[\rho(\vec{r})] = \int \rho(\vec{r}) \epsilon_{XC}^{LDA} \rho(\vec{r}) d\vec{r}$ ).

On the other hand, regarding the LYP correlation functional, it was developed based on the correlation energy of the helium atom and it is finally combined with the B88X exchange functional, introduced by Becke under the *General Gradient Approximation* (GGA), which introduces a non-homogeneous  $\rho$  to take into account the density gradient instead of only the electron density at a located position, such as in the LDA. The exchange functional is defined as follows:

$$E_X^{GGA}[\rho(\vec{r})] = \int \rho^{\frac{4}{3}}(\vec{r}) f(x) d\vec{r} \quad (6)$$

Here the gradient expansion,  $f$ , introduced by Becke (B88X) is written as:

$$f(x) = C_X + \beta \left( \frac{x^2}{1 + 6\beta x \operatorname{arcsinh}(x)} \right) \quad (7)$$

where  $\beta$  is calculated by a fit to atomic exchange energies of several noble gases. Thus, this combination led to the well-known B3LYP functional, employed in the abovementioned equation (4).

Hence, the structures of all the species studied in this Ph.D. (reactants, intermediates, transition states, and products) are first optimized in the framework of the density functional theory (DFT) employing this functional, because of the good quality of the predicted structural and spectroscopic parameters linked to a very low computational cost (Koch 2001). In addition, Grimme's D3 dispersion interactions are included (Grimme et al. 2010), which significantly enhance the description of the dispersion forces (long-range interactions).

Afterward, walking the thin line between accuracy and computational cost, we find the double hybrid functionals. These functionals sits on the seamline between wavefunction and DFT methods, since they add non-local electron correlation effects to a standard hybrid functional by second-order perturbation theory (Grimme 2006). Schwabe & Grimme (2007) added an empirical dispersion term (DFT-D) to the energy expression and left the electronic part of the functional untouched, in order to correctly

estimate long-range dispersion interactions. In general, this approach offers a similar quality for many chemical-physical properties as high-level coupled-cluster treatments (described below). We have used double-hybrid B2PLYPD3 methods to compute geometry optimizations for all the previously calculated B3LYP structures from the conformational screening.

Also, at the respective B2PLYPD3 optimized geometries, harmonic vibrational frequencies have been calculated to apply zero-point-energy corrections to the electronic energy. In some cases, to aid in a possible experimental identification by IR spectroscopy, we have computed anharmonic vibrational frequencies at the B2PLYPD3 or MP2 level (see below) of theory using the second-order perturbation treatment (VPT2) (Hoy et al. 1972, Barone et al. 2007). The treatment includes a full cubic force field (CFF) and semi-diagonal quartic force constants. In addition, vibration–rotation interaction constants can also be evaluated, from the CFF calculations, allowing for correction of rotational constants, including vibrational effects. For this purpose, the CFOUR (Stanton et al. 2013) package was employed.

### 3.2.b. *Ab initio methods*

These methods involve a mathematical treatment of the systems relying exclusively on quantum mechanics, thus solving the Schrödinger equation without the assumption of any external parameter. They can provide very precise information about systems in which quantum effects are relevant. Their main drawback is that it is not feasible to apply them to large sets of particles, given the enormous complexity and computational cost that these would reach (Levine 1991). Sometimes approximations are used within the quantum treatment to be able to compute larger systems. Examples of such approximations are the Born-Oppenheimer (Oppenheimer 1927) discussed above and the Hartree-Fock treatment (Fock 1930, Slater 1930).

An effective Hamiltonian is proposed, and the integrals are evaluated using exclusively the values of universal constants, without using any type of experimental data. We can formulate the following classification:

- Hartree-Fock (HF) or Self Consistent field (SCF) methods: The HF method is the most common starting point for most of the electronic structure methods. It considers that the electrons do not interact among them and, therefore, the interelectronic repulsion is considered as an average.
- Post-HF methods: they include the electronic correlation using variational (CI, MCSCF), perturbative (MP) or coupled cluster (CC) approaches.

The Moller-Plesset (MP) perturbative post-HF methods (Moller & Plesset 1934; Szabo et al 1996) have been extensively used in this Ph.D. (highlighting the second-order perturbative method, MP2). However, there are cases where a higher precision is required than that obtained using MP2 a perturbative method. The solution lies in the use of the Coupled Cluster (CC) methods (Purvis & Barlett 1982; Raghavachari et al 1989). These methods essentially take the basic Hartree-Fock molecular orbital method and uses an exponential *cluster operator* to construct multi-electron wavefunctions that account for the electron correlation. The cluster operator ( $T$ ) can be defined as the sum of single ( $T_1$ ), double ( $T_2$ ), triple ( $T_3$ ) ...etc., excitation operators as follows:

$$T = T_1 + T_2 + T_3 + \dots \quad (8)$$

$T_i$  is described by the following expression:

$$T_1 = \sum_{a,r} t_a^r a_a^\dagger a_a \quad (9a)$$

$$T_2 = \sum_{a>b} \sum_{r>s} t_{ab}^{rs} a_s^\dagger a_b a_r^\dagger a_a \quad (9b)$$

where  $t$  describes the cluster amplitudes.

Moreover, in the CC theory the wave function is an exponential function, in which the exponential part can be expanded in a Taylor series:

$$\begin{aligned} \Psi_{CC} = \exp(T) \Psi_{HF}; \quad \exp(T) = 1 + T_1 + T_2 + T_3 + \dots \\ + \frac{1}{2!} T_1^2 + T_1 T_2 + \frac{1}{3!} T_1^3 + \frac{1}{4!} T_1^4 + \frac{1}{2!} T_1^2 T_2 + T_1 T_3 + \dots, \end{aligned} \quad (10)$$

Thus, the  $\Psi_{CC}$  equation can be rewritten as:

$$\Psi_{CC} = \prod (1 + T_1 + T_2 + T_3 + \dots) \Psi_{HF} \quad (11)$$

However, the evaluation of the complete CC wave function, which satisfies the Schrödinger equation  $\hat{H} \exp(T) |\Psi_{HF}\rangle = E \exp(T) |\Psi_{HF}\rangle$ , is not computationally affordable, and it is therefore necessary to truncate T. The resulting  $\Psi_{CC}$  should include double excitations to recover the electron correlation. For instance, when  $T = T_1 + T_2$ , the CC method is named as CCSD.

Hence,  $\Psi_{CCSD}$  can be written as:

$$\Psi_{CCSD} = \exp(T_1 + T_2) \Psi_{HF} = (1 + T_1 + T_2 + \frac{1}{2} (T_1 + T_2)^2 + \dots) \Psi_{HF} \quad (12)$$

where several terms, the single ( $t_a^r$ ) and double ( $t_{ab}^{rs}$ ) cluster amplitudes are unknown. Then, the so-called *unlinked* or *projected coupled cluster equations* should be used in order to compute the correlation energy ( $E_{corr}$ ) and T, employing projections over  $|\Psi_{HF}\rangle$  and excited determinants over  $|\Psi_{HF}\rangle$  as follows:

$$E_{corr} = \langle \Psi_{HF} | \hat{W} \exp(T_1 + T_2) | \Psi_{HF} \rangle \quad (13a)$$

$$E_{corr} = \langle \Psi_{HF} | \hat{W} \exp\left(1 + T_2 + \frac{1}{2} T_1^2\right) | \Psi_{HF} \rangle \quad (13b)$$

Here  $\hat{W} = \hat{H} - E_0$  and the energy exclusively depends on single and double amplitudes, disregarding the T truncation. Regarding the cluster amplitudes, exemplified for the single cluster amplitude:

$$\langle \Psi_a^r | \hat{W} \exp(T_1 + T_2) | \Psi_{HF} \rangle = E_{corr} \langle \Psi_a^r | \exp(T_1 + T_2) | \Psi_{HF} \rangle \quad (14a)$$

$$\langle \Psi_a^r | \hat{W} \exp\left(1 + T_2 + \frac{1}{2} T_1^2 + T_1 T_2 + \frac{1}{3!} T_1^3\right) | \Psi_{HF} \rangle = t_a^r E_{corr} \quad (14b)$$

where  $t_a^r$  only depends on single, double and triple excitations and are finally optimized using an iterative method.

Finally, in case an even higher quality method is required, triple excitations can be introduced using perturbation theory [CCSD(T)]. Nevertheless, its high CPU time scales with system size ( $N^7$  for CCSD(T) Watts et al. 1993), and its applicability is usually limited to small molecules. Although most of the molecular systems under study contain more than ten atoms (relatively large molecules), in this thesis we have carried out coupled-cluster calculations, i.e., several geometry optimizations have been performed at the CCSD/cc-pVTZ level (see Chapter VI). Moreover, the CCSD(T) method (Raghavachari et al. 1989) have also been used in conjunction with the aug-cc-pVTZ basis set on the MP2/aug-cc-pVTZ or B2PLYPD3/aug-cc-pVTZ optimized geometries to perform single point computations in order to obtain more accurate energies

### 3.3. Basis sets

The basis set is a compilation of mathematical functions combined to properly describe the molecular orbitals. There two types of basis functions extensively employed in quantum-chemical computations: the Salter Type Orbitals (STO) and the Gaussian Type Orbitals (GTO). Nowadays, the later ones are the most commonly used, since they have an analytical solution for the three- and four-centre two-electron integrals, and are represented as follows:

$$\chi_{\zeta, l_x, l_y, l_z}(x, y, z) = N x^{l_x} y^{l_y} z^{l_z} e^{-\zeta r^2} \quad (15)$$

Nevertheless, these functions offer a poor description of electron-nucleus interactions, being therefore mandatory to use a linear combination of GTO functions, also named Contracted Gaussian Type Orbital (CGTO) functions. Depending on the number of functions that are used, the basis sets can be classified into several types or groups that will be selected depending on the chemical nature of the system under study as well as the required computational time. Starting from the so-called minimum set, or the indispensable set to represent all the orbitals of the molecule, more functions can be added to the base set thus increasing the precision of the calculations.

### 3.3.1. Pople Basis Sets

In 1969, the minimal basis sets ( $\phi_{\text{STO-MG}}$ ) were developed, which exclusively use the minimum number of functions to contain all the electrons of the ground state of each single atom (Hehre 1986)). However, these basis sets are not capable of describing the polarization of the electron density. Thus, the split basis set were introduced, which involve more than a contracted basis function. In particular, the Pople split basis set, which follows the notation X-YZG, separates the orbitals in core orbitals (where X represent the number of primitive Gaussians containing each core atomic orbital basis function, being the orbitals which contribute the most to the electronic energy) and valence orbitals (where Y and Z indicates that they are described with more than one contracted basis functions, and are involved in the reactivity of the system). Also, within the Pople basis set's notation, a hyphen symbol is employed to split the primitives (Gaussian functions) used for the expansion of the core and valence orbitals, respectively.

Moreover, polarization functions (usually functions with high angular momentum) can be added to enhance the description of the polarization of the electron density. For instance, the 6-31G (d, p) implies  $p$  functions for hydrogen atoms and  $d$  functions for heavy ones. In addition, in some cases it is of interest to describe non-covalent or Van der Waals interactions, phenomena that usually occur at long range distances. Thus, to correctly describe these molecular systems, functions including small exponents are needed (wide shape). Within the Pople basis set's nomenclature, diffuse functions are labeled with a "+" symbol.

### 3.3.2. Dunning Basis Sets

The Dunning Basis Set, also known as correlation consistent basis set (ccBS), are optimized for post Hartree-Fock (HF) methods and include single and double excitations. They are named aug-cc-pVXZ, where the prefix aug (*augmented*) shows that the base is augmented with diffuse functions, the acronym cc (*correlation-consistent*) indicates the correlation consistency of the basis set and the *cardinal number*,  $X$ , refers to the inclusion of polarization functions in the valence shell (double, triple, etc ...) Zeta ( $\zeta$ ) (Dunning 1989; Woon & Dunning 1993). This basis set can be systematically enhanced by increasing the size of the basis ( $X$ ), leading to a convergence to the basis set limit. Also, this limit can be reached by using an extrapolation approach (see Chapter VI).

All in all, in this Ph.D. the Pople 6-311++G(d, p) (Ditchfield et al. 1971) and the correlation-consistent polarized valence triple- $\zeta$  Dunning cc-pVTZ and aug-cc-pVTZ basis sets have been mainly used.

### 3.4. *Hardware Resources*

All the computational resources employed in the present thesis belong to the Computational Chemistry Group (CCG) (<http://www3.uva.es/ccg>) and the “Grupo de Espectroscopía Molecular” (GEM) ([www.gem.uva.es](http://www.gem.uva.es)) from the University of Valladolid. A cluster has been arranged for intensive scientific computation, which is made up of 8 Opteron 246 dual core nodes and 8 Xeon E5450 Quad core nodes with two interconnected processors through a Gigabit server. In addition, the group has access to an HP C-110 Workstation (64 Mb RAM and 4 Gb disk) shared with other research groups that is currently obsolete. There is also a 4.5 TB capacity hard drive for storing programs and simulation results, and an uninterruptible power supply. Very recently a new computation server has been acquired in the GEM group with 128 CPUs and 512 Gb of RAM memory.

## 4. THEORETICAL TOOLS FOR ROTATIONAL SPECTROSCOPY

Rotational spectra, like most spectra, can be considered as a “puzzle”. These spectroscopic puzzles are built of different patterns that, once resolved, will lead to very useful experimental data regarding the three-dimensional structure of the molecule. The analysis of a rotational spectrum is an iterative process which implies a fit of the experimental frequencies to the eigenvalues of the Hamiltonian that describes the molecular rotation. The selection of a particular Hamiltonian will rely on different factors, such as the symmetry properties of the molecule, the need to add additional centrifugal distortion parameters for the correct analysis of the spectrum, as well as the presence of splitted lines due to nuclear quadrupole coupling interactions. All these factors will be employed as spectroscopic tools and will be essential to achieve a proper conformational identification. In the following subsections, we will briefly describe each of them (for additional details see (Gordy 1970; Kroto 1992, Hollas 1998).

#### 4.1. Rotational constants

Under the Born-Oppenheimer approximation and considering the Eckart axes system (Eckart 1968) to minimize the rotational-vibrational interaction, the rotational Hamiltonian can be defined as follows:

$$\begin{aligned}
 H_{rot} = T &= \frac{1}{2} \sum_i m_i (\vec{\omega} \wedge \vec{r}_i) (\vec{\omega} \wedge \vec{r}_i) = \\
 &= \frac{1}{2} [\omega_x^2 \sum_i m_i (y_i^2 + z_i^2) + \omega_y^2 \sum_i m_i (x_i^2 + z_i^2) + \omega_z^2 \sum_i m_i (x_i^2 + y_i^2) \\
 &\quad - 2\omega_y \omega_z \sum_i m_i (y_i z_i) - 2\omega_x \omega_z \sum_i m_i (x_i z_i) - 2\omega_x \omega_y \sum_i m_i (x_i y_i)]
 \end{aligned} \tag{16}$$

Taken into account the definition of the moments of inertia and their products, the rotational Hamiltonian can be written as:

$$H_{rot} = \frac{1}{2} [I_{xx}\omega_x^2 + I_{yy}\omega_y^2 + I_{zz}\omega_z^2 + 2I_{xy}\omega_x\omega_y + 2I_{xz}\omega_x\omega_z + 2I_{yz}\omega_y\omega_z] \tag{17}$$

which can be described as the following matrix equation:

$$H_{rot} = \frac{1}{2} (\omega_x \ \omega_y \ \omega_z) \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{xy} & I_{yy} & I_{yz} \\ I_{xz} & I_{yz} & I_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \tag{18}$$

The matrix corresponding to the moments of inertia is symmetric and hermitian, and their components are all of them real (tensor of inertia). Therefore, after the corresponding diagonalization, we obtain:

$$H_{rot} = \frac{1}{2} I_a \omega_a^2 + \frac{1}{2} I_b \omega_b^2 + \frac{1}{2} I_c \omega_c^2 = \frac{P_a^2}{2I_a} + \frac{P_b^2}{2I_b} + \frac{P_c^2}{2I_c} \tag{19}$$

where the angular momentum,  $P = I \omega$ , and the moments of inertia ( $I_a, I_b, I_c$ ) can be interpreted as  $I_\alpha = \sum m_i \cdot r_{i,\alpha}^2$ . The axes system obtained is the so-called system of the *principal axes of inertia* (a, b, c) of the molecule.



Thus, equation (19) will describe the surface of an inertial ellipsoid, with which we are able to classify the molecules attending to the effects of the molecular symmetry on the moments of inertia: linear tops ( $I_a = 0, I_b = I_c$ ), spherical tops ( $I_a = I_b = I_c$ ), oblate symmetric tops ( $I_a = I_b < I_c$ ), prolate symmetric tops ( $I_a < I_b = I_c$ ) or asymmetric tops ( $I_a \neq I_b \neq I_c$ ).

Afterward, going back to equation (19), the rotational Hamiltonian can be rewritten as follows:

$$\hat{H}_R = A \hat{P}_a^2 + B \hat{P}_b^2 + C \hat{P}_c^2 \quad (20)$$

where A, B and C are the rotational constants, which are inversely proportional to the moments of inertia and highly depend on the mass distribution of the molecule, and  $\hat{P}_a^2$ ,  $\hat{P}_b^2$  and  $\hat{P}_c^2$  are the angular momentum operators in the already defined system of the principal axes of inertia of the molecule. For the symmetric top case, ( $\hat{P}_z$  (choosing z as the symmetry axis) and  $\hat{P}^2$  commute with the Hamiltonian, since both of them are constant through the molecular rotation, The resulting eigenvalues are defined as:

$$\text{Prolate top: } E_{J,K} = h [BJ(J + 1) + (A - B)K^2] \quad (21)$$

$$\text{Oblate top: } E_{J,K} = h [BJ(J + 1) + (C - B)K^2]$$

where  $J = 0, 1, 2, \dots$  y  $K = 0, \pm 1, \pm 2, \dots \pm J$ . Therefore, the energy levels are degenerated in K.

Finally, in the case of asymmetric tops, the three principal moments of inertia are different from each other ( $I_a \neq I_b \neq I_c$ ). Consequently, the energy levels can not be described using simple expressions. The quantum number K no longer is a sufficient parameter to describe the energy levels, since it is not constant through the molecular rotation ( $\hat{P}_z$  no longer commutes with the Hamiltonian). To overcome this fact, the solution lies in the use of a combination of the  $K_a$  (o  $K_l$ ) and  $K_c$  (o  $K_l$ ) parameters, which represent the projection of the total angular momentum with respect to the symmetry axis of the molecule in the symmetric prolate and oblate limit situations, respectively.

Moreover, we can estimate the asymmetric degree of a molecules using the Ray constant,  $\kappa$ , which is defined as:

$$\kappa = \frac{(2B - A - C)}{(A - C)} \quad (22)$$

In the limit case,  $\kappa = -1$  represents the prolate symmetric top, while in the oposite case el  $\kappa = +1$  represents the oblate symmetric top. Also, for  $\kappa = 0$  the molecule will present the higher possible assymmetric degree.

#### 4.2. Dipole moments

The electric dipole moment of the molecule – in particular its projection along the molecular axes a, b, c ( $\mu_a, \mu_b, \mu_c$ ) - is closely related to the intensity of the rotational transitions. In order to evaluate the line intensity, we need to analyze the transition moment integral, defined as follows:

$$\langle \psi_1 | \mu_{a,b,c} | \psi_2 \rangle \quad (23)$$

where  $\psi_1$  and  $\psi_2$  are the inital and final states, respectively and, for instance,  $\mu_a$  is the projection of the dipole moment along the  $a$  axis. Hence, if the direct product of the irreducible symmetric representation of the elements of the transition moment integral contains the totally symmetric representation  $[(\psi_1) \otimes (\mu_{a,.}) \otimes \Gamma(\psi_2)]$ , the transition is allowed. In Table 4.1 the selection rules derived from this expression are shown.

**Table 4.1** Selection rules for the rotational transitions in an assymmetric top.

Dipole moment component	Transition type	$\Delta J$	$\Delta K_a$	$\Delta K_c$
$\mu_a$	<i>a - type</i>	$0, \pm 1$	$0, \pm 2, \dots$	$\pm 1, \pm 3, \dots$
$\mu_b$	<i>b - type</i>	$0, \pm 1$	$\pm 1, \pm 3, \dots$	$\pm 1, \pm 3, \dots$
$\mu_c$	<i>c - type</i>	$0, \pm 1$	$\pm 1, \pm 3, \dots$	$0, \pm 2, \dots$

Moreover, the computed absolute values of the electric dipole moment components will allow us to estimate the polarization power needed to optimally polarize the transitions belonging to a particular species. For instance, we will employ a relatively low polarization power to measure molecules with high dipole moment values. Thus, we can profit from this fact, play around with the polarization power and finally use the magnitude of the dipole moment components as a very useful tool for the conformational identification.

### 4.3. Nuclear quadrupole coupling constants

For molecules containing one or more nuclei with a nuclear spin angular momentum,  $I$ , greater than  $1/2$ , such nuclei show a non-spherical or asymmetric distribution of the nuclear charge distribution, leading to a non-vanished nuclear quadrupole moment ( $eQ$ ) (Gordy & Cook 1970). This will be the case of most biomolecules and many interstellar compounds which contains, at least, one nitrogen atom. In this context, its most abundant isotope,  $^{14}\text{N}$ , shows a nuclear spin  $I = 1$ . This quadrupolar moment interacts with the electric field gradient created by the rest of the molecular charges. Consequently, the nuclear spin,  $I$ , and the overall angular momentum of the molecule,  $J$ , are coupled, giving rise to a splitting in the rotational energy levels. The quantum number that describes the energy levels resulting for this coupling is denoted as  $F$  ( $F = I+J$ ), which can adopt values according to:

$$F = J + I, J + I - 1, \dots, J - I \quad (24)$$

As a consequence, rotational transitions show the so-called nuclear quadrupole hyperfine structure (HPS), whose resolution entails the use of high-resolution spectroscopic techniques (for an example see the rotational study of **cyanoacetamide** and **cyanoacetic acid**, Chapter V).

Hence, the rotational Hamiltonian employed to analyze the spectrum of these sort of molecules is built by two different terms: the rigid rotor Hamiltonian ( $H_R$ ), presented in the previous subsections, and an additional term that considers the nuclear quadrupole coupling effect ( $H_Q$ );

$$H = H_R + H_Q \quad (25)$$

$$H_Q = \frac{eQq_{ij}}{2J(2J-1)I(2I-1)} \left[ 3(IJ)^2 + \frac{3}{2}IJ - I^2J^2 \right]$$

where  $q_{ij}$  ( $i, j = a, b, c$ ) are the coefficients that describe the average electronic field gradient in terms of the coordinates in the principal axis orientation, which are related to the components of the nuclear quadrupole coupling tensor,  $\chi$ :

$$\chi_{ij} = eQq_{ij}; \chi = \begin{pmatrix} \chi_{aa} & \chi_{ab} & \chi_{ac} \\ \chi_{ba} & \chi_{bb} & \chi_{bc} \\ \chi_{ca} & \chi_{cb} & \chi_{cc} \end{pmatrix} \quad (26)$$

From the analysis of the rotational spectra and for nuclei that present a relatively weak coupling, such as the  $^{14}\text{N}$ , we can obtain, from the fit, the values of the diagonal elements of the tensor,  $\chi_{aa}$ ,  $\chi_{bb}$  and  $\chi_{cc}$  using the SPFIT/SPCAT programme package (Pickett 1991), due to their linear dependence  $\chi_{aa} + \chi_{bb} + \chi_{cc} = 0$ . These values are extremely sensitive to the electronic environment of the  $^{14}\text{N}$  nucleus and, therefore, they can be also used as spectroscopic tools for the conformational identification.

#### 4.4. Centrifugal distortion constants

In most of the real cases of study, the molecules are not completely rigid and, therefore, they can not be described using a rigid rotor approximation. This approach surmises that the positions of all atoms are fixed during the rotation motion and thus the potential energy does not change. However, there are plenty of molecules that are quite flexible. This flexibility implies a distortion of the nuclear positions during the rotation motion. What is more, this fact has an even more noticeable effect as the speed of the rotation increases (for higher  $J$  values), when the centrifugal forces contribute to increase the bond lengths and, consequently, the moments of inertia decrease. Hence, in general, the analysis of a rotational spectra at higher frequencies (higher  $J$ ), requires to take into account the effect of the distortion in the Hamiltonian ( $H_D$ ):

$$H = H_R + H_D \quad (27)$$

Watson formulated two different reductions of the rotational Hamiltonian (Watson 1977) the symmetric S-reduction and the asymmetric A-reduction, to describe the distortion of asymmetric molecules. In this Ph.D., we have employed both reductions, depending on the symmetry of the system under study, highlighting the use of the A-reduction for the study of **glycolamide** (Chapter IV) and **n- and i-butyraldehyde** (Chapter VI), where the centrifugal distortion constant up to the sextic terms were determined. In Table 4.2 we provide a list of the centrifugal distortion constants (up to the octic order) for both reductions.

**Table 4.2** Centrifugal distortion constants for the Watsons's A- and S-reduced Hamiltonians.

A Reduction	S Reduction
Quartic order	
$\Delta_J$	$D_J$
$\Delta_{JK}$	$D_{JK}$
$\Delta_K$	$D_K$
$\delta_J$	$d_1$
$\delta_k$	$d_2$
Sextic order	
$\Phi_J$	$H_J$
$\Phi_{JK}$	$H_{JK}$
$\Phi_{KJ}$	$H_{KJ}$
$\Phi_K$	$H_K$
$\phi_J$	$h_1$
$\phi_{JK}$	$h_2$
$\phi_K$	$h_3$
Octic order	
$L_J$	$L_J$
$L_{JK}$	$L_{JK}$
$L_{JK}$	$L_{JK}$
$L_{KKJ}$	$L_{KKJ}$
$L_K$	$L_K$
$l_J$	$l_1$
$l_{JK}$	$l_2$
$l_{KJ}$	$l_3$
$l_K$	$l_4$

## 5. REFERENCIAS

- Alonso, E. R., Kolesniková, L., Tercero, B., et al., *ApJ*, **2016**, 832, 42
- Alonso, E. R., Kolesniková, L., Białkowska-Jaworska, E., et al., *ApJ*, **2018**, 861, 70.
- Alonso, E.R. *Biomolecules and Interstellar molecules: Structure, Interactions and Spectroscopic Characterization*, **2018**, URI: <http://uvadoc.uva.es/handle/10324/30210>
- Alonso, E. R., León, I., Alonso, J.L., in *Dev. Phys. Theor. Chem.* (Ed.: E.R.B.T.-I.I.I.B.N.B.S. Bernstein), Elsevier, **2021**, pp. 93–141.
- Alonso, J. L., Lorenzo, F. J., López, J. C., et al. *Chem. Phys.*, 1997, 218, 267
- Alonso, J. L., Cocinero, E. J., Lesarri, A., et al., *Angew. Chemie Int. Ed.* **2006**, 45, 3471–3474.
- Alonso, J. L., Lozoya, M. A, Peña, I., et al., *Chem. Sci.* **2014**, 5, 515–522.
- Alonso, J. L., López, J. C., in *Top. Curr. Chem.* (Eds.: A.M. Rijs, J. Oomens), Springer International Publishing, Cham, **2015**, pp. 335–401.
- Andresen, U., Dreizler, H., Grabow, J.U., Stahl, W., *Rev Sci Instrum*, **1990**, 61:3694.
- Balle, T. J., Campbell, E. J., Keenan, M. R., and Flygare, W. H., *The Journal of Chemical Physics*, **1979**, 71, 2723–2724.
- Balle, T. J. and Flygare, W. H., *Review of Scientific Instruments*, **1981**, 52, 33–45.
- Becke, A.D. 1993, *J. Chem. Phys.*, 98, 5648-5652.
- Bergevin, J., Wu, T.-H., Yeak, J., *Nat. Commun.* **2018**, 9, 1273.
- Bermúdez, C., Mata, S., Cabezas, C., Alonso, J.L., *Angew. Chem. Int. Ed.* **2014**, 53, 11015 –11018.
- Blanco, S., López, J.C., Alonso, J.L., Ottaviani, P., Caminati, W., *J Chem Phys*, **2003**, 119:880
- Blanco, S., Sanz, M.E., López, J.C., Alonso, J.L., *PNAS*, **2007**, 104, 51, 20183-20188
- Brown, G.G., Dian B.C., Douglass, K.O., Geyer, S.M., Shipman, S.T., Pate, B.H, *Rev Sci Instrum*, **2008**, 79:053103.
- Cabezas, C., Cernicharo, J., Alonso, J. L., et al., *Astrophys. J.* **2013**, 775, 133.
- Cabezas, C., Varela, M., Alonso, J. L., *Angew. Chemie* **2017**, 129, 6520–6525.
- Daly, A., Kolesniková, L., Mata, S., Alonso, J., & J. L., *J. Mol. Spectrosc.*, **2014**, 306, 11.
- Ditchfield, R., Hehre, W., Pople, J., *J. Chem. Phys.* **1971**, 54, 724-728.
- Dunning, T. H., *J. Chem. Phys.*, **1989**, 90, 1007–1023.
- Endo, Y., Kohguchi, H., Ohshima, Y., *Faraday Discuss*, **1994**, 97:341.
- Eckart, C., *Phys. Rev.* **1968**, 47, 552.
- Fock, V. *Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. Z. f. Phys.* **1930**, 61, 126.
- Gaussian 09, Revision B.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K.

N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, **2009**.

Gaussian 16, Revision B.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Gaussian, Inc., Wallingford CT, **2016**.

Gonzalez, C.; Schlegel, H.B., *J. Chem. Phys.*, **1989**, *90*, 2154–2161.

Gonzalez, C.; Schlegel, H. B., *J. Phys. Chem.*, **1990**, *94*, 5523–5527.

Gordy, W., Cook, R. L. *Microwave Molecular spectra*, Interscience Pub, **1970**, New York.

Grabow, J. Stahl, W. and Dreizler, H., *Review of Scientific Instruments*, **1996**, *67*, 4072–4084.

Graner G, Hirota E, Iijima T, Kuchitsu K, Ramsay DA, Vogt J, Vogt N, *Structure data of free polyatomic molecules*. Basic data. Springer, **1998**, Berlin, 214 pp.

Grimme, S., *J. Chem. Phys.* **2006**, *124*, 034108– 16.

Grimme, S., Antony, J., Ehrlich, S. and Krieg, S., *J. Chem. Phys.* **2010**, *132*, 154104.

Halgren, T. A. *J. Comp. Chem.*; **1996**, 490-519.

Hayes, J. M., and Small, G. J., *Anal. Chem.* **1983**, *55*, 565A.

Hehre, W. J., *Ab-initio Molecular Orbital Theory*; Wiley-Interscience: **1986**.

Hohenberg, P. & Kohn, W. *Phys Rev.* **1964**, *136*, 864.

Hollas, J. M. *High Resolution Spectroscopy*, Jonh Wiley & Sons, United Kingdom, **1998**.

Hoy, A. R., Mills, I. M., & Strey, G., *Molecular Physics*, **1972**, *24*, 1265

Kerse, C., Kalaycıoğlu, H., Elahi, P., et al. Ilday, *Nature* **2016**, *537*, 84–88.

Kisiel, Z., Pszczółkowski, L., Medvedev, I. R., et al., *J. Mol. Spectrosc.*, **2005**, *233*, 231.

Koch, W., Holthausen, M. C., *A Chemist Guide to Density Functional Theory*, 2<sup>nd</sup> ed. Weinheim, Germany: Wiley-VCH, **2001**.

Kolesníková, L., León, I., Alonso, E. R., et al., *J. Phys. Chem. Lett.*, **2019**, *10*, 1325

Kretschmer, U., Consalvo, D., Knaack, A., Schade, W., Stahl, W., Dreizler, H., *Mol Phys*, **1996**, *87*:1159.

Kroto, H. W. *Molecular Rotation Spectra*, Dover Publications Inc., New York, **1992**.

Lee, C.; Yang, W.; Parr, R. G., *Phys. Rev. B: Condens. Matter Mater. Phys.*, **1988**, *37*, 785–789.

Lesarri, A., Mata, S., López, J.C., Alonso, J.L., *Rev Sci Instrum*, **2003**, *74*:4799.

- León, I., Alonso, E. R., Mata, S., et al., *Phys Chem Chem Phys* (Incorporating Faraday Transactions), **2017**, 19, 24985.
- León, I., Alonso, E. R., Mata, S. et al., *Angew. Chemie - Int. Ed.* **2019**, 58, 16002–16007.
- Levine, I.N. *Quantum Chemistry*. Englewood Cliffs, New jersey: Prentice Hall, **1991**, pp. 455–544.
- Levy, D. H., Wharton, L. and Smalley, R. E. *In Chemical and Biochemical Applications of Lasers*, Vol. II (Moore, C. B., ed.), pp. 1-41, Academic Press, New York, **1977**.
- Levy, D. H., *Science*, **1981**, 214, 263 55
- Lovas, F.J., Kawashima, Y., Grabow, J.-U., Suenram, R.D., Fraser, G.T., Hirota, E., *Astrophys J Lett*, **1995**, 455: L201.
- Mata, S., Peña, I., Cabezas, C., López, J.C., Alonso, J.L., *J Mol Spectrosc*, **2012**, 280:91–96.
- McCarthy, M.C., Thaddeus, P., *Chem Soc Rev*, **2001**, 30:177.
- McCarthy, M.C., Lee, K.L.K., Carroll, P.B., et al., *J Phys Chem A*, **2020**, 124 (25):5170-5181.
- Moller, C. Plesset, M. S. *Phys. Rev.* **1934**, 46, 618.
- Mori, T., Suma, K., Sumiyoshi, Y., & Endo, J., *J. Chem. Phys.*, **2009**. 130, 204308.
- Oppenheimer, M. Born, J. R. *Zur Quantentheorie der Molekeln*. Ann. Der Phys. **1927**, 84, 457.
- Peng, C. and Schlegel, H. B., *J. Chem.*, 33, **1993**, 449-54
- Peng, C., Ayala, P. Y., Schlegel, H. B. and Frisch, M. J., *J. Comp. Chem.*, 17, **1996**, 49-56.
- Peña, I. Mata, S., Martín et al. *Phys Chem Chem Phys*, **2013**, 7; 15(41):18243-8.
- Peña, I., Cabezas, C., Alonso, J. L., *Angew. Chemie - Int. Ed.* **2015**, 54, 2991–2994.
- Pickett, H. M., *Journal of Molecular Spectroscopy*, **1991**, 148, 371-377.
- Pople, J. A., Head-Gordon, M. Raghavarachi, K.. *The Journal of Chemical Physics. American Institute of Physics.* **1987**, 87, 5968–35975.
- Purvis G. D., Bartlett, R. J., *The Journal of Chemical Physics*, **1982**, 76.
- Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A., *Chem. Phys. Lett.*, **1989**, 157, 479–483.
- Russo, R. E., Mao, X., Gonzalez, J. J., et al., *Anal. Chem.* **2013**, 85, 6162–6177.
- Schwabe, T., & Grimme, S., *Physical Chemistry Chemical Physics* (Incorporating Faraday Transactions), **2007**, 9, 3397
- Slater, J.C., *Note on Hartree's Method*. *Phys. Rev.* **1930**, 35, 210.
- Stafe, M., Marcu, A., Puscas, N., *Pulsed Laser Ablation of Solids*, Springer, **2014**.
- Stanton, J.F., Gauss, L. Cheng, M.E. et al. and the integral packages *MOLECULE* (Almlöf, J., and Taylor, P.R.), *PROPS* (Taylor, P.R.), *ABACUS* (T. Helgaker, H.J. Aa. Jørgensen, J.J., and Olsenm J.), and ECP routines by Mitin A. V., and van Wüllen, C. For the current version, see <http://www.cfour.de>.
- Suenram, R.D., Lovas, F.J., Fraser, G.T., *J Mol Spectrosc*, **1988**, 127:472 30
- Suenram, R.D., Lovas, F.J., Fraser, G.T., Matsumura, K., *J Chem Phys*, **1990**, 92:4724 40.
- Sutter, D.H., Dreizler, H., *Z. Naturforsch A. Phys. Sci.*, **2001**, 56:425.



Szabo, A., and Ostlund, N.S., *Mod. Quant. Chem.: Introduction to advanced electronic structure theory*. New York: Dover Publishing, Mineola, **1996**.

Walker, K.A., Gerry, M.C.L., *J Mol Spectrosc*, **1997**, 182:178.

Watts, J.D., Gauss, J., Bartlett, R.J., *J. Chem. Phys.*, **1993**, 98, 8718-8733.

Watson, J. K. G. *Vibrational Spectra and Structure*, Vol. 6 , Elsevier: Amsterdam, **1977**.

Welzel, A., Stahl, W. *Phys Chem Chem Phys*, **1999**, 1:5109 31.

Woon, D. E.; Dunning, T. H., *J. Chem. Phys.*, **1993**, 98, 1358–1371.



## IV. EXPERIMENTAL AND THEORETICAL STUDY OF INTERSTELLAR GLYCINE ISOMERS

Adapted from Sanz-Novo et al. 2019, *ACS Earth Space Chem.* 3, 1170-1181; Sanz-Novo et al. 2020, *Astronomy & Astrophysics*, 639, A135, and Sanz-Novo et al. 2021 in prep.



The fundamental question whether glycine exists or not in the interstellar medium is one of the most pursued targets in astrochemistry. Despite several attempts to observe glycine in the ISM have been reported, its detection has never been confirmed. Moreover, observation of amino acids and its most important isomers in

the ISM and in solar system bodies should be of crucial importance for revealing the chemistry that may have led to life's origin. In this chapter, a preliminar computational study on the reaction of protonated hydroxylamine and acetic acid has been carried out, leading to different protonated glycine isomers with  $[\text{H}_6\text{C}_2\text{O}_2\text{N}]^+$  molecular formula. We found feasible process under interstellar conditions which leads to protonated acetohydroxamic acid. Therefore, the neutral analogue is suggested as a candidate for interstellar detection. Afterward, as a prerequisite step for astronomical identification, we recorded the rotational spectra of acetohydroxamic acid ( $\text{CH}_3\text{C}(\text{O})\text{NHOH}$ ) and glycolamide ( $\text{CH}_2(\text{OH})\text{C}(\text{O})\text{NH}_2$ ), relevant glycine isomers and also some of the simplest derivatives of acetamide (e.g., one hydrogen atom is replaced by a hydroxyl group), which is a well-known interstellar molecule. In this context, we employed a battery of state-of-the-art rotational spectroscopic techniques in the frequency and time domain to provide direct experimental frequencies of the ground vibrational state of these glycine isomers in the centimeter-, millimeter- and submillimeter-wavelength regions (only for glycolamide) in order to enable their identification in the interstellar medium. Then, we used the spectral line survey named Exploring Molecular Complexity with ALMA (EMoCA), which was performed toward the star forming region Sgr B2(N) with ALMA to search for glycolamide in space. We also searched for glycolamide toward Sgr B2(N) with the Effelsberg radio telescope. The astronomical spectra were analyzed under the local thermodynamic equilibrium approximation. Finally, we used the gas-grain chemical kinetics model MAGICKAL to interpret the results of the astronomical observations.

## 1. INTRODUCTION

Amino acids are essential components of living systems since they play a crucial role as building blocks of proteins. They were the first prebiotic molecules to be identified in the Miller experiment (Miller et al. 1953, Johnson et al. 2008) from the reaction of simple organic molecules. Also, taking into account their presence in some chondritic meteorites (Pizzarello et al. 2006, 2010; Burton et al. 2012), amino acids should be among the easiest biomonomers to synthesize (Ruiz-Mirazo et al. 2014). Despite several radioastronomical searches, so far, the smallest amino acid, glycine, has not been conclusively identified in the interstellar medium (ISM) (Synder et al. 2005; Cunningham et al. 2007; Jones et al. 2007; Hollis et al. 2003).

There is a question as to whether glycine will ever be detectable even if present, given the difficulties for its possible detection in space (Irvine et al. 1998). One of these difficulties arises from its rotational spectrum features with relatively weak lines due to its large molecular partition function. Secondly, amino acids are highly susceptible to UV photo destruction, and they will likely be destroyed during the lifetime of a typical interstellar cloud (Ehrenfreund et al. 2001). Consequently, they only survive in shielded environments. Finally, from a chemical-physics point of view, it might happen that efficient synthetic routes for the production of amino acids under interstellar conditions do not exist.

In this regard, as previously mentioned, exothermic ion-molecule reactions between positive ions and neutral molecules play a crucial role in the synthesis of organic interstellar molecules in the gas phase since these processes tend to have no activation energy barriers (Herbst et al. 2001). In particular, the feasibility of the gas-phase formation of protonated glycine from the reaction of acetic acid and protonated hydroxylamine was studied by Bohme and co-workers (Snow et al. 2007; Blagojevic et al. 2003). Both reactants are already detected interstellar molecules. Acetic acid ( $\text{CH}_3\text{C}(\text{O})\text{OH}$ ) was detected in the 3 mm region toward the massive star-forming region Sgr B2(N-LMH) (Mehring et al. 1997) as well in the additional hot core source W51e2 (Remijan et al. 2002). Meanwhile, hydroxylamine ( $\text{NH}_2\text{OH}$ ), which is considered as one of the main precursors in the formation of amino acids in the interstellar medium, has been recently identified toward the quiescent molecular cloud G+0.693-0.027 located in the Galactic Center (Rivilla et al. 2020).

In this chapter, following Jeanvoine et al. (2018) conclusions (see Section 5.1 of Chapter II: Introduction) we have carried out a computational study of the potential energy surfaces (PES) corresponding to the reactions of formation of different products, all of them with mass-overcharge ratio,  $m/z$  76 and  $[\text{H}_6\text{C}_2\text{O}_2\text{N}]^+$  molecular formula. These products together with a water molecule are obtained from the reaction of protonated hydroxylamine and acetic acid. For each one of the reaction products we provide theoretical spectroscopic data that could guide their plausible detection in a terrestrial laboratory. In addition, we also give information for their neutral counterparts, with the aim to aid their laboratory or astronomical detection by radioastronomy or infrared (IR) spectroscopy as well as to analyse their behaviour upon protonation.

Moreover, it is of great astrochemical interest to study peptide-like bond molecules, given that this type of bond can play an essential role in the generation of proteins in the early stages of Earth (Belloche et al. 2017, Colzi et al. 2021). The simplest molecule containing a peptide bond, formamide ( $\text{NH}_2\text{CHO}$ ), was detected in the interstellar medium (ISM) in the 1970s in Sagittarius B2 (Rubin et al. 1971) and later in Orion KL (Turner, 1991). One of the simplest derivatives of formamide that also contains a peptide bond, acetamide ( $\text{CH}_3\text{C}(\text{O})\text{NH}_2$ ), has already been detected in the ISM (Hollis et al. 2006; Halfen et al. 2011). Very recently,  $\text{HNCO}$ ,  $\text{HC}(\text{O})\text{NH}_2$ ,  $\text{CH}_3\text{NCO}$ ,  $\text{CH}_3\text{C}(\text{O})\text{NH}_2$ , and  $\text{CH}_3\text{NHCHO}$  have been detected for the first time towards the hot core G31.41+0.31, outside the Galactic centre (Colzi et al. 2021). Following this, we propose the study of acetoxyamic acid ( $\text{CH}_3\text{C}(\text{O})\text{NHOH}$ ) and glycolamide ( $\text{CH}_2(\text{OH})\text{C}(\text{O})\text{NH}_2$ ) as potential interstellar molecules since they fit in very nicely within the chemistry of the ISM

Hydroxamic acids (HA) constitute a family of natural compounds obeying the formula ( $\text{RCONR}'\text{OH}$ ; R, R' = H, aryl, or alkyl) and usually exhibiting biological activity (Miller 1989; Gupta 2013). HA are N-hydroxy substituted derivatives of amines and involve, like amides, the fragment of the simplest protein structure  $\text{HNC}=\text{O}$ . There has been increasing interest in their roles as specific enzymes inhibitors, since HA possess a moiety that is as a key structural element in many highly potent and selective inhibitors against a variety of biocatalysts (Muri et al. 2002; Marmion et al. 2004), being compounds of potential chemotherapeutics aiming at cardiovascular diseases, Alzheimer's disease, and HIV (Pal. & Saha 2012). Among them, acetoxyamic acid (AHA), apart from being an interesting glycine isomer, it has been clinically used under the name of Lithostat

for the treatment of urinary tract infections by urease inhibition (Griffith et al. 2017) and shows many applications as a chelating agent in coordination chemistry (Farkas et al. 2001; Codd 2008). Nevertheless, besides the FTIR studies by matrix isolation techniques (Saldyka & Mielke 2003, 2007 and 2018), it has only been investigated in the condensed phases by X-ray spectroscopy (Bracher & Small 1970), remaining unexplored in the isolation conditions of the gas phase.

On the other hand, glycolamide was detected experimentally by Nuevo et al. (2010) when they simulated conditions relevant to the interstellar medium and Solar System icy bodies such as comets, further increasing the astrophysical appealing of this system. In their experimental setup, a condensed  $\text{CH}_3\text{OH} : \text{NH}_3 = 11 : 1$  ice mixture was UV irradiated at about 80 K and urea, glycolic acid, glycerol, glycerolic acid, and glycerol amide were also found as products. Also, reactions involving glycolamide are relevant, since it can participate in polycondensation reactions, leading to the formation of N-(2-amino-2-oxoethyl)-2-hydroxyacetamide, a derivative of glycine dipeptide (Szóri et al. 2011), already investigated by high-resolution rotational spectroscopy (Cabezas et al. 2017).

Furthermore, glycolamide stands as a challenging problem for gas-phase millimeter-wave rotational studies due to its high melting point (122–126 °C), which is related to condensation problems. Therefore, its spectrum so far has remained unexplored in these high-frequency regions (see previous Stark-modulation study, Maris 2004). High-quality data at mm/submm wavelengths are also necessary for the eventual astronomical detection of glycolamide with observational facilities such as the radio telescopes of the Institut de RadioAstronomie Millimétrique (IRAM) or the Atacama Large Millimeter/submillimeter Array (ALMA). This motivated us to record the millimeter and submillimeter-wave spectrum of glycolamide up to 460 GHz using the heated-cell millimeter-wave spectrometer at the University of Valladolid.

Despite the importance of these molecules as appealing candidates in the ISM, scarce rotational information of the isolated systems has been yet reported. In this work, we use a battery of state-of-the-art rotational spectroscopic techniques in the frequency and time domain to derive a precise set of spectroscopic constants for acetohydroxamic acid and glycolamide. The  $^{14}\text{N}$  nuclear quadrupole coupling constants need to be determined together with the rotational constants to correctly reproduce the spectrum, which is a prerequisite for its identification in astronomical spectra. This is especially true for

observations in low-frequency regions accessible with the Green Bank Telescope (GBT), the Yebes 40-m or the Effelsberg 100 m telescopes, for which the hyperfine components of transitions with low angular momentum number  $J$  are generally spread over several MHz.

One of the best astronomical sources to search for complex organic molecules in the interstellar medium is the giant molecular cloud complex Sagittarius B2 (Sgr B2), located in the galactic center region at a distance of 8.2 kpc (Reid et al. 2019). This star forming region contains the protocluster Sgr B2(N) in which several hot molecular cores are embedded (e.g., Bonfand et al. 2019). Many complex organic molecules were first detected toward Sgr B2(N) (see, e.g., McGuire 2018, for a census). Several spectral line surveys were done toward Sgr B2(N) over the past four decades. We use one of the most recent ones, the imaging spectral line survey EMOCA, which stands for Exploring Molecular Complexity with ALMA (Belloche et al. 2016), performed with ALMA, to search for glycolamide and compare its abundance to other related organic molecules. We also targeted the whole Sgr B2(N) region with the Effelsberg telescope to search for glycolamide in its extended molecular envelope.

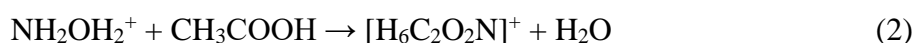
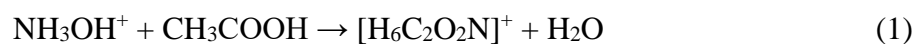
## 2. RESULTS AND DISCUSSION

### 2.1 *Theoretical study of neutral and protonated glycine isomers*

#### 2.1.1 *Stability and formation processes*

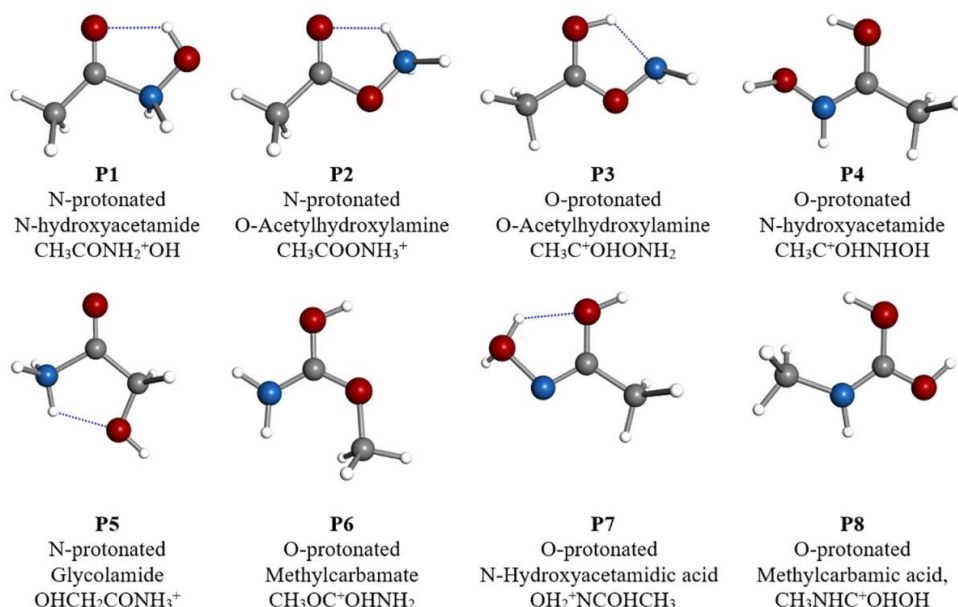
According to the Snow et al. 2007 experiments both isomers of protonated hydroxylamine (the most stable one,  $\text{NH}_3\text{OH}^+$ , and the high energy  $\text{NH}_2\text{OH}_2^+$ ) could be obtained in the protonation process of hydroxylamine using  $\text{CH}_5^+$  as protonating agent. Neither of them reacts with  $\text{H}_2$  under space conditions (Largo et al. 2009) and therefore both, if present in the interstellar medium, should be able to react with other molecules such as acetic acid.

In principle, acetic acid may react with both isomers of protonated hydroxylamine to produce protonated glycine and water, although other different products also with  $[\text{H}_6\text{C}_2\text{O}_2\text{N}]^+$  molecular formula could be formed. Both processes can be summarized as follows:



Given the spin multiplicity of the reactants,  $\text{NH}_3\text{OH}^+$  ( $^1\text{A}'$ ),  $\text{NH}_2\text{OH}_2^+$  ( $^1\text{A}$ ),  $\text{CH}_3\text{COOH}$  ( $^1\text{A}'$ ), the reaction takes place on the singlet potential energy and all of the products are in their singlet electronic state ( $^1\text{A}$ ).

In the dynamics simulations study by Jeanvoine et al. (2018) eight products, denoted as P1-P8, are obtained from the reaction of protonated hydroxylamine with acetic acid. The chemical structures of the P1-P8 reaction products are schematized in Fig. 1. Before computation of reaction energies, we performed a detailed conformational analysis, for each one of the P1-P8 reaction products. It should be noted that, in the case of P1 and P3 products, we have found lowest-lying conformational structures including intramolecular hydrogen bonds that are somewhat different to those considered in the previous work.



**Figure 1.** Chemical structures of  $[\text{H}_6\text{C}_2\text{O}_2\text{N}]^+$  (P1-P8) isomers optimized at MP2/aug-cc-pVTZ level of theory.

In Table 1, we collect the relative energies, with respect to reactants, of the possible products that can be formed in the reaction between either  $\text{NH}_3\text{OH}^+$  or  $\text{NH}_2\text{OH}_2^+$  and  $\text{CH}_3\text{COOH}$  computed at different levels of theory. At the CCSD(T) level of theory, hydroxylamine protonated in its oxygen atom,  $\text{NH}_2\text{OH}_2^+$ , was found to lie  $25.4 \text{ kcal mol}^{-1}$  higher in energy than the most stable isomer,  $\text{NH}_3\text{OH}^+$ . Thus, this value will be the difference of the reaction energies of the (2) and (1) processes computed at this level of theory. The energy barrier corresponding to the  $\text{NH}_3\text{OH}^+ \rightarrow \text{NH}_2\text{OH}_2^+$  isomerization



process is 50.71 kcal/mol at the CCSD(T) level of theory, consequently, both isomers,  $\text{NH}_3\text{OH}^+$  and  $\text{NH}_2\text{OH}_2^+$ , could coexist when hydroxylamine is protonated. For a detailed description of the computational methodology employed in this work see Section 3 of Chapter III.

**Table 1.** Relative energies (referred to reactants), including zero point corrections, in kcal mol<sup>-1</sup>, obtained at different levels of theory, for the reaction between protonated hydroxylamine and acetic acid yielding protonated glycine (GlyH<sup>+</sup>) and its structural isomers (P1-P8).

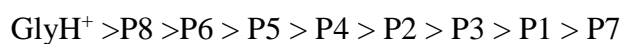
Reaction <sup>a</sup>	B3LYP <sup>b</sup>	MP2 <sup>c</sup>	CCSD <sup>d</sup>	CCSD(T) <sup>e</sup>
<b>GlyH<sup>+</sup> + H<sub>2</sub>O</b>	-51.77	-59.12	-56.08	-56.29
<b>GlyH<sup>+</sup> + H<sub>2</sub>O</b>	-76.59	-85.41	-83.62	-81.69
<b>P8 + H<sub>2</sub>O</b>	-50.87	-54.86	-53.66	-52.86
<b>P8 + H<sub>2</sub>O</b>	-75.69	-81.15	-81.20	-78.26
<b>P6 + H<sub>2</sub>O</b>	-48.56	-51.38	-50.70	-49.88
<b>P6 + H<sub>2</sub>O</b>	-73.38	-77.67	-78.24	-75.28
<b>P5 + H<sub>2</sub>O</b>	-31.09	-38.13	-35.75	-36.02
<b>P5 + H<sub>2</sub>O</b>	-55.91	-64.43	-63.30	-61.41
<b>P4 + H<sub>2</sub>O</b>	-11.49	-12.74	-12.84	-12.83
<b>P4 + H<sub>2</sub>O</b>	-36.31	-39.03	-40.38	-38.23
<b>P2 + H<sub>2</sub>O</b>	-8.64	-11.86	-9.68	-10.95
<b>P2 + H<sub>2</sub>O</b>	-33.46	-38.15	-37.23	-36.35
<b>P3 + H<sub>2</sub>O</b>	-6.66	-7.59	-7.62	-8.38
<b>P3 + H<sub>2</sub>O</b>	-31.48	-33.89	-35.16	-33.78
<b>P1 + H<sub>2</sub>O</b>	5.42	0.89	2.71	1.52
<b>P1 + H<sub>2</sub>O</b>	-19.41	-25.40	-24.83	-23.88
<b>P7 + H<sub>2</sub>O</b>	15.85	14.60	14.43	13.29
<b>P7 + H<sub>2</sub>O</b>	-8.97	-11.69	-22.54	-12.11

<sup>a</sup>The first entry refers to the  $\text{NH}_3\text{OH}^+ + \text{CH}_3\text{COOH}$  reaction and the second one to the  $\text{NH}_2\text{OH}_2^+ + \text{CH}_3\text{COOH}$  reaction. <sup>b</sup>Electronic energy calculated at the B3LYP/cc-pVTZ. <sup>c</sup>Electronic energy calculated at the MP2/aug-cc-pVTZ levels. <sup>d</sup>Electronic energy calculated at the CCSD/aug-cc-pVTZ//MP2/aug-cc-pVTZ level. <sup>e</sup>Electronic energy calculated at the CCSD(T)/aug-cc-pVTZ//MP2/aug-cc-pVTZ level.

Excepting for the formation of the P1 (N-protonated N-hydroxyacetamide or acetohydroxamic acid,  $\text{CH}_3\text{CONH}_2^+\text{OH}$ ) and P7 (O-protonated N-Hydroxyacetamidic acid,  $\text{CH}_3\text{COHNOH}_2^+$ ) products from the reaction of  $\text{NH}_3\text{OH}^+$  with acetic acid, all reactions studied including those giving protonated glycine are exothermic processes. The most favorable process, from thermodynamic arguments, is the formation of protonated

glycine ( $\text{GlyH}^+$ ,  $\text{NH}_3^+\text{CH}_2\text{COOH}$ ). However, the P8 (O-protonated Methylcarbamic acid,  $\text{CH}_3\text{NHC}(\text{OH})_2^+$ ) and the P6 (O-protonated Methylcarbamate,  $\text{CH}_3\text{OC}(\text{OH})^+\text{NH}_2$ ) products were found to be very close in energy to the most stable isomer ( $3.43 \text{ kcal mol}^{-1}$  and  $6.41 \text{ kcal mol}^{-1}$  respectively at the CCSD(T) level). The P5 (N-protonated Glycolamide,  $\text{HOCH}_2\text{CONH}_3^+$ ) product was located  $20.28 \text{ kcal mol}^{-1}$  (at the CCSD(T) level) above protonated glycine and the remaining reaction products were predicted to lie more than  $40 \text{ kcal mol}^{-1}$  above protonated glycine.

Regardless of the level of theory used, the stability order of the isomers with  $[\text{H}_6\text{C}_2\text{O}_2\text{N}]^+$  molecular formula studied in this work at CCSD(T)/aug-cc-pVTZ level is (> means more stable than):



If a comparison is made among the results obtained at different level of theory, we found a stabilization of the P8 and P6 isomers when B3LYP methodology is employed. At this level of theory, protonated glycine and the P8 isomers are almost isoenergetic, with  $\text{GlyH}^+$  being slightly preferred by  $0.9 \text{ kcal mol}^{-1}$ . On the other hand, the CCSD(T) level tends to predict lower relative energies than the MP2 one.

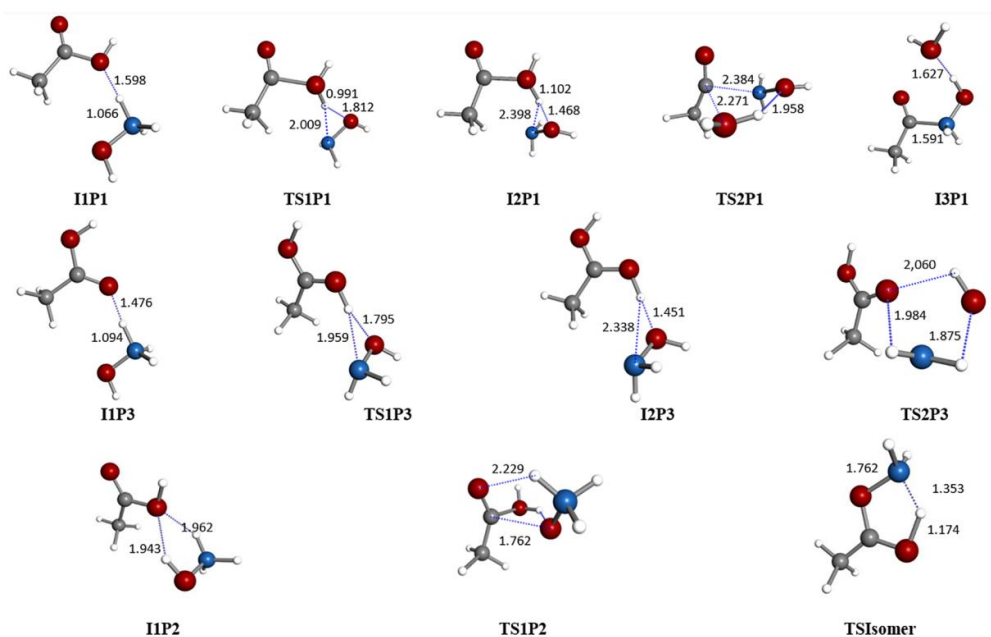
A large energy difference was found between the two ends of the isomeric system, where protonated glycine,  $\text{GlyH}^+$ , (the most stable) and P7 (the less stable) differ by approximately  $70 \text{ kcal mol}^{-1}$ . This difference could be expected, and it mainly arises from the different chemical nature of isomers. The multiple topological dispositions of the eleven atoms that make up the different structures allow to generate species with different functional groups. In addition, the arrangement of the NH and OH groups leads to a wide range of possibilities for the formation of intramolecular hydrogen bonds that significantly stabilize this chemical system.

The relative stabilities of some of the protonated glycine isomers were previously computed at the B3LYP/6-311G(d,p) level of theory (Lattelais et al 2011). From that study, protonated glycine was predicted to be the most stable isomer of the family with the  $\text{CH}_3\text{NHC}(\text{OH})_2^+$  (P8) and the  $\text{CH}_3\text{OC}(\text{OH})^+\text{NH}_2$  (P6) isomers located  $2.2$  and  $4.7 \text{ kcal mol}^{-1}$ , respectively, higher in energy.

Even though protonated glycine is the most stable species among  $[\text{H}_6\text{C}_2\text{O}_2\text{N}]^+$  molecular formula species, the chemical dynamics simulations study of the reaction between protonated hydroxylamine and acetic acid found that P1 and P2 products were

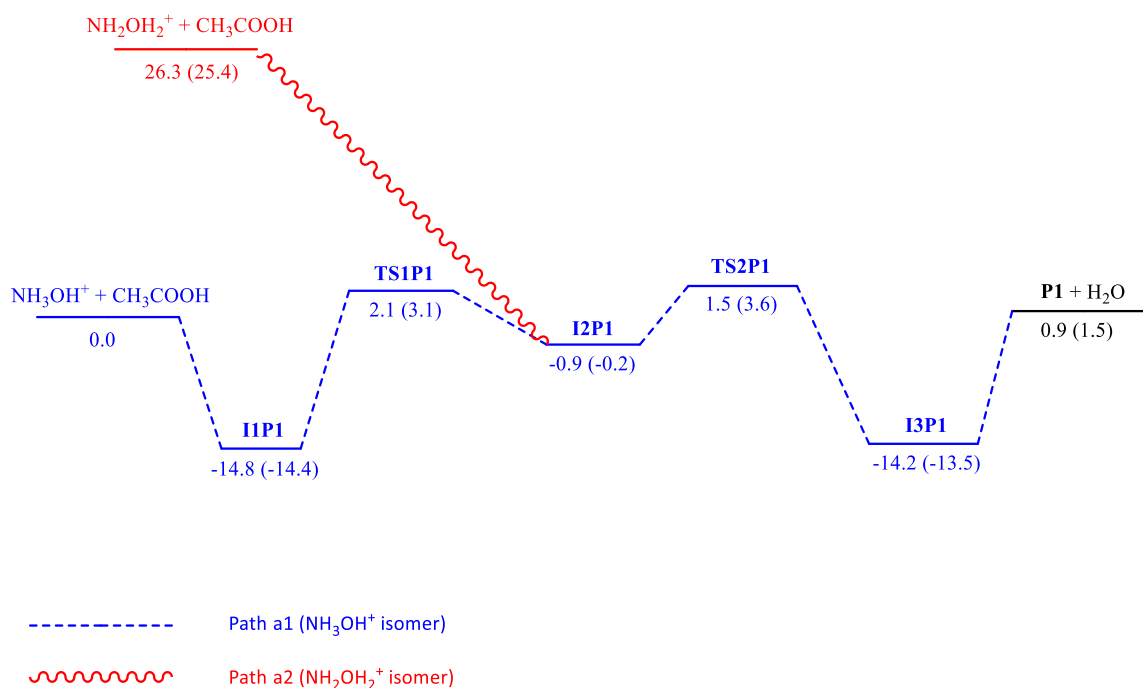
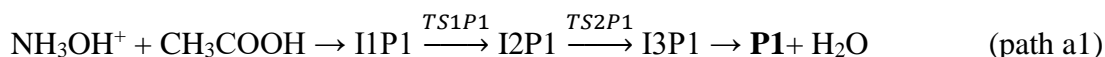
the most abundant isomers. It should be noted that these products are located 57.81 and 47.91 kcal mol<sup>-1</sup> respectively, at the CCSD(T) level of theory, higher in energy than protonated glycine. Only a small number of P3 and P5 products were observed at the two temperatures (5 and 300 K) considered in Jeanvoine et al 2018. Therefore, it should be useful to carry out a full exploration of the PES, to analyze the feasibility of the reactions studied as a source of the P1-P3 species in the interstellar medium. Given the conditions of interstellar medium, essentially low temperatures and low densities, interstellar plausible reactions should be exothermic and barrier free processes.

Fig. 2 depicts the MP2/aug-cc-pVTZ optimized geometries for the intermediate species and transition states, with [H<sub>8</sub>C<sub>2</sub>O<sub>3</sub>N]<sup>+</sup> molecular formula, involved in the formation process of the P1-P3 isomers. The energy profiles for these reactions computed at MP2 and CCSD(T) levels of theory are represented in Fig. 3-5. In these representations, the energy of reactants is taken as a reference and the notation I1PX, I2PX, ..., TS1PX, TS2PX, ... (X=1,2,3) is used for naming intermediates and transition states, respectively, involved in the corresponding reactions. As can be inferred from the figures, in general, a good agreement between the results obtained with the MP2 and CCSD(T) methodologies was found.



**Figure 2.** Intermediates and transition states involved on the reaction of formation of the P1-P3 isomers. Geometries are optimized at the MP2/aug-cc-pVTZ level of theory. Bond lengths are given in Angstroms.

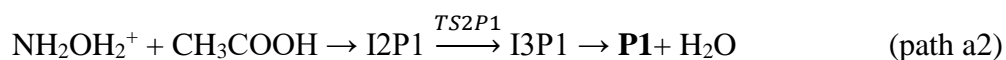
The energy profiles for the reactions of formation of the P1 isomer are shown in Fig. 3. When the most stable isomer of protonated hydroxylamine,  $\text{NH}_3\text{OH}^+$ , is considered the reaction starts with the formation of the I1P1 intermediate, which is the result of the interaction between the hydroxylic oxygen of acetic acid with one of the hydrogen atoms of the  $\text{NH}_3$  group of protonated hydroxylamine. The exothermic formation of this first intermediate ( $14.4 \text{ kcal mol}^{-1}$  at the CCSD(T) level) produces an energy reservoir that is used as the reaction proceeds toward the products. Once I1P1 is formed, a hydrogen migration from nitrogen to oxygen takes place through the TS1P1 transition state which is located  $3.1 \text{ kcal mol}^{-1}$  ( $1560 \text{ K}$ ) above reactants (at the CCSD(T)) level) giving the I2P1 intermediate. This intermediate evolves to the I3P1 intermediate through the TS2P1 transition state, which is located,  $3.6 \text{ kcal mol}^{-1}$  ( $1812 \text{ K}$ ), at the CCSD(T) level, above reactants. In the TS2P1 transition state, a water molecule is coordinated to both  $\text{NH}_2\text{OH}$  and the highly electrophilic ion  $\text{COCH}_3^+$ . Thus, when the  $\text{NH}_2$  group is properly oriented a N-C bond is formed, giving the I3P1 intermediate. This intermediate finally evolves to the formation of the P1 product and a water molecule. The process can be summarized as follows:



**Figure 3.** Energy profile, in kcal mol<sup>-1</sup>, for the reactions of protonated hydroxylamine with acetic acid producing **P1** computed at the MP2/aug-cc-pVTZ and CCSD(T)/aug-cc-pVTZ (in parentheses) levels of theory. Zero-point vibrational energy computed at the MP2/aug-cc-pVTZ level is included.

As can be seen in Fig. 3, the global process is slightly endothermic ( $\Delta E=1.5$  kcal mol<sup>-1</sup> at the CCSD(T) level), and has a small but non-negligible activation barrier of 3.6 kcal mol<sup>-1</sup> (1812 K) at the CCSD(T) level of theory. This barrier precludes this mechanism from taking place in the interstellar medium.

The first step in the mechanism of the formation of the P1 product from the less stable isomer of protonated hydroxylamine is the direct interaction between one of the hydrogen atoms bonded to oxygen of NH<sub>2</sub>OH<sub>2</sub><sup>+</sup> and the hydroxylic oxygen of CH<sub>3</sub>COOH giving the I2P1 intermediate. This complex finally evolves, through TS2P1, to produce P1 and H<sub>2</sub>O as it is schematized in path a2:

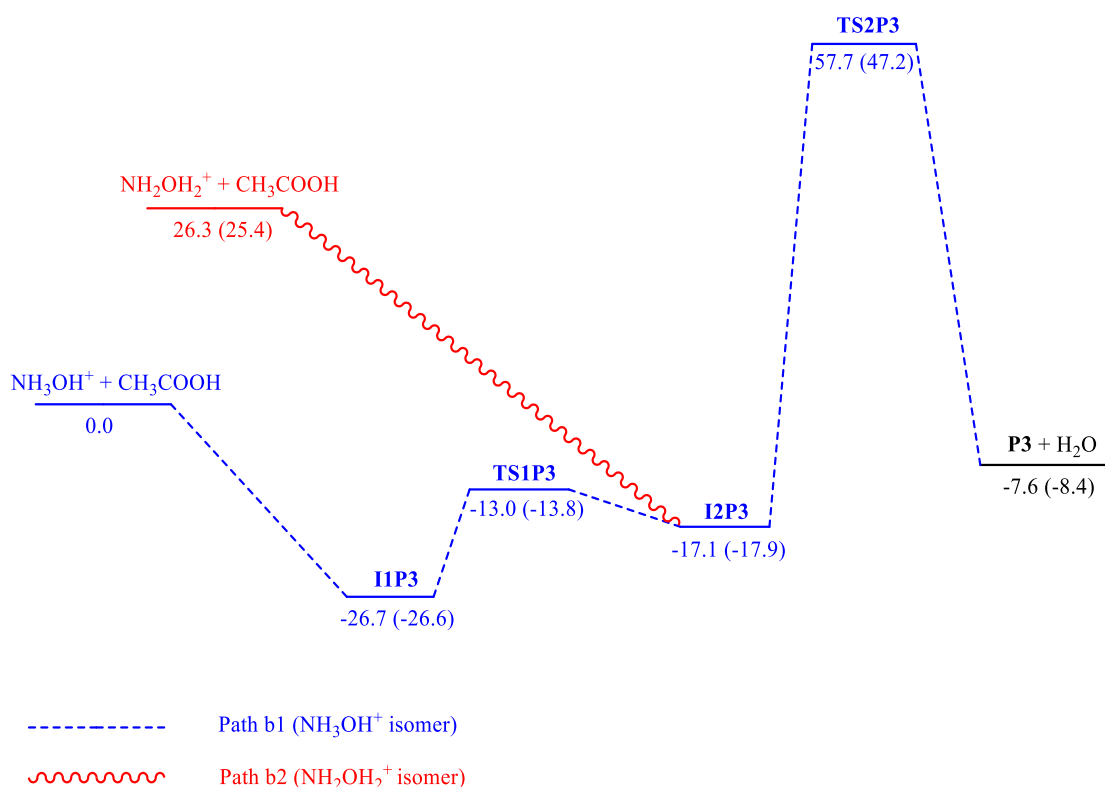


Pathway a2 is a clearly exothermic process ( $\Delta E=-23.9$  kcal mol<sup>-1</sup> at the CCSD(T) level) with no net activation barrier since now the TS2P1 transition state lies 21.8 kcal mol<sup>-1</sup> below reactants. Therefore, the formation of the P1 isomer could be feasible under interstellar conditions from the reaction of the less stable isomer of protonated hydroxylamine and acetic acid.

We will consider in advance the analysis of the PES corresponding to the formation of the P3 isomer because the process of formation of the P2 isomer can be derived from that of the P3 product.

Regarding formation of the P3 isomer, depicted in Fig. 4, the mechanism of the reaction, involving the most stable isomer of protonated hydroxylamine, starts with the approach of one of the hydrogen atoms bonded to the nitrogen of NH<sub>3</sub>OH<sup>+</sup> to the carbonyl oxygen of acetic acid, giving rise to the I1P3 intermediate. Once I1P3 is obtained, a proton transfer from the NH<sub>3</sub> group to the carboxylic oxygen, through the TS1P3 transition state, occurs forming the I2P3 complex. As can be seen from Fig. 4, TS1P3 is located lower in energy than the reactants (-13.8 kcal mol<sup>-1</sup> at the CCSD(T) level). From an appropriate

spatial arrangement of atoms in the I2P3 intermediate, a N-O bond cleavage and a simultaneous approach of the NH<sub>2</sub> group to the carbonyl oxygen takes place through the TS2P3 transition state which is located 47.2 kcal mol<sup>-1</sup> (23752 K) above reactants (at the CCSD(T) level). This concerted process will eventually result in the formation of P3. This process can be summarized as:

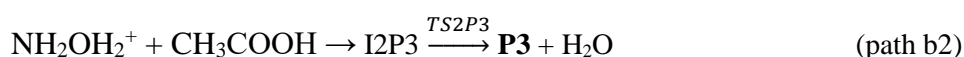


**Figure 4.** Energy profile, in kcal mol<sup>-1</sup>, for the reactions of protonated hydroxylamine with acetic acid producing P3 computed at the MP2/aug-cc-pVTZ and CCSD(T)/aug-cc-pVTZ (in parentheses) levels of theory. Zero-point vibrational energy computed at the MP2/aug-cc-pVTZ level is included.

From this energy profile it seems that production of the P3 isomer from the lowest-lying isomer of protonated hydroxylamine even though to be an exothermic process ( $\Delta E = -8.4$  kcal mol<sup>-1</sup> at the CCSD(T) level) involves a large energy barrier (23752 K),

since TS2P3 lies clearly higher in energy than reactants and it will not be feasible under interstellar conditions.

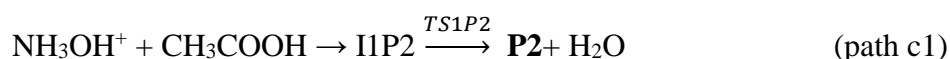
When the high energy isomer of protonated hydroxylamine is considered, the reaction starts with the direct approach of one of the hydrogen atoms bonded to oxygen of  $\text{NH}_2\text{OH}_2^+$  to the carbonylic oxygen of  $\text{CH}_3\text{COOH}$  giving the I2P3 intermediate. From this intermediate, as in the b1 pathway, the P3 isomer can be obtained from the migration of the  $\text{NH}_2$  group and subsequent elimination of a water molecule through the TS2P3 transition state. TS2P3 is now located  $21.8 \text{ kcal mol}^{-1}$  ( $10970 \text{ K}$ ) above reactants at the CCSD(T) level. The process can be represented as:



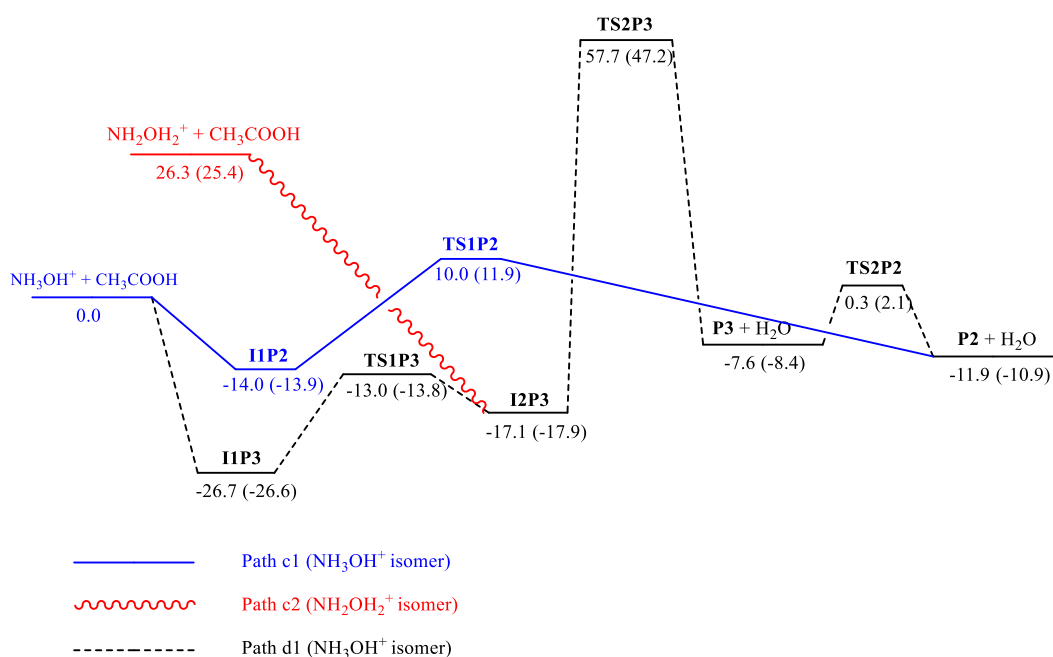
Therefore, the formation of the P3 isomer from the reaction of the less stable isomer of protonated glycine and acetic acid is an exothermic process ( $\Delta E = -33.8 \text{ kcal mol}^{-1}$ ) however, a net activation barrier of  $21.8 \text{ kcal mol}^{-1}$  was found and this process will not be allowed under interstellar conditions.

It should be noted that the PES for the P3 isomer was previously studied (Jobst et al. 2008) in an experimental and theoretical work using the CBS-QB3 complete basis set model chemistry (Montgomery et al. 2000). From that study, the existence of a relationship with glycolaldehyde through a dissociative process is suggested.

In Fig. 5 the PES corresponding to the formation of the P2 isomer from the two isomers of protonated hydroxylamine is schematized. If the reaction is initiated from  $\text{NH}_3\text{OH}^+$ , the mechanism starts with a concerted approach between both one of the hydrogen atoms bonded to nitrogen and the hydroxylic hydrogen of  $\text{NH}_3\text{OH}^+$  to the hydroxylic oxygen of  $\text{CH}_3\text{COOH}$  giving the I1P2 intermediate. This intermediate evolves towards the products formation ( $\text{P2} + \text{H}_2\text{O}$ ) through the TS1P2 transition state. In this transition state, the hydrogen atom of the hydroxyl group of  $\text{NH}_3\text{OH}^+$  is transferred to the hydroxylic oxygen of  $\text{CH}_3\text{COOH}$ , preforming a water molecule. Simultaneously, the highly reactive oxygen atom of the  $\text{NH}_3\text{O}$  group interacts with the carbonyl carbon atom of acetic acid generating the P2 product in a concerted process where a water molecule is released. The process can be schematized as follows:



As can be seen from Fig. 5, the reaction of formation of the P2 isomer, from the lowest-lying isomer of protonated hydroxylamine through the c1 pathway, is an exothermic process ( $\Delta E = -10.9 \text{ kcal mol}^{-1}$  at the CCSD(T) level) that involves a significant energy barrier since TS1P2 lies  $11.9 \text{ kcal mol}^{-1}$  (at the CCSD(T) level) than the reactants.



**Figure 5.** Energy profile, in  $\text{kcal mol}^{-1}$ , for the reactions of protonated hydroxylamine with acetic acid producing **P2** and **P3** computed at the MP2/aug-cc-pVTZ and CCSD(T)/aug-cc-pVTZ (in parentheses) levels of theory. Zero-point vibrational energy computed at the MP2/aug-cc-pVTZ level is included.

If we compare the b1 and c1 pathways, it can be seen that in the former an additional energy is required to release the hydroxylic hydrogen of  $\text{NH}_3\text{OH}^+$  whereas in the latter pathway a most favourable proton transfer occurs. This energy difference correlates with the energy associated to the following two processes (Jeanvoine et al. 2018):





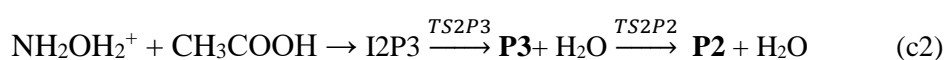


Another possibility to obtain the P2 isomer, from  $\text{NH}_3\text{OH}^+$ , implies firstly, the formation of the P3 product through the b1 pathway. Once P3 is obtained, secondly, the hydrogen atom migration from the oxygen atom to the nitrogen one, through the TS2P2 transition state, leads to the most stable isomer P2. This  $\text{P3} \rightarrow \text{P2}$  isomerization process implies an energy barrier of  $10.5 \text{ kcal mol}^{-1}$  (at the CCSD(T) level). The overall process can be summarized as:



The TS2P2 transition state is located  $2.1 \text{ kcal mol}^{-1}$  higher in energy than reactants, and production of the P3 isomer involves a large energy barrier ( $23752 \text{ K}$ ) thus, the d1 path will have the same characteristics as the c1 pathway.

The reaction of formation of the P2 product from the highest energy isomer of protonated hydroxylamine starts with the interaction between one of the hydrogen atoms bonded to nitrogen in  $\text{NH}_2\text{OH}_2^+$  and the carbonylic oxygen of acetic acid giving the I2P3 intermediate. This complex, following the b1 pathway evolves to the P3 product and then it isomerizes into the most stable P2 product through the TS2P2 transition state:



The process represented by the c2 pathway is clearly exothermic ( $\Delta E = -36.3 \text{ kcal mol}^{-1}$  at the CCSD(T) level) but has a significant activation barrier ( $21.8 \text{ kcal mol}^{-1}$  at the CCSD(T) level) and consequently it should not be relevant under interstellar conditions.

Our results show that the formation of P1,  $\text{CH}_3\text{CONH}_2^+\text{OH}$ , is the only favourable product from this reaction. It should be noted, that protonated complex organic molecules (COMs) have not been yet detected in ISM, it seems that they could evolve to neutral species by dissociative recombination. Therefore, the neutral isomer,  $\text{OHNHCOCH}_3$ , might be a candidate molecule to be searched for in the interstellar medium.

### 2.1.2 Structure and theoretical spectroscopic parameters of $[H_6C_2O_2N]^+$ and $[H_5C_2O_2N]$ isomers

Structural data for the  $[H_6C_2O_2N]^+$  isomers are given as an Appendix, in Table A1 (Section A). In Table 2, we provide the relevant spectroscopic parameters to rotational spectroscopy, together with computed dipole moment components along their principal inertial axes, for the P1-P8 isomers. Equilibrium rotational constants (A, B, C) were computed at the MP2/aug-cc-pVTZ level and the corresponding rotational constants for the ground vibrational state were calculated from the vibration–rotation coupling constants. We have also included in the table centrifugal distortion parameters in the symmetrically reduced Hamiltonian ( $\Delta J$ ,  $\Delta K$ ,  $\Delta_{JK}$ ,  $\delta_J$  and  $\delta_K$ ).  $[H_6C_2O_2N]^+$  isomers possess one nucleus with the quadrupole moment,  $^{14}N$  ( $I = 1$ ). Hence, each rotational transition of P1-P8 isomers carry the nuclear quadrupole hyperfine pattern expected for the presence of a  $^{14}N$  nucleus. Thus, in Table 2, the diagonal elements of the  $^{14}N$  nuclear quadrupole coupling tensor ( $X_{aa}$ ,  $X_{bb}$  and  $X_{cc}$ ) are given.

**Table 2.** Spectroscopic parameters for the P1 – P8 isomers computed at the MP2/aug-cc-pVTZ level of theory

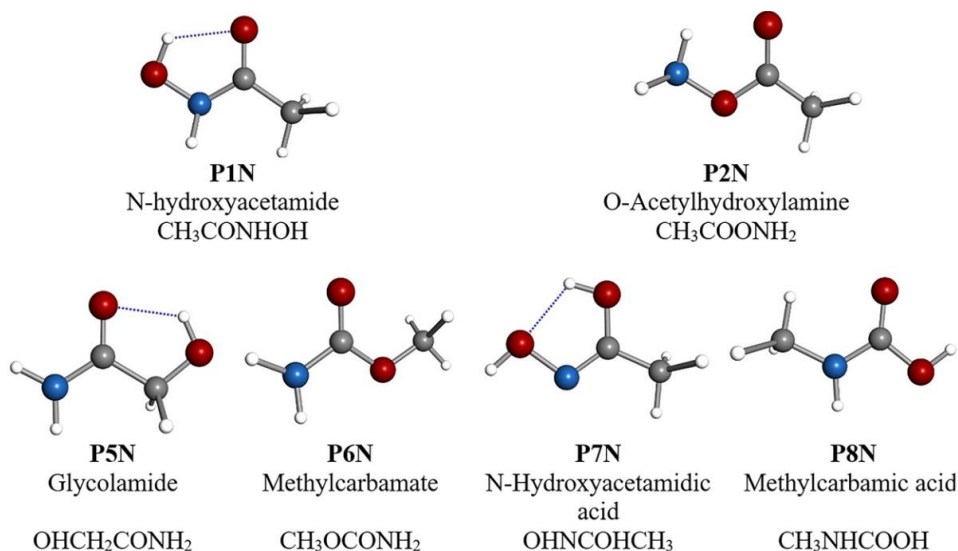
Parameter	P1	P2	P3	P4	P5	P6	P7	P8
A <sup>a</sup>	9767	10218	10478	10161	8524	9996	10140	10290
B	3858	4126	4162	4027	4192	4280	4043	4054
C	2859	3045	3085	2961	2906	3057	2957	29638
A <sub>0</sub> <sup>b</sup>	9652	10173	10419	10087	8423	9907	10046	10188
B <sub>0</sub>	3799	4086	4141	4000	4174	4230	4008	4020
C <sub>0</sub>	2819	3021	3068	2938	2896	3023	2930	2937
$\Delta_J$ <sup>c</sup>	0.90	0.72	0.61	0.60	0.72	0.77	0.64	0.68
$\Delta_K$	7.78	5.35	4.78	6.38	-27.02	2.02	6.57	2.00
$\Delta_{JK}$	2.11	4.37	5.23	3.88	34.59	3.67	3.24	4.87
$\delta_J$	0.23	0.19	0.16	0.16	0.18	0.22	0.16	0.18
$\delta_K$	0.98	-5.66	-4.16	2.13	-14.68	2.74	2.06	2.17
$\chi_{aa}$ <sup>d</sup>	1.60	0.26	5.40	-1.20	-0.43	1.02	6.10	2.04
$\chi_{bb}$ - $\chi_{cc}$	1.10	0.35	-3.44	1.79	-0.10	5.51	0.05	4.40
$ \mu_a  /  \mu_b $ /	1.4 / 3.8	4.7 / 1.4 /	1.5 / 0.7	0.7 / 2.1	1.3 / 3.2	0.8 / 1.7	1.4 / 1.8	1.4 / 0.4
$ \mu_c $ <sup>e</sup>	/ 0.0	0.0	/ 0.0	/ 1.1	/ 0.1	/ 0.0	/ 1.5	/ 0.0
P <sub>c</sub> <sup>f</sup>	3.046	3.020	2.919	2.208	3.84	1.655	1.957	1.624

<sup>a</sup> A, B, and C represent the rotational constants of the equilibrium structure (in MHz). <sup>b</sup> A<sub>0</sub>, B<sub>0</sub>, and C<sub>0</sub> are the rotational constants of the ground vibrational state, v=0 (in MHz). <sup>c</sup>  $\Delta_J$ ,  $\Delta_K$ ,  $\Delta_{JK}$ ,  $\delta_J$  and  $\delta_K$  represent the quartic centrifugal distortion constants (in KHz). <sup>d</sup>  $\chi_{aa}$ ,  $\chi_{bb}$  and  $\chi_{cc}$  represent the elements of the  $^{14}N$  nuclear quadrupole coupling tensor (in MHz). <sup>e</sup>  $|\mu_a|$ ,  $|\mu_b|$  and  $|\mu_c|$  are the absolute values of the electric dipole moment components (in D). <sup>f</sup> P<sub>c</sub> is the planar inertial moment (in u Å<sup>2</sup>). Conversion factor: 505379.1 MHz – 1 u Å<sup>2</sup>.

Planar moments of inertia,  $P_c$ , have been also included in Table 2. They can be obtained from rotational constants through the relation  $P_c = h/(16\pi^2)(-1/A_0 + 1/B_0 + 1/C_0)$ , and represent the mass distribution outside the  $ab$  inertial plane. As can be seen in Table 2, the planar moments for the P1-P8 isomers are, in all cases, relatively significant in magnitude and different from zero indicating that all P1-P8 isomers are not entirely rigid nor planar structures. On the other hand, the values of the computed dipole moments are high enough to allow for the observation of their corresponding rotational spectra using standard micro- and milli-meter-wave spectroscopy instruments.

All reaction products were predicted to be asymmetric tops. This characteristic together with the presence of one  $^{14}\text{N}$  nucleus with a quadrupole moment ( $I = 1$ ) results in a complex rotational spectrum.

To help in the possible identification of P1-P8 isomers in the gas phase through IR spectroscopy, on Table A2 in Section A.1 of Appendix A, we give their predicted MP2/aug-cc-pVTZ harmonic and anharmonic vibrational frequencies. For all the P1-P8 isomers, anharmonic frequencies are lower than the corresponding harmonic ones and the largest differences between harmonic and anharmonic frequencies were found, in general, in the frequencies corresponding to stretching modes.



**Figure 6.** Chemical Structures of  $[\text{H}_5\text{C}_2\text{O}_2\text{N}]$  isomers optimized at MP2/aug-cc-pVTZ level of theory

Once P1-P8 ionic species were studied, the next step was to obtain spectroscopic information for the corresponding neutral species denoted as P1N-P8N. All of them are structural isomers of glycine and have  $[\text{H}_5\text{C}_2\text{O}_2\text{N}]$  molecular formula. In this case, six

different chemical species were analysed, since both the P1N and P4N isomers and the P2N and P3N isomers converge, respectively, in the same neutral species (P1N = P4N, P2N = P3N). The six isomeric species studied together with neutral glycine are presented in Fig. 6.

Again, for each isomer a previous conformational analysis was performed. In Table 3, the relative energies for the most stable conformers of the  $[\text{H}_5\text{C}_2\text{O}_2\text{N}]$  isomers obtained at different levels of theory are given. As can be observed from the table, the stability order is not modified when going from the B3LYP level to the CCSD(T) one. The most stable structure is the P8N isomer (N-Methylcarbamic acid,  $\text{CH}_3\text{NHC}(\text{O})\text{OH}$ ) and the less stable one corresponds to the P7N structure (N-Hydroxyacetamidic acid,  $\text{CH}_3\text{C}(\text{O})\text{NHOH}$ ). It should be noted that, among neutral isomers, glycine (Gly) is not the most stable species, it is situated in stability between P6N (Methylcarbamate,  $\text{CH}_3\text{OC}(\text{O})\text{NH}_2$ ) and P5N (Glycolamide,  $\text{CH}_2\text{OHC}(\text{O})\text{NH}_2$ ) isomers. Regardless of the level of calculation employed, the stability order of the isomers considered in this work with  $[\text{H}_5\text{C}_2\text{O}_2\text{N}]$  molecular formula is (> means more stable than):



The energetic difference between the ends of the series is now lower than in the protonated counterpart. The high energy isomer, P7N, is  $45.5 \text{ kcal mol}^{-1}$  (at the CCSD(T) level) higher in energy than the most stable one P8N.

Lattelais et al. 2011 in their study of the isomers of glycine, found that the most stable isomer was not glycine, but N-methylcarbamic acid, (P8N). Following in energy were found methylcarbamate (P6N) and glycine (Gly) which are located  $4.9$  and  $8.8 \text{ kcal mol}^{-1}$ , respectively, higher in energy than the most stable isomer at the CCSD(T)/cc-pVQZ//B3LYP/6-311G(d,p) level.

**Table 3.** Relative energies, including zero point corrections, in kcal mol<sup>-1</sup>, for glycine (Gly) and its structural isomers.

Isomer	B3LYP <sup>a</sup>	MP2 <sup>b</sup>	CCSD <sup>c</sup>	CCSD(T) <sup>d</sup>
<b>P8N</b>	0.00	0.00	0.00	0.00
<b>P6N</b>	4.29	5.26	4.64	4.64
<b>Gly</b>	10.44	9.05	8.98	8.44
<b>P5N</b>	11.02	10.16	9.92	9.67
<b>P2N</b>	43.51	46.36	44.96	43.47
<b>P1N</b>	44.32	46.77	46.03	44.66
<b>P7N</b>	47.30	48.23	46.94	45.45

<sup>a</sup>Electronic energy calculated at the B3LYP/cc-pVTZ level. <sup>b</sup>Electronic energy calculated at the MP2/aug-cc-pVTZ level. <sup>c</sup>Electronic energy calculated at the CCSD/aug-cc-pVTZ//MP2/aug-cc-pVTZ level. <sup>d</sup>Electronic energy calculated at the CCSD(T)/aug-cc-pVTZ//MP2/aug-cc-pVTZ level.

Proton transfer processes are common in the interstellar medium, therefore we have computed proton affinities for the neutral isomers and are showed, at different level of theory, in Table 4. As can be seen from the Table, the values of proton affinities are relatively high if comparison is made to those of some abundant interstellar molecules. For example, the proton affinities of H<sub>2</sub>, CO, and C<sub>2</sub>H<sub>2</sub> are 100, 143, and 152 kcal mol<sup>-1</sup>, respectively. Therefore, neutral isomers should react quite easily in proton-rich interstellar media to give the corresponding protonated species. As supporting material, in Table A3 of Appendix A (Section A.1), we give geometrical parameters for the [H<sub>5</sub>C<sub>2</sub>O<sub>2</sub>N] isomers.

Rotational spectroscopic parameters for neutral isomers are given in Table 5. As in their protonated counterparts, all neutral isomers are asymmetric tops and their rotational spectrum should show the quadrupole hyperfine pattern that can be expected for those of molecules with a <sup>14</sup>N atom. Dipole moments values for the neutral isomers were high enough to allow for their characterization by micro- and millimeter-wave spectroscopy. To sum everything up, the present study should serve as a first step to detect these molecular systems in the laboratory. Finally, the calculated harmonic and anharmonic vibrational frequencies and IR intensities for the neutral isomers are given in Table A4 in Section A.1 of Appendix A. This information could be useful for their eventual detection in the gas phase through IR spectroscopy.

**Table 4.** Proton affinities, including zero point corrections, in kcal mol<sup>-1</sup>, for [H<sub>5</sub>C<sub>2</sub>O<sub>2</sub>N] isomers.

Reaction	B3LYP <sup>a</sup>	MP2 <sup>b</sup>	CCSD <sup>c</sup>	CCSD(T) <sup>d</sup>
<b>P8N + H<sup>+</sup> → P8</b>	201.37	196.44	200.67	199.15
<b>P6N + H<sup>+</sup> → P6</b>	203.35	198.22	202.35	200.80
<b>P5N + H<sup>+</sup> → P5</b>	192.60	189.87	192.69	191.97
<b>P2N + H<sup>+</sup> → P2</b>	202.66	199.80	201.66	200.71
<b>P2N + H<sup>+</sup> → P3b</b>	200.67	195.53	199.59	198.14
<b>P1N + H<sup>+</sup> → P4d</b>	206.31	201.12	205.89	203.78
<b>P1N + H<sup>+</sup> → P1b</b>	189.40	187.46	190.33	189.43
<b>P7N + H<sup>+</sup> → P7a</b>	181.95	175.21	180.86	178.45

<sup>a</sup>Electronic energy calculated at the B3LYP/cc-pVTZ level. <sup>b</sup>Electronic energy calculated at the MP2/aug-cc-pVTZ level. <sup>c</sup>Electronic energy calculated at the CCSD/aug-cc-pVTZ//MP2/aug-cc-pVTZ level. <sup>d</sup>Electronic energy calculated at the CCSD(T)/aug-cc-pVTZ//MP2/aug-cc-pVTZ level

**Table 5.** Spectroscopic parameters for the [H<sub>5</sub>C<sub>2</sub>O<sub>2</sub>N] isomers computed at the MP2/aug-cc-pVTZ level of theory.

Parameter	P1N	P2N	P5N	P6N	P7N	P8N
A <sup>a</sup>	10600	10465	10449	10673	10114	10695
B	4140	4236	4077	4429	4133	4109
C	3045	3102	2987	3196	2988	3025
A <sub>0</sub> <sup>b</sup>	10529	10383	10361	10593	10012	10625
B <sub>0</sub>	4099	4193	4015	4389	411	4063
C <sub>0</sub>	3015	3073	2955	3164	2962	2989
Δ <sub>J</sub> <sup>c</sup>	0.72	0.76	0.76	0.77	0.67	0.65
Δ <sub>K</sub>	7.73	6.66	2.29	3.83	5.80	-5.15
Δ <sub>JK</sub>	3.22	4.14	6.31	3.05	5.21	12.65
δ <sub>J</sub>	0.19	0.21	0.20	0.21	0.17	0.17
δ <sub>K</sub>	1.62	-0.29	2.050	2.39	3.16	-16.2
X <sub>a</sub> <sup>d</sup>	4.04	4.39	5.93	2.21	5.06	2.55
X <sub>b</sub> - X <sub>c</sub>	5.70	3.83	0.09	6.07	-3.55	6.59
μ <sub>a</sub>   /  μ <sub>b</sub>   /  μ <sub>c</sub>   <sup>e</sup>	1.7 / 2.7 / 0.8	0.8 / 1.1 / 1.3	3.2 / 3.0 / 0.0	0.1 / 2.2 / 0.6	2.5 / 0.0 / 0.0	1.2 / 2.2 / 0.0
P <sub>c</sub> <sup>f</sup>	1.835	2.373	1.812	1.564	1.365	1.436

<sup>a</sup>A, B, and C represent the rotational constants of the equilibrium structure (in MHz). <sup>b</sup>A<sub>0</sub>, B<sub>0</sub>, and C<sub>0</sub> are the rotational constants of the ground vibrational state, v=0 (in MHz). <sup>c</sup>Δ<sub>J</sub>, Δ<sub>K</sub>, Δ<sub>JK</sub>, δ<sub>J</sub> and δ<sub>K</sub> represent the quartic centrifugal distortion constants (in KHz). <sup>d</sup>X<sub>a</sub>, X<sub>b</sub> and X<sub>c</sub> represent the elements of the <sup>14</sup>N nuclear quadrupole coupling tensor (in MHz). <sup>e</sup>|μ<sub>a</sub>|, |μ<sub>b</sub>| and |μ<sub>c</sub>| are the absolute values of the electric dipole moment components (in D). <sup>f</sup>P<sub>c</sub> is the planar inertial moment (in u Å<sup>2</sup>). Conversion factor: 505379.1 MHz – 1 u Å<sup>2</sup>.

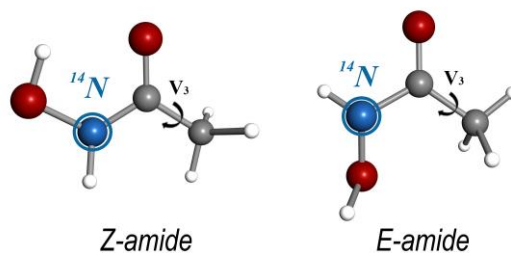
## 2.2 *Experimental rotational study of acetohydroxamic acid*

### 2.2.1 *Jet-cooled broadband spectrum*

After we found a feasible process to form protonated N-hydroxyacetamide under interstellar conditions, we suggested the neutral counterpart, N-hydroxyacetamide, also known as acetohydroxamic acid (for now on AHA), as a candidate for interstellar detection and recorded its rotational spectrum. Acetohydroxamic acid is a highly hygroscopic solid with a melting point of 88-90 °C and very low vapor pressure, standing as a challenging problem for high-resolution rotational spectroscopy. However, our laser ablation technique (see Chapter III) enables an ultra-fast vaporization process of organic and biological molecules from the solid state to the gas phase to further probe them in the isolated conditions of a supersonic jet (Alonso et al. 2014, Alonso & López 2015). Thus, the experimental approach followed in this study allowed us to record the rotational spectra of this glycine isomer for the first time.

AHA is a relatively flexible molecule that exhibits *Z*(*zusammen, cis*)-*E*(*entgegen, trans*) isomerization due to a rotation around the C-N bond (shown in Figure 6). In the gas phase the *Z* and *E* forms are of relatively similar stability, but their interconversion is hindered by barriers greater than 25 kcal/mol (Saldyka & Mielke 2007).

Before starting the experimental study and to guide the spectral search, we extended the previous theoretical calculations (Section 2.1 of the present Chapter) to the low energy conformers of AHA (see Table A2.1 of Appendix A, Section A.2). The presence of a <sup>14</sup>N nucleus in our molecule is expected to be a helpful tool to identify structures, since it introduces hyperfine rotational probes at defined sites of AHA, depending on the chemical environment of the <sup>14</sup>N nucleus. Thus, the treatment and interpretation of the quadrupole hyperfine structure of AHA is needed to unveil, conclusively, the observed species. Moreover, for each conformer we computed a relaxed potential-energy scan, choosing the HCCO torsional angle as the driving coordinate to study the methyl internal rotation. Note that intermediate barrier height values were predicted (see Table 6), pointing to a splitting in the same order of magnitude as that expected for the <sup>14</sup>N quadrupole hyperfine structure.



**Figure 6.** Structures of aceto-hydroxamic acid conformers, showing the presence of a  $^{14}\text{N}$  nucleus.

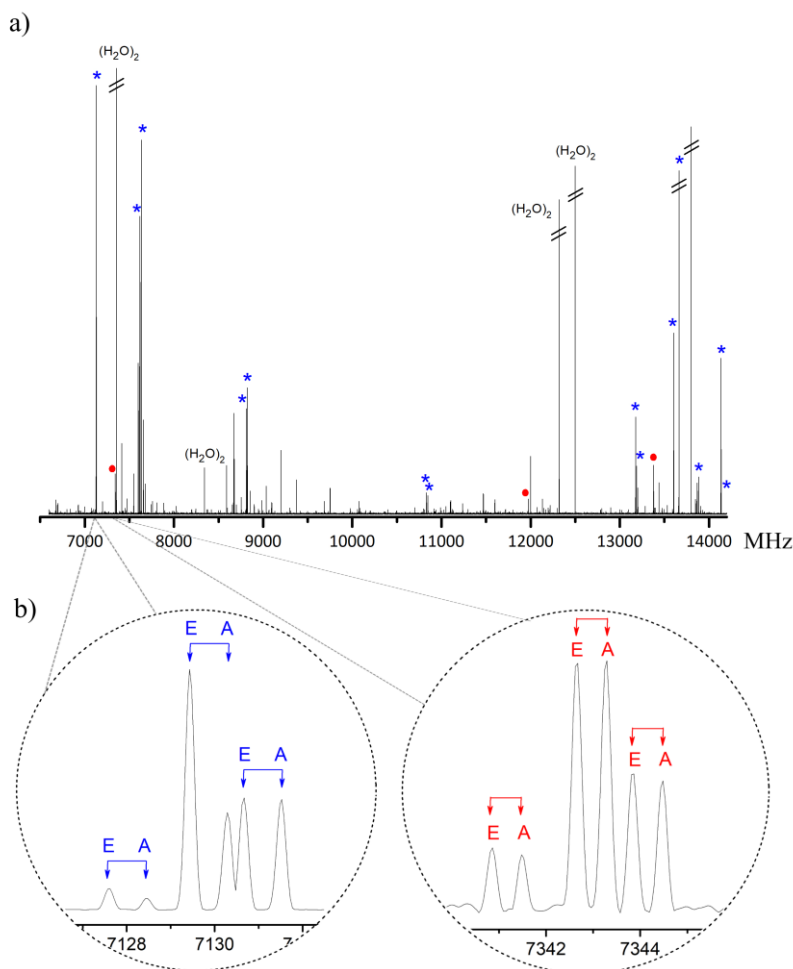
The broadband spectrum of AHA from 6.8 to 14.2 GHz is shown in Figure 7 (a). According to the predictions of Table 6 the two plausible amide conformers are expected to be near-prolate asymmetric tops with a nonzero  $\mu_a$  and  $\mu_b$  component of the electric dipole moment. In this context, their  $R$ -branch,  $\mu_a$ -type spectra are expected to show distinguishing arrangements of individual  $K_a$  transitions spaced approximately  $B + C$ . Furthermore, all the observed transitions were split into several close hyperfine components showing the characteristic pattern due to a  $^{14}\text{N}$  nucleus, which confirms the presence of a single nitrogen atom in the observed species. Moreover, additional splittings of several MHz were observed in most of the transitions, which significantly complicated the data analysis (an illustration is given in Figure 7(b)). Given that AHA is a closed shell molecule, no other fine or hyperfine structure in the rotational spectra except that arising from the coupling between the internal and overall rotation is expected to occur. Thus, we attributed the splitting to the internal rotation of the methyl group causing the occurrence of the  $A$ – $E$  doublets (Cabezas et al. 2012). Taking this into account, we were able to detect two sets of strong  $\mu_a$ -type  $R$ -branch progressions  $(J+1)_0, J+1 \leftarrow J_{0, J}$  and  $(J+1)_1, J+1 \leftarrow J_{1, J}$ , corresponding to the  $A$ -symmetry state of rotamer **I** (an illustration is given in Figure 7b). Afterwards, employing the XIAM program (Hartwig and Dreizler 1996, Kolesnikova 2017),  $E$  transitions were predicted using the theory calculated barrier to the internal rotation  $V_3$ . The subsequent iterative fitting and prediction procedure allowed us to observe new  $a$ -type transitions and to extend our measurements to the  $R$ -branch  $b$ -type spectrum. Notice,  $c$ -type lines were also predicted but not observed. Although the resolution of our broadband CP-FTMW technique was sufficient to partly resolve the  $^{14}\text{N}$  nuclear quadrupole hyperfine patterns for some transitions (see Figure 6 (b)), in a first step, we did not struggle to analyze the hyperfine structure. Instead, we measured and fitted (Pickett 1991) the intensity-weighted mean of the hyperfine line



cluster to a rigid rotor Hamiltonian to generate a preliminary set of rotational constants ( $A = 10642.7$ ,  $B = 4106.3$  and  $C = 3024.2$ ) (in MHz) for rotamer **I**.

After excluding the lines belonging to rotamer **I**, another weaker progression of  $aR$ -branch transitions was discovered and assigned to another rotamer, labeled as **II**. The assignments were confirmed by predicting and measuring new  $a$ - and  $b$ -type lines (additional rotational lines), which are also subjected in some cases to a different  $A$ - $E$  splitting. After the corresponding fitting, a first set of constants ( $A = 8951.5$ ,  $B = 4320.7$  and  $C = 3022.8$ ) (in MHz) were derived from a rigid rotor analysis.

The fundamental basis for the identification of the observed rotamers is derived from the values of the rotational constants. The comparison of the experimental and the predicted rotational parameters in Table 6 shows that the observed rotational constants are in good agreement with those calculated at the B2PLYPD3 level for conformers **Z** and **E**. Moreover, the changes in the experimental values of the rotational constants between rotamers **I** and **II** are  $\Delta A = \Delta A_{\mathbf{I}} - \Delta A_{\mathbf{II}} \approx 1691.1$  MHz,  $\Delta B \approx -214.4$  MHz and  $\Delta C \approx 1.4$  MHz. These values are in good agreement with the differences between the rotational constants of the amide conformers **Z** and **E**;  $\Delta A = \Delta A_{\mathbf{Z}} - \Delta A_{\mathbf{E}} \approx 1646.4$  MHz,  $\Delta B \approx -205.0$  MHz, and  $\Delta C \approx 5.8$  MHz. This fact enables the identification of rotamers **I** and **II** as the lowest-lying energy lowest lying energy **Z** and **E** AHA conformers, respectively. In conclusion, using the abovementioned arguments, an unambiguous conformational characterization of AHA was achieved.

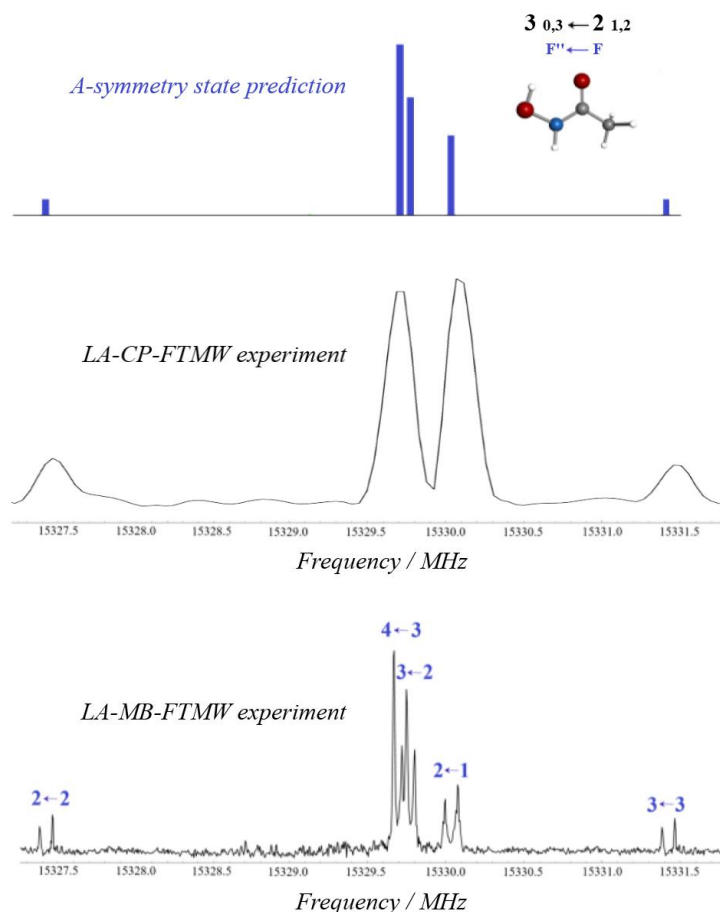


**Figure 7.** (a) Broadband LA-CP-FTMW spectrum of acetohydroxamic acid from 6.8 to 14.2 GHz. We depict strong *a*-type R-branch  $(J+1)_{0J+1} \leftarrow J_{1J}$  and  $(J+1)_{1J+1} \leftarrow J_{0J}$  transitions of rotamer **I** (**Z**-AHA, in blue) as well as some transitions of rotamer **II** (**E**-AHA, in red). b) Zoom-in on the broadband spectrum showing the  $1_{01} \leftarrow 0_{00}$  rotational transition for both acetohydroxamic acid species showing the *A*–*E* doublets because of the internal rotation of the methyl group. The intensity is given in arbitrary units.

### 2.2.2 *Fine and hyperfine structure: narrowband rotational spectrum*

A more puzzling task was the resolution of the fine and hyperfine structure of acetohydroxamic acid. First, the *A*–*E* splitted structure attributed to the methyl internal rotation motion can be generally used as a molecular “fingerprint” to search for a molecule in an astronomical line survey (Cernicharo et al. 2016, Belloche et al. 2017). Moreover, it is known that for interstellar searches using low-frequency surveys, i.e., those conducted toward cold dark molecular clouds such as TMC-1, the interpretation of the hyperfine patterns could be a critical drawback for deciphering the radio astronomical observations (McCarthy & McGuire 2021). As an example, the analysis of the hyperfine

structure of several N-bearing species, such as N-protonated isocyanic acid ( $\text{H}_2\text{NCO}^+$ , Rodríguez-Almeida et al. 2021) which has been detected towards G+0693-0.027 along with  $\text{C}_2\text{H}_5\text{NCO}$ , was required for its conclusive line-by-line identification. Also, these low-frequency datasets are lately suggested among the best targets to search for large molecular systems (McGuire et al. 2021), further supporting centimeter-wave studies of yet unknown interstellar candidates. Hence, we employed our narrowband LA-MB-FTMW spectrometer (Bermúdez et al. 2014) to completely resolve both fine and hyperfine structures. Initially, we used the SPFIT/SPCAT programme package (Pickett 1991) to analyze hyperfine patterns for the previously measured *A* and *E* lines of *Z*-AHA (see Figure 8 for an example). This scrutiny followed by measuring additional hyperfine components of several *a*- and *b*-type *R*-branch lines. Once we completed the analysis, new *R*-branch transitions belonging to the *E*-conformer were identified.



**Figure 8.** Partially resolved hyperfine pattern for the  $3_{0,3} \leftarrow 2_{1,2}$  transition of the *A*-symmetry ground state of *Z*-AHA (LA-CP-FTMW experiment) and completely resolved hyperfine structure using the LA-MB-FTMW experiment, where each transition appears as Doppler doublets due to the sub-Doppler resolution of the spectrometer. The simulated spectrum is also depicted for comparison (in blue). The intensity is given in arbitrary units.

**Table 6.** Theoretical and experimental spectroscopic parameters for the *A*-symmetry state of the observed conformers of AHA.

	Experiment		Theory <sup>[a]</sup>	
	Rotamer I (DR)	Rotamer II	Z-AHA	E-AHA
<i>A</i> / MHz <sup>[b]</sup>	10642.6888 (21) <sup>[j]</sup>	8951.527 (6)	10634.1	8987.7
<i>B</i> / MHz	4106.2983 (12)	4320.672 (25)	4114.3	4319.3
<i>C</i> / MHz	3024.1996 (10)	3022.795 (20)	3032.8	3027.0
$\Delta_J$ / kHz <sup>[c]</sup>	0.667 (48)	-	-	-
$\Delta_{JK}$ / kHz	5.16 (35)	-	-	-
$\delta_J$ / kHz	0.173 (49)	-	-	-
$\mu_a / \mu_b / \mu_c / D$ <sup>[d]</sup>	Yes / Yes / No	Yes / Yes / No	1.6 / 2.8 / 0.7	2.9 / 1.2 / 0.8
$\chi_{aa}$ / MHz <sup>[e]</sup>	4.0872 (33)	3.973 (10)	4.197	4.131
$\chi_{bb}$ / MHz	1.0959 (50)	1.812 (17)	1.027	1.873
$\chi_{cc}$ / MHz	-5.1831 (50)	-5.785 (17)	-5.225	-6.003
N. A lines <sup>[f]</sup>	65	18	-	-
N. E lines	33	14	-	-
$V_3$ / cm <sup>-1</sup> <sup>[g]</sup>	255.7 (5)	413 (81)	250.8	403.2
$\sigma$ / kHz <sup>[h]</sup>	8.1	10.8	-	-
$\Delta E_{TOT}$ / cm <sup>-1</sup> <sup>[i]</sup>	-	-	0/0	772.8

[a] Theoretical calculations at B2PLYPD3/aug-cc-pVTZ level of theory. [b] *A*, *B*, and *C* represent the rotational constants (in MHz); [c]  $\Delta_J$  is the quartic centrifugal distortion constant (in KHz); [d]  $\mu_a$ ,  $\mu_b$ , and  $\mu_c$  are the electric dipole moment components (in D). [e]  $\chi_{aa}$ ,  $\chi_{bb}$  and  $\chi_{cc}$  are the diagonal elements of the <sup>14</sup>N nuclear quadrupole coupling tensor (in MHz), [f] N is the number of hyperfine components; [g]  $V_3$  is the barrier to the methyl internal rotation motion (in cm<sup>-1</sup>); [h]  $\sigma$  is the root mean square (rms) deviation of the fit (in kHz); [i]  $\Delta E_{TOTAL}$  is the electronic energy (in cm<sup>-1</sup>), taking into account the zero-point vibrational energy (ZPE) computed at the same level. [j] Standard error in parentheses in units of the last digit.

In a quest to extend the measurements of the *Z*-form (global minimum) to higher frequencies and allow for an eventual direct comparison with the Yebes 40-m and Effelsberg 100-m observations or with the GBT line surveys, we used our double-resonance (DR) configuration, which also enables the resolution of the hyperfine components (see additional details in Chapter III: Methodology). Subsequently, we

performed a fit (Pickett 1991) including our broadband and narrowband microwave data as well as the double resonance lines of the *A* symmetry state. We used the *A*-reduction Hamiltonian in the *I'* representation (Watson 1977), supplemented with an additional term that accounts for the nuclear quadrupole coupling interactions (Foley 1947; Robinson & Cornwell 1953). The Hamiltonian was set up in the coupled basis set ( $F = J + I$ ) to label the energy levels involved in each rotational transition in terms of the quantum numbers  $J$ ,  $K_a$ ,  $K_c$ , and  $F$ . The effective derived spectroscopic parameters for *Z*-AHA, including the  $^{14}\text{N}$  quadrupole coupling constants are listed in the first column of Table 6. We also provide a complete list included all the measured lines for *Z*- and *E*-amide AHA in Tables A2.1-A2.2 of Section A.2 (Appendix A)

Furthermore, the same sets of *A* and *E* lines were used to experimentally determine the internal rotation barrier  $V_3$  for both conformers using the inertial axis method, which is already implemented in the XIAM programme (Hartwig and Dreizler 1996, Kolesnikova 2017). *Z*-AHA presents a barrier height of  $255.4(4) \text{ cm}^{-1}$ , which is in great agreement with the theoretically predicted data (see Table 6). As appeared for the *Z*-amide form, *E*-AHA also exhibits a splitting due to the methyl internal rotation, which is in this case a bit more hindered due to the crowded environment of the methyl group ( $413(19) \text{ cm}^{-1}$ ). Nevertheless, it should be noted that in the XIAM fit we were not able to include all the measured lines since some of the showed an unresolved *A-E* splitting.

Finally, in Table 7 we provide the rotational ( $Q_r$ ) and vibrational ( $Q_v$ ) partition functions of the *A*-symmetry state of *Z*-acetohydroxamic acid. We used SPCAT (Pickett 1991) to compute the values of  $Q_r$  from first principles at the typical temperatures as applied in the JPL database (Pickett et al. 1998). The vibrational part,  $Q_v$ , was estimated employing a harmonic approximation and Eq.~3.60 of Gordy (1970) where all the vibrational modes below  $1000 \text{ cm}^{-1}$  were taken into account. The frequencies of the normal modes were computed using B2PLYPD3 double-hybrid calculations. Thus, the total partition function,  $Q_{tot}$  can be calculated as the product of  $Q_r$  and  $Q_v$ .

**Table 7.** Rotational and vibrational partition functions of the A-symmetry state of Z-acetohydroxamic acid

Temperature (K)	$Q_r^{(a)}$	$Q_v^{(b)}$
9.38	1269.4024	1.0000
18.75	3581.2976	1.0003
37.50	10117.2564	1.0178
75.00	28602.4865	1.1682
150.00	80902.3854	1.9489
225.00	148663.4611	3.8325
300.00	228949.0677	8.0398

**Notes.**<sup>(a)</sup>  $Q_r$  is the rotational partition function. It does not take the hyperfine splitting into account. We used  $J = 150$  as the maximum  $J$  value <sup>(b)</sup>  $Q_v$  is the vibrational partition function. The total partition function of the molecule (with the hyperfine splitting taken into account) is  $Q_r \times Q_v$ .

To sum everything up, we have investigated the rotational spectrum of acetohydroxamic acid in the microwave region (up to 40 GHz). We have provided very accurate laboratory data as a preliminary step to search for the molecule toward different regions of the ISM. A line catalogue including the A-symmetry state and also the  $^{14}\text{N}$  hyperfine structure has been generated to be eventually compared with observational datasets such as the Green Bank Observatory (GBT) surveys (McCarthy et al. 2021), the Yebes 40-m QUIJOTE line survey (Cernicharo et al. 2021) as well as the Effelsberg 100-m observations (see Section 2.5 of this Chapter).

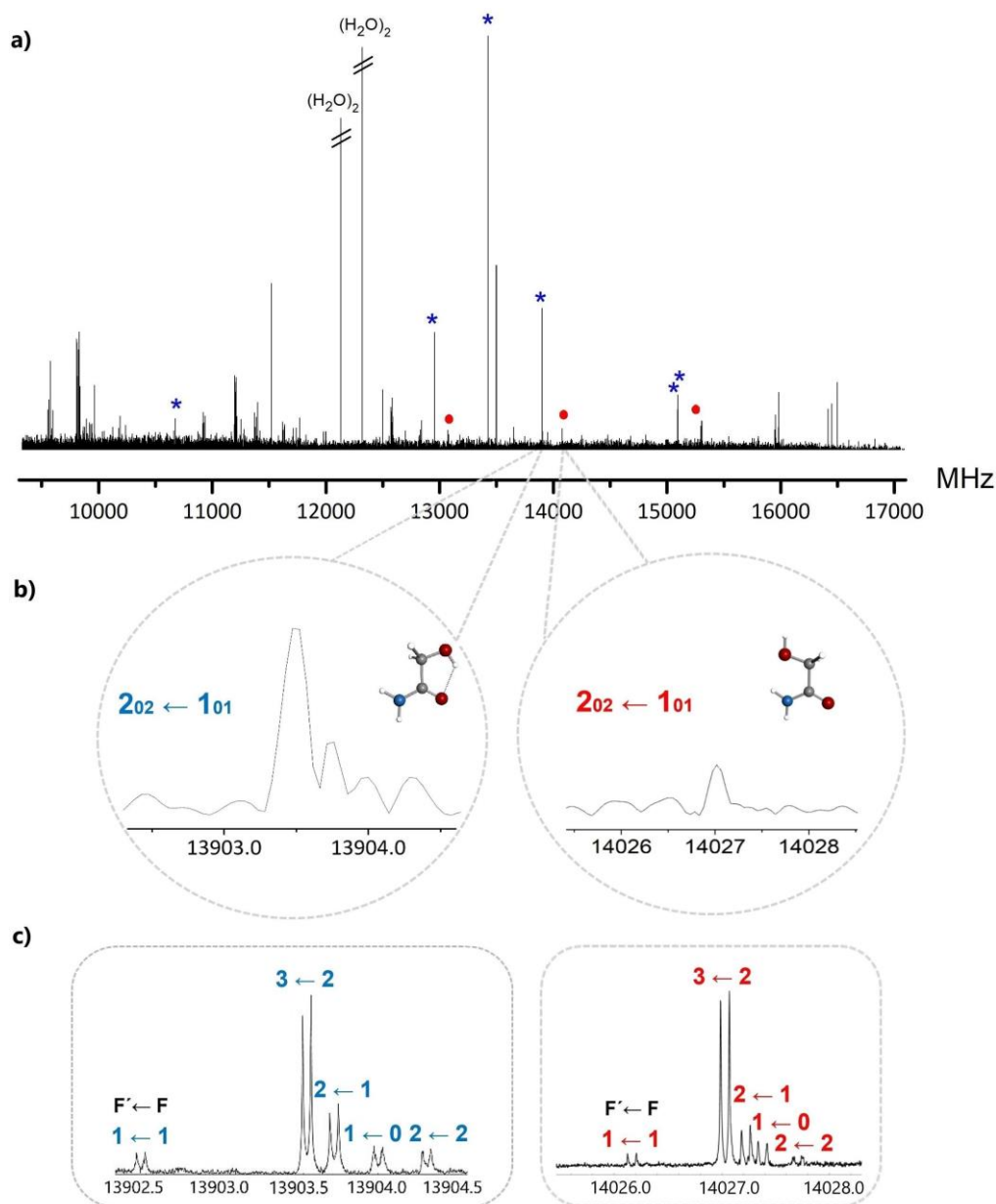
Regarding a future millimeter-wave study of acetohydroxamic acid, which should be the following step in our study, this molecule does not have much vapor pressure (a laser ablation procedure laser ablation was used to transfer the solid into the gas phase). Therefore, a conventional millimeter-wave investigation has not been accessible to date, even with the new heating cell (see Chapter III: methodology). Nevertheless, this experimental microwave data will enable a confident search of acetohydroxamic acid toward low frequency surveys, which are lately suggested as the best datasets to look for large molecules (McGuire et al. 2021).

## 2.3 *Experimental rotational study of glycolamide*

### 2.3.1 *Jet-cooled rotational spectrum of glycolamide*

Once we finished the experimental study of acetohydroxamic acid we proceeded to record the microwave spectrum of glycolamide, our second target molecule. The CP-FTMW rotational spectrum of glycolamide is shown in Fig. 9a. It appears dominated by intense water dimer (H<sub>2</sub>O)<sub>2</sub> signals and other lines belonging to known decomposition products that were identified and discarded from the analysis. At the first stage of the line assignment, predictions based on the spectroscopic constants from Maris (2004) were used. Initial inspection of the broadband spectrum was directed to search for a set of intense *R*-branch  $\mu_a$ -type rotational transitions (with  $J' \leftarrow J'' = 2 \leftarrow 1$ ) arising from *syn*-glycolamide predicted with significant  $\mu_a$  electric dipole-moment component. The initial assignment was extended to other intense *b*-type transitions. After an iterative procedure of fittings and predictions, a total of six rotational transitions were measured. They were least-squares fitted to a rigid rotor Hamiltonian (Pickett 1991). After excluding these lines from the spectral analysis, another weaker progression of <sup>a</sup>*R*-branch transitions was discovered and assigned to *anti*-glycolamide.

Rotational transitions exhibit the typical hyperfine structure arising from the interaction of the electric quadrupole moment of the <sup>14</sup>N ( $I = 1$ ) nucleus with the electric field gradient created at the site of the quadrupolar nucleus by the rest of the electronic and nuclear charges of the molecule. As shown in Fig. 1b, the resolution attained with our broadband CP-FTMW technique was not sufficient to fully resolve the hyperfine patterns. At this point, we took advantage of the high resolution of our narrowband MB-FTMW spectrometer. The analysis began with the measurement of a total of 5 hyperfine components by interpreting the quadrupole coupling pattern for the  $2_{02} \leftarrow 1_{01}$  rotational transition (see Fig. 9c) and then extended to other *a*- and *b*-type *R*-branch transitions. Once the analysis of *syn*-glycolamide was completed, new <sup>b</sup>*R*-type transitions belonging to *anti*-glycolamide were identified.



**Fig. 9.** (a) Broadband CP-FTMW spectrum of glycolamide from 9.2 to 17.1 GHz. We depict strong *a*- and *b*- type *R*-branch transitions of *syn*-glycolamide (in blue) as well as some transitions of *anti*-glycolamide (in red). (b) Zoom-in section of the broadband spectrum showing the rotational transition  $2_{02} \leftarrow 1_{01}$  of both *syn*-glycolamide and *anti*-glycolamide. (c) The same rotational transition measured in the MB-FTMW experiment, displaying the high resolution of our spectrometer. The resonance frequency is determined by the arithmetic mean of two Doppler components. The MB-FTMW spectrum was obtained by averaging 200 experimental cycles (four free induction decay signals per cycle).



**Table 8.** Experimental spectroscopic parameters for *syn*- and *anti*- glycolamide.

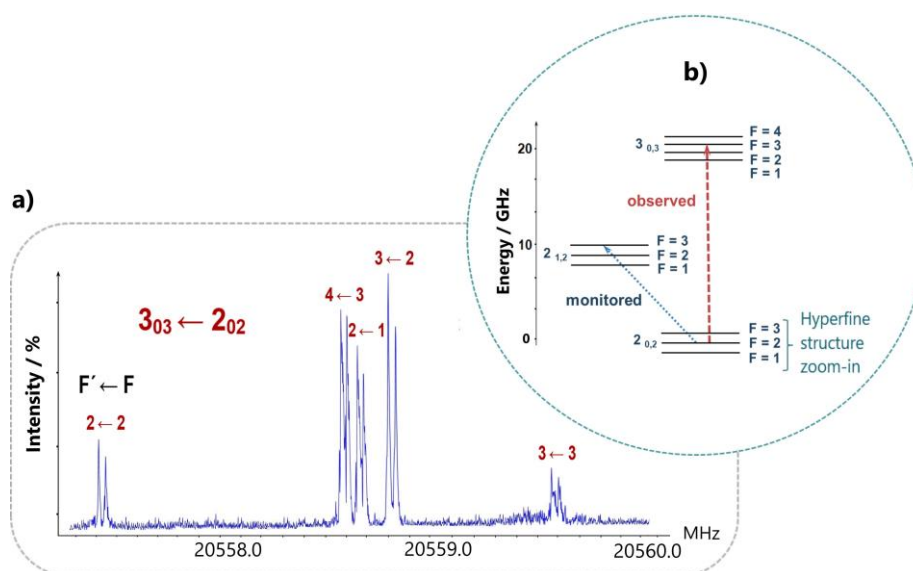
Parameters	<i>Syn</i> (DR)	<i>anti</i>
A <sup>[a]</sup> / MHz	10454.2640 (48) <sup>[e]</sup>	9852.6242 (22)
B / MHz	4041.15663 (24)	4125.1664 (15)
C / MHz	2972.07394 (20)	2967.57976 (75)
$\chi_{aa}$ <sup>[b]</sup>	2.050 (4)	1.618 (5)
$\chi_{bb}$	1.930 (6)	2.020 (10)
$\chi_{cc}$	-3.981 (6)	-3.638 (10)
$\Delta_J$ / kHz	0.7723 (59)	0.700 (38)
$\Delta_K$ / kHz	2.224 (74)	3.45 (28)
$\Delta_{JK}$ / kHz	6.98 (46)	3.3 (8)
$\delta_j$ / kHz	0.2061 (21)	0.27 (2)
$\delta_k$ / kHz	2.082 (65)	-
N <sup>[c]</sup>	97	46
$\sigma$ <sup>[d]</sup>	21.3	15.6

**Notes.** [a] *A*, *B*, and *C* represent the rotational constants; [b]  $\chi_{aa}$ ,  $\chi_{bb}$  and  $\chi_{cc}$  are the diagonal elements of the  $^{14}\text{N}$  nuclear quadrupole coupling tensor. [c] *N* is the number of measured transitions. [d]  $\sigma$  is the root mean square (rms) deviation of the fit (in kHz). [e] Standard error in parentheses in units of the last digit.

Contrary to *syn*-glycolamide, *anti*-glycolamide exhibits a weak *b*-type rotational spectrum and we were only able to measure *b*-type lines during the MB-FTMW experiment. After iterative fittings and predictions, a set of low-*J* rotational transitions along with a selection of well-resolved nuclear hyperfine components from the previous Stark study (Maris 2004) was adequately weighted and fitted (Pickett 1991) to a semirigid-rotor Hamiltonian in the *A*-reduction and the *F* representation  $H_R^{(A)}$  (Watson 1977), supplemented with a term to take into account the quadrupole interaction  $H_Q$  (Foley 1947; Robinson & Cornwell 1953), namely  $H = H_R^{(A)} + H_Q$ . The Hamiltonian was set up in the coupled basis set  $I + J = F$ , so the energy levels involved in each transition are labeled with the quantum numbers *J*,  $K_a$ ,  $K_c$ , and *F*. The final set of rotational constants and nuclear quadrupole coupling constants are listed in the second column of Table 8.

Seeking for an extension of our measurements to higher frequencies, we used the abovementioned double-resonance technique with which also hyperfine components can be resolved. An example of the double resonance spectrum is shown in Fig. 10, where the *b*-type transition,  $2_{12} \leftarrow 2_{02}$  is monitored and five hyperfine components of the *a*-type

transition  $3_{03} \leftarrow 2_{02}$ , which is connected to the lower level of the millimeter-wave transition, have been observed. Subsequently, we performed a fit (Pickett 1991) including our microwave and double resonance (DR) lines as well as a selection of transitions from Maris (2004). We used the same Hamiltonian mentioned above. The derived spectroscopic parameters for *syn*-glycolamide are listed in the first column of Table 1. Also, in Table A3.1 of Appendix A (Section A.3) we provide a comparison between the experimental and *ab initio* predicted quadrupole coupling constants.

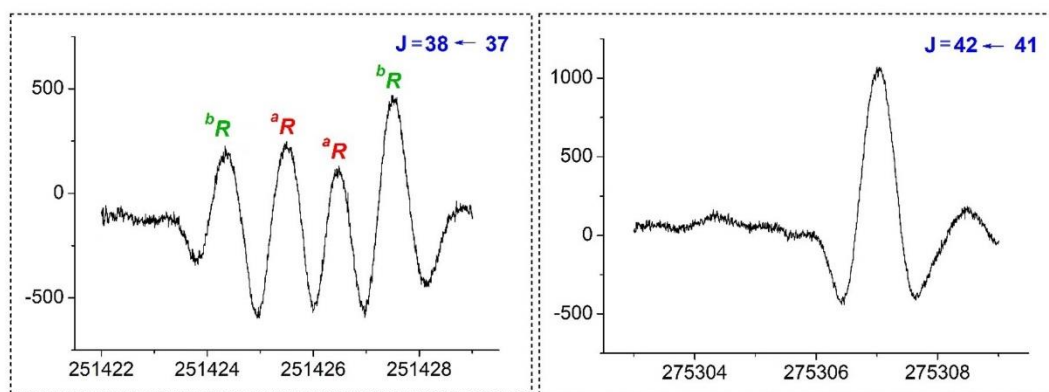


**Fig. 10.** (a) Double-resonance spectrum of the  $3_{03} \leftarrow 2_{02}$  rotational transition of *syn*-glycolamide showing five different hyperfine components. The  $2 \leftarrow 2$  and  $3 \leftarrow 2$  lines were observed monitoring the hyperfine component  $F = 2$  of the rotational transition  $2_{12} \leftarrow 2_{02}$ , while the  $2 \leftarrow 1$  was observed monitoring the hyperfine component  $F = 1$  component, and both  $4 \leftarrow 3$  and  $3 \leftarrow 3$  were observed monitoring the  $F = 3$  component of the same rotational transition; (b) Schematic energy level diagram (in GHz) relative to the  $2_{02}$  energy level of *syn*-glycolamide from the determined molecular constants. The hyperfine splittings are shown. Arrows indicate an example of the observed (dashed arrows in garnet) and monitored (dotted arrows in blue) transitions.

### 2.3.2 Millimeter and submillimeter-wave spectra of glycolamide

Afterwards, predictions based on the spectroscopic constants from the high-resolution data of the previous section were used. Groups of *a*-type *R*-branch transitions dominate the millimeter-wave spectrum and can be easily identified due to their typical  $(B+C)$  periodicity. These groups are interspersed by *b*-type *R*-branch transitions that do not cluster into any characteristic pattern. However, as *J* increases, the rotational energy levels

with the lowest  $K_a$  quantum numbers become nearly degenerate, and pairs of  $b$ -type with corresponding  $a$ -type transitions involving these energy levels form one quadruply degenerate line. For instance, in Fig. 11 we depict the coalescence for the quartets corresponding to  $K_a = 4$  and 5. Taking this into account, intense  $^aR$ -branch transitions involving  $K_a = 0, 1$  energy levels were searched for at first. Subsequently, higher  $K_a$   $R$ -branch  $a$ - and  $b$ -type transitions together with several  $Q$ -branch transitions were assigned. However, it is worth noticing that discrepancies up to 60 MHz were found at first between the experimental  $Q$ -branch and also the high  $K_a$  rotational transitions compared to initial predictions, emphasizing once more the primary interest of this study.



**Figure 11.** Pairs of  $a$ - and  $b$ -type transitions of syn-glycolamide with  $K_a = 4, 5$  at  $J'' = 37$  (left panel) that coalesce into a quadruply degenerate line at  $J'' = 41$  (right panel) during the mmw experiment while going up in frequency ( $b$ -type transitions colored in green and  $a$ -type transitions in red, respectively). The broadening is higher than that of conventional room-temperature experiments due to the effect of the temperature of the cell. The  $x$  axis is labeled in MHz

Finally, 1480 new transition lines were measured (together with the previous MW and DR-MW measurements correctly weighted) and the ranges of  $J$  and  $K_a$  quantum numbers were extended up to 51 and 33, respectively. The fits and predictions were made in terms of Watson's  $A$ -reduced Hamiltonian in  $I'$ -representation with the Pickett's SPFIT/SPCAT program suite (Pickett 1991). The presently available data allowed the determination of accurate rotational constants, the full set of quartic and almost all the sextic centrifugal distortion constants. What is more, despite the broadening due to the temperature effect of the cell, the RMS deviation of the fit is considerably small (around 50 kHz). The spectroscopic constants are presented in Table 9, and the measured transitions are reported in Table 11. The final spectroscopic set decisively improves the

previous free-jet Stark modulated experimental data by up to three orders of magnitude (Maris 2004). In conclusion, the millimeter and submillimeter-wave spectrum of the most stable conformer of glycolamide has been studied up to 460 GHz to provide accurate laboratory reference spectra that are needed to compare directly against the observational data in different regions of the ISM.

**Table 9.** Ground-state spectroscopic constants of Glycolamide (A-Reduction, I<sup>r</sup>-Representation).

Constant	This study	Previous free jet experiment [g]
A [b] / MHz	10454.26296 (48) [f]	10454.260 (6)
B / MHz	4041.15605 (24)	4041.163 (4)
C / MHz	2972.07260 (20)	2972.083 (5)
$ \mu_a ,  \mu_b ,  \mu_c $ [b] / D	Yes / Yes / No	Yes / Yes / No
$\chi_{aa}$ [c]	2.05153 (19)	2.3 (2)
$\chi_{bb}$	1.93027 (31)	2.2 (1)
$\chi_{cc}$	-3.98181 (31)	-3.5 (1)
$\Delta_J$ / kHz	0.77546 (14)	0.80 (2)
$\Delta_K$ / kHz	2.35393 (55)	2.36 (6)
$\Delta_{JK}$ / kHz	6.47368 (90)	6.3 (2)
$\delta_J$ / kHz	0.204628 (23)	0.202 (5)
$\delta_k$ / kHz	2.0742 (21)	1.8 (2)
$\Phi_J$ / mHz	-0.492 (36)	-
$\Phi_{JK}$ / Hz	-0.02302 (28)	-
$\Phi_K$ / Hz	0.04078 (77)	-
$\phi_{JK}$ / Hz	-0.01456 (71)	-
$\phi_K$ / Hz	-0.1638 (21)	-
N [d]	1480	121
$\sigma$ [e]	50	34

[a]  $A$ ,  $B$ , and  $C$  represent the rotational constants; [b]  $\mu_a$ ,  $\mu_b$ , and  $\mu_c$  are the electric dipole moment components. [c]  $\chi_{aa}$ ,  $\chi_{bb}$  and  $\chi_{cc}$  are the diagonal elements of the  $^{14}\text{N}$  nuclear quadrupole coupling tensor. [d]  $N$  is the number of measured transitions. [e]  $\sigma$  is the root mean square (rms) deviation of the fit (in kHz). [f] Standard error in parentheses in units of the last digit. [g] Taken from Maris 2004.

We provide the rotational ( $Q_r$ ) and vibrational ( $Q_v$ ) partition functions of glycolamide in Table 10. The values of  $Q_r$  were calculated from first principles at different temperatures, using the Picket program (Pickett 1991). The maximum value of the  $J$  quantum number of the energy levels taken into account to calculate the partition function is 200. The vibrational part,  $Q_v$ , was estimated using a harmonic approximation and a simple formula that corresponds to Eq. (3.60) of Gordy & Cook (1970). The frequencies

of the normal modes were obtained from double-hybrid calculations of the harmonic force field using the B2PLYPD3 method and the aug-cc-pVTZ basis set (see Table A.2 of the Appendix). The values of  $Q_v$  were calculated by taking the ten lowest vibrational modes into account. The full partition function,  $Q_{tot}$ , is thus the product of  $Q_r$  and  $Q_v$ .

**Table 10.** Rotational and vibrational partition functions of glycolamide.

Temperature (K)	$Q_r^{(a)}$	$Q_v^{(b)}$
2.72	68.8745	1.0000
8.00	342.4694	1.0000
9.38	434.0770	1.0000
18.75	1224.6939	1.0024
37.50	3459.8737	1.0529
75.00	9781.5471	1.3302
100.00	15059.0883	1.6225
150.00	27667.5424	2.5317
225.00	50841.4225	5.2978
300.00	78229.5551	11.5094

**Notes.** <sup>(a)</sup>  $Q_r$  is the rotational partition function. It does not take the hyperfine splitting into account. <sup>(b)</sup>  $Q_v$  is the vibrational partition function. The total partition function of the molecule (without hyperfine splitting) is  $Q_r \times Q_v$ .

**Table 11.** Sample list of the measured transitions in the ground-state of glycolamide.

Technique	J'	K <sub>a</sub> '	K <sub>c</sub> '	F'	J''	K <sub>a</sub> ''	K <sub>c</sub> ''	F''	$\nu_{obs}^{(a)}$	$\nu_{calc}^{(b)}$	$\nu_{obs} - \nu_{calc}^{(c)}$
MB-FTMW	2	0	2	1	1	0	1	1	13902.46900	13902.46519	0.00381
	2	0	2	3	1	0	1	2	13903.52500	13903.52425	0.00075
	2	0	2	1	1	0	1	0	13904.00500	13904.00377	0.00123
	2	0	2	2	1	0	1	1	13903.70900	13903.70734	0.00166
	2	0	2	2	1	0	1	2	13904.32500	13904.32278	0.00222
DR-MB-FTMW	3	0	3	2	2	0	2	2	20557.42000	20557.43477	-0.01477
	3	0	3	4	2	0	2	3	20558.60000	20558.59481	0.00519
	3	0	3	2	2	0	2	1	20558.69000	20558.67692	0.01308
	3	0	3	3	2	0	2	2	20558.82000	20558.82911	-0.00911
	3	0	3	3	2	0	2	3	20559.64000	20559.62765	0.01235
mmW	17	0	17	-	16	1	16	-	103873.60062	103873.57857	0.02204
	17	1	17	-	16	1	16	-	103874.60897	103874.61022	-0.00126
	17	0	17	-	16	0	16	-	103875.68167	103875.65985	0.02182
	17	1	17	-	16	0	16	-	103876.71342	103876.69150	0.02192
	38	5	34	-	37	4	33	-	251427.51339	251427.53733	-0.02394
	38	4	34	-	37	4	33	-	251426.47975	251426.49068	-0.01093
	38	5	34	-	37	5	33	-	251425.49693	251425.49090	0.00603
	38	4	34	-	37	5	33	-	251424.35071	251424.32875	0.02195
	48	31	17	-	48	30	18	-	419757.55974	419757.52028	0.03946
	48	31	18	-	48	30	19	-	419757.55974	419757.52028	0.03946
	49	31	18	-	49	30	19	-	419568.26082	419568.23007	0.03075
	49	31	19	-	49	30	20	-	419568.26082	419568.23007	0.03075
	32	17	15	-	31	16	16	-	453552.81658	453552.80102	0.01555
	32	17	16	-	31	16	15	-	453552.81658	453552.80099	0.01559

**Notes.** Upper and lower state quantum numbers are indicated by ' and ', respectively. <sup>(a)</sup> Observed frequency. <sup>(b)</sup> Calculated frequency. <sup>(c)</sup> Observed minus calculated frequency.

## 2.4 Search for glycolamide toward Sgr B2(N2) with ALMA

### 2.4.1 Observations

We used the spectroscopic predictions derived in Section 2.3.2 to search for glycolamide toward the hot molecular core Sgr B2(N2) located in the giant molecular cloud Sgr B2. We used the EMoCA imaging spectral line survey performed toward the protocluster Sgr B2(N) with ALMA. As mentioned in Sect. 1, Sgr B2(N) is a high-mass star forming region with a very rich chemistry that is part of the Sagittarius B2 giant molecular cloud in which many complex organic molecules were discovered for the first time in the ISM (see, e.g., Menten 2004, and Sect. 5.3). A detailed account of the observational setup and data reduction of the EMoCA survey can be found in Belloche et al. (2016). In short, the survey covers the frequency range from 84.1 GHz to 114.4 GHz with a spectral resolution of 488 kHz (1.7-1.3 km s<sup>-1</sup>) and a median angular resolution of 1.6''. The field was centered at  $(\alpha, \delta)_{J2000} = (17^{\text{h}}47^{\text{m}}19^{\text{s}}87, -28^{\circ}22'16''0)$ . Here we analyze the spectrum at the peak position of Sgr B2(N2) located at  $(\alpha, \delta)_{J2000} = (17^{\text{h}}47^{\text{m}}19^{\text{s}}86, -28^{\circ}22'13''4)$

As described in Belloche et al. (2016), the spectrum was modeled with the software Weeds (Maret et al. 2011) under the assumption of local thermodynamical equilibrium (LTE), which is an appropriate assumption given the high densities characterizing the region where the hot-core emission is detected ( $>1 \times 10^7 \text{ cm}^{-3}$ , see Bonfand et al. 2019). A best-fit synthetic spectrum of each molecule was derived separately, and the contributions of all identified molecules were then added together. Each species was modeled with a set of five parameters: size of the emitting region, column density, temperature, linewidth, and velocity offset with respect to the assumed systemic velocity of the source (74 km s<sup>-1</sup>).

### 2.4.2 Nondetection of glycolamide

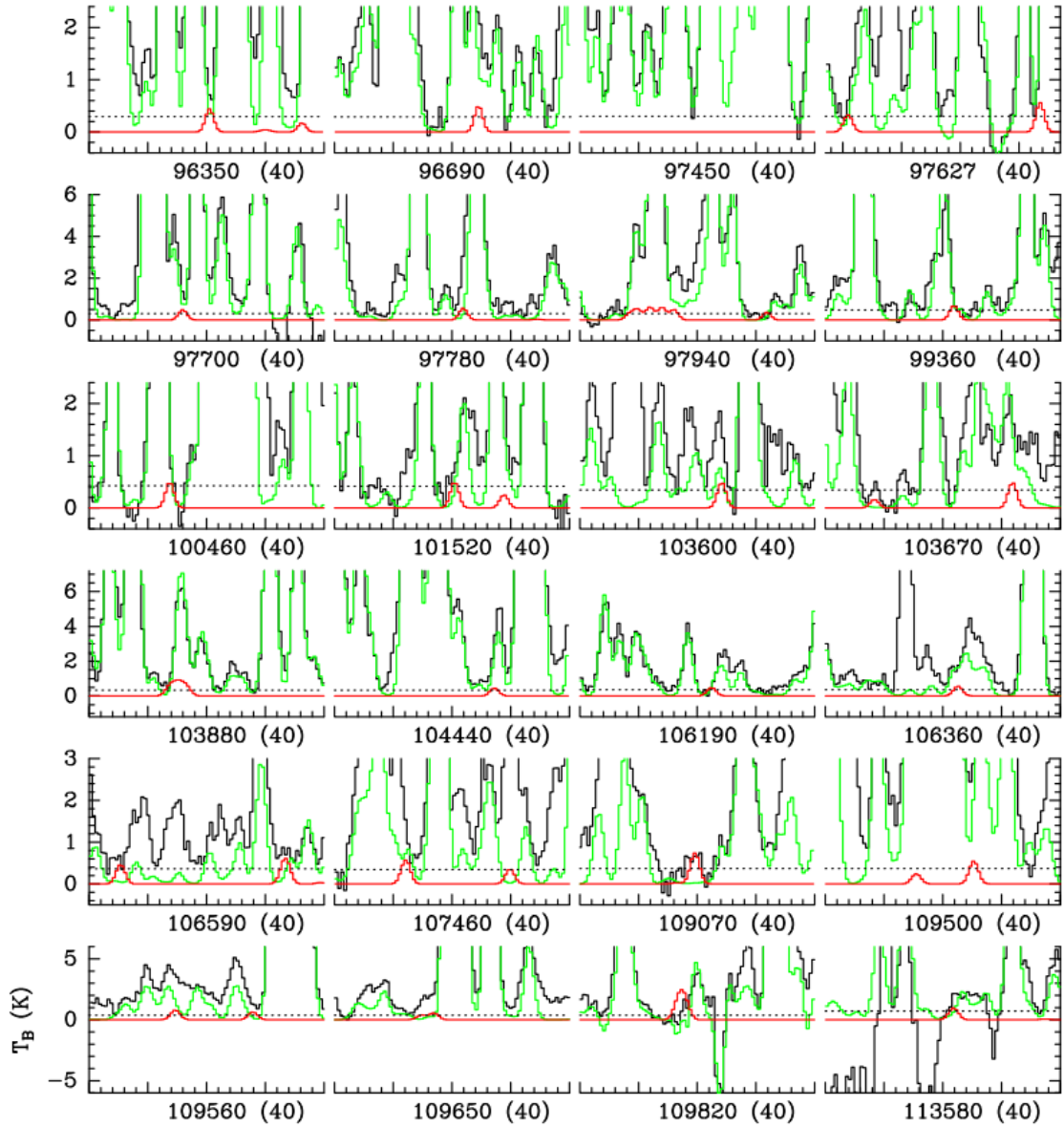
We did not find clear evidence for the presence of glycolamide in Sgr B2(N2). Fig. 12 shows in red the synthetic spectrum that we computed to obtain the upper limit on the glycolamide column density. In order to model the emission of glycolamide, we assumed the same parameters as those derived by Belloche et al. (2017) for acetamide (CH<sub>3</sub>C(O)NH<sub>2</sub>): an emission size of 0.9'', a temperature of 180 K, a linewidth of 5 km s<sup>-1</sup>, and no velocity offset with respect to the systemic velocity. The red spectrum shown in Fig. 4 was computed with a glycolamide column density of  $2.4 \times 10^{16} \text{ cm}^{-2}$  (see Table

12).

The ALMA spectrum displayed in Fig. 4 shows a  $\sim 5\sigma$  line matching one line of the synthetic spectrum of glycolamide at  $\sim 109070$  MHz, but several issues at other frequencies suggest that this is likely fortuitous. One major problem is indeed that the glycolamide synthetic spectrum at  $\sim 103\,875$  MHz, if added to the contribution of all identified species, would significantly overestimate the signal detected with ALMA at this frequency, which is already accounted for (even slightly overestimated) by the contributions of ethanol and  $\text{CH}_2^{13}\text{CHCN}$ . The latter species both have many lines detected in the survey and well reproduced by our best-fit LTE model (see Belloche et al. 2016; Müller et al. 2016). There is thus no room left for glycolamide emission at this frequency.

Another problem occurs at  $\sim 109817$  MHz, where the red spectrum overestimates the detected signal. However, as can be seen in the green spectrum, there are several absorption components tracing diffuse and translucent clouds along the line of sight that are expected from  $\text{C}^{18}\text{O } J = 1-0$  in this frequency range (see, e.g., Thiel et al. 2019), in particular two of them very close to the glycolamide transition frequency. The strength of these  $\text{C}^{18}\text{O}$  absorption components is difficult to constrain. Therefore, the inconsistency of the glycolamide spectrum at this frequency may not be significant.

All in all, the fact that only one line of glycolamide matches a detected line and the fact that there is no room left for glycolamide emission at  $\sim 103\,875$  MHz leads us to conclude that glycolamide is not detected in Sgr B2(N2). We thus consider the red spectrum shown in Fig. 4 as an upper limit to the emission of glycolamide in Sgr B2(N2). Also very recently, glycolamide has been search for towards the hot core G31.41+0.31 with the ALMA 3 mm spectral survey GUAPOS (**G**31.41+0.31 **U**nbiased **A**LMa **s**Pectral **O**bservational **S**urvey) without success (Colzi et al. 2021).



**Fig. 12.** Synthetic LTE spectrum of glycolamide (in red) used to derive the upper limit to its column density, overlaid on the ALMA spectrum of Sgr B2(N2) (in black) and the synthetic spectrum that contains the contributions from all the species (but not from glycolamide) that we have identified so far in this source (in green). The dotted line in each panel indicates the  $3\sigma$  noise level. For clarity purposes, the  $x$ -axis of each panel is only labeled with its central frequency followed by the value of the displayed frequency range in parentheses, both in MHz. The transitions of glycolamide that are weaker than  $3\sigma$  and the ones which are heavily contaminated by other species are not shown.

### 2.4.3 Comparison to other molecules

In order to put the column density upper limit derived for glycolamide into a broader astrochemical context, we used the EMOCA (Exploring Molecular Complexity with



ALMA) survey to determine the column densities of other species detected toward Sgr B2(N2) that can be compared to glycolamide. We selected acetaldehyde ( $\text{CH}_3\text{CHO}$ ), glycolaldehyde ( $\text{CH}_2(\text{OH})\text{CHO}$ ), and acetamide ( $\text{CH}_3\text{C}(\text{O})\text{NH}_2$ ) because acetamide and glycolamide are structurally related in the same way as acetaldehyde and glycolaldehyde: in both cases, the substitution of a hydrogen atom in the methyl group of the former species with an hydroxyl group gives the latter species. The column density derived for acetamide in Sgr B2(N2) was already reported in Belloche et al. (2017).

For both acetaldehyde and glycolaldehyde, we used the spectroscopic predictions (entries 44003 version 3 and 60006 version 2, respectively) available in the Jet Propulsion Laboratory (JPL) spectroscopic catalog (Pickett et al. 1998). The acetaldehyde entry is based on Kleiner et al. (1996) with data in the range of our survey from Kleiner et al. (1991, 1992). The ground state data of glycolaldehyde are based on Butler et al. (2001), those of the excited data on Widicus Weaver et al. (2005). Recently, we became aware of an extensive study of the vibrational spectrum of glycolaldehyde by Johnson et al. (2013). The vibrational energies are higher than those used for the JPL catalog which were based on relative intensities of vibrational states in the rotational spectrum of glycolaldehyde. As a result, the vibrational correction factor to the partition function at 170 K is only 1.38 instead of 1.50. We made the appropriate correction to the column density derived below. For the  $^{13}\text{C}$  isotopologs of acetaldehyde, we used spectroscopic predictions from Margulès et al. (2015). For the  $^{13}\text{C}$  isotopologs of glycolaldehyde, we used predictions (entries 61513 and 61514, both version 1) available in the Cologne Database for Molecular Spectroscopy (Endres et al. 2016). The  $^{13}\text{C}$  entries of glycolaldehyde are based on Haykal et al. (2013). The acetamide entry is based on Ilyushin et al. (2004).

Many transitions of acetaldehyde and glycolaldehyde in their vibrational ground state are detected toward Sgr B2(N2) (see Figs. A.1 and A.8 in [Appendix A](#)). We also detect transitions from within the first two torsionally excited states of acetaldehyde (Figs. A.2 and A.3 in [Appendix A](#)) and transitions from within the first two vibrationally excited states of glycolaldehyde (Figs. A.9 and A.10 in [Appendix A](#)). There is an issue with the transition of  $\text{CH}_3\text{CHO}$   $\nu_t = 2$  at  $\sim 90\,910$  MHz, and we suspect that this results from an inaccurate rest frequency. The frequency computed for this transition by Kleiner et al. (1996) is indeed somewhat higher than the frequency listed in the JPL catalog:  $90\,910.914 \pm 0.204$  MHz versus  $90\,910.330 \pm 0.101$  MHz. The former would better match the EMOCA spectrum.

**Table 12.** Parameters of our best-fit LTE model of acetaldehyde, glycolaldehyde, and acetamide, and column density upper limit for glycolamide, toward Sgr B2(N2).

Molecule	Status <sup>(a)</sup>	N <sub>det</sub> <sup>(b)</sup>	Size <sup>(c)</sup> ( $''$ )	T <sub>rot</sub> <sup>(d)</sup> (K)	N <sup>(e)</sup> (cm <sup>-2</sup> )	F <sub>vib</sub> <sup>(f)</sup>	$\Delta V$ <sup>(g)</sup> (km s <sup>-1</sup> )	$\Delta V_{\text{off}}$ <sup>(h)</sup> (km s <sup>-1</sup> )	N <sub>ref</sub> / N <sup>(i)</sup>
<i>Acetaldehyde</i>									
CH <sub>3</sub> CHO, $\nu = 0^*$	d	19	1.1	160	5.3 (17)	1.08	5.6	0.0	1
$\nu_t = 1$	d	13	1.1	160	5.3 (17)	1.08	5.6	0.0	1
$\nu_t = 2$	d	2	1.1	160	5.3 (17)	1.08	5.6	0.0	1
<sup>13</sup> CH <sub>3</sub> CHO, $\nu = 0$	t	2	1.1	160	2.4 (16)	1.01	5.6	0.0	22
$\nu_t = 1,$	n	0	1.1	160	2.4 (16)	1.01	5.6	0.0	22
CH <sub>3</sub> <sup>13</sup> CHO, $\nu = 0$	n	0	1.1	160	2.4 (16)	1.01	5.6	0.0	22
$\nu_t = 1,$	n	0	1.1	160	2.4 (16)	1.01	5.6	0.0	22
<i>Glycolaldehyde</i>									
CH <sub>2</sub> (OH)CHO, $\nu = 0^*$	d	21	1.0	170	1.1 (17)	0.92	5.4	1.0	1
$\nu_1 = 1$	t	2	1.0	170	1.1 (17)	0.92	5.4	1.0	1
$\nu_2 = 2$	t	1	1.0	170	1.1 (17)	0.92	5.4	1.0	1
<sup>13</sup> CH <sub>2</sub> (OH)CHO, $\nu = 0$	n	0	1.0	170	< 1.6 (16)	1.37	5.4	1.0	>6.7
CH <sub>2</sub> (OH) <sup>13</sup> CHO, $\nu = 0$	n	0	1.0	17	< 2.1 (16)	1.37	5.4	1.0	<5.4
<i>Acetamide</i>									
CH <sub>3</sub> C(O)NH <sub>2</sub> , $\nu = 0^*$	d	10	0.9	180	1.4 (17)	1.23	5.0	1.5	1
$\nu_t = 1$	d	8	0.9	180	1.4 (17)	1.23	5.0	1.5	1
$\nu_t = 2$	d	5	0.9	180	1.4 (17)	1.23	5.0	1.5	1
$\Delta\nu_t \neq 0$	t	0	0.9	180	1.4 (17)	1.23	5.0	1.5	1
<i>Glycolamide</i>									
CH <sub>2</sub> (OH)C(O)NH <sub>2</sub> , $\nu = 0$	d	0	0.9	180	< 2.4 (16)	3.38	5.0	0.0	-

**Notes.** The parameters for acetamide were published in Belloche et al. (2017). <sup>(a)</sup>d: detection, t: tentative detection, n: nondetection. <sup>(b)</sup>Number of detected lines (conservative estimate, see Sect. 3 of Belloche et al. 2016). One line of a given species may mean a group of transitions of that species that are blended together. <sup>(c)</sup>Source diameter (*FWHM*). <sup>(d)</sup>Rotational temperature. <sup>(e)</sup>Total column density of the molecule. x (y) means  $x \times 10^y$ . An identical value for all listed vibrational and torsional states of a molecule means that LTE is an adequate description of the vibrational and torsional excitation. <sup>(f)</sup>Correction factor that was applied to the column density to account for the contribution of vibrationally excited states, in the cases where this contribution was not included in the partition function of the spectroscopic predictions. For glycolaldehyde, see explanation in Sect. 4.3. <sup>(g)</sup>Linewidth (*FWHM*). <sup>(h)</sup>Velocity offset with respect to the assumed systemic velocity of Sgr B2(N2),  $V_{\text{sys}} = 74 \text{ km s}^{-1}$ . <sup>(i)</sup>Column density ratio, with  $N_{\text{ref}}$  the column density of the previous reference species marked with a \*.

On the basis of a first-guess LTE synthetic spectrum for each molecule, we selected the transitions that are not significantly contaminated by other species to measure the size of the emission of acetaldehyde and glycolaldehyde. The emission is compact around Sgr B2(N2) and we obtained median sizes (HPBW) of  $1.1''$  and  $1.0''$ , respectively. The transitions that are not much contaminated were then used to produce population diagrams (Figs. A.12a and A.13a). Following the method described in Belloche et al. (2016), we corrected the diagrams for optical depth and for the contamination due to the other species using our synthetic spectrum that includes the contribution of all species identified so far (Figs. A.12a and A.13b). A fit to the corrected population diagrams yields temperatures of  $166 \pm 17$  K and  $164 \pm 6$  K for acetaldehyde and glycolaldehyde, respectively (see Table 13). As explained in Sect. 3 of Belloche et al. (2016), the uncertainties on the rotational temperatures are purely statistical and do not take into account the residual contamination by species that have not been identified so far. The best-fit spectra shown in Figs. A.1–A.3 and Figs. A.8–A.10 were computed with temperatures of 160 K and 170 K, respectively, consistent within the uncertainties with the values derived from the fits to the population diagrams.

On the basis of the parameters obtained for acetaldehyde and glycolaldehyde, we also searched for emission of their  $^{13}\text{C}$  isotopologs, assuming a  $^{12}\text{C}/^{13}\text{C}$  isotopic ratio in the range 20–25 as usually found for complex organic molecules in Sgr B2(N2) (see, e.g., Belloche et al. 2016; Müller et al. 2016). We tentatively detect  $^{13}\text{CH}_3\text{CHO}$  in its vibrational ground state (Fig. B.4), with an isotopic ratio of 22. Assuming the same isotopic ratio for both  $^{13}\text{C}$  isotopologs, which is commonly the case in Sgr B2(N2) for COMs (Belloche et al. 2016; Margulès et al. 2016; Müller et al. 2016), yields a synthetic spectrum of  $\text{CH}_3^{13}\text{CHO}$  consistent with the observed spectrum (Fig. A.6). However, no transition is sufficiently free of contamination from emission of other species to claim a detection of this isotopolog. Still, we include it in our complete model because its emission contributes significantly to the signal detected at several frequencies. Transitions from the first torsionally excited state of both isotopologs contribute to the observed spectra at the  $\sim 3\sigma$  level at most (see Figs. A.5 and A.7). The  $^{13}\text{C}$  isotopologs of glycolaldehyde are not detected. We report upper limits to their column densities in Table 12. Their nondetection is consistent with the  $^{12}\text{C}/^{13}\text{C}$  typical ratio obtained for other COMs in Sgr B2(N2).

**Table 13.** Rotational temperatures derived from population diagrams toward Sgr B2(N2).

Molecule	States <sup>(a)</sup>	T <sub>fit</sub> <sup>(b)</sup> (K)
CH <sub>3</sub> CHO	v = 0, v <sub>t</sub> = 1, v <sub>t</sub> = 2	166 (17)
CH <sub>2</sub> (OH)CHO	v = 0, v <sub>1</sub> = 1, v <sub>2</sub> = 1	163.8 (6.1)

**Notes.** <sup>(a)</sup>Vibrational states that were taken into account to fit the population diagram. <sup>(b)</sup>The standard deviation of the fit is given in parentheses. As explained in Sect. 3 of Belloche et al. (2016), this uncertainty is purely statistical and should be viewed with caution. It may be underestimated.

From Table 12 we derive the following column density ratios in Sgr B2(N2):  $[\text{CH}_3\text{CHO}]/[\text{CH}_2(\text{OH})\text{CHO}] = 4.8$  and  $[\text{CH}_3\text{C}(\text{O})\text{NH}_2]/[\text{CH}_2(\text{OH})\text{C}(\text{O})\text{NH}_2] > 5.8$ . The lower limit on the latter ratio is only slightly higher than the former ratio. We also find that glycolamide is at least five times less abundant than glycolaldehyde ( $[\text{CH}_2(\text{OH})\text{CHO}]/[\text{CH}_2(\text{OH})\text{C}(\text{O})\text{NH}_2] > 4.6$ ).

## 2.5 Search for glycolamide toward Sgr B2(N) with Effelsberg

At low radio frequencies, in the cm-wavelength range, transitions of many molecules including COMs are often seen in absorption against the strong nonthermal (synchrotron) radiation (Hollis et al. 2007) of the Sgr B2 complex, while others appear in emission depending on the excitation process (e.g., Menten 2004; Hollis et al. 2004b, 2006, and see Sect. 2.4.3 below). In the following we describe our search for several low- $J$  transitions of glycolamide within in the 4–8 and 12–17 GHz frequency ranges.

### 2.5.1 Observations

We performed measurements (project id: 08-20) toward Sgr B2(N) with the S45mm and S20mm receivers of the 100 m radio telescope at Effelsberg, Germany<sup>1</sup> on 28 March and 3 April 2020. The targeted position was  $\alpha_{J2000} = 17^{\text{h}}47^{\text{m}}19.8^{\text{s}}$ ,  $\delta_{J2000} = -28^{\circ}22'17''$ . This is identical to the position used by Hollis et al. (2006) for their acetamide observations and within a few arc seconds of the position of Sgr B2(N2) hot core for which the ALMA spectra discussed in Sect. 2.3 were extracted. We used the position-switching mode with an off-source position located 600 east of Sgr B2(N). The S45 mm receiver is a broad-

<sup>1</sup> The 100 m telescope at Effelsberg is operated by the Max-PlanckInstitut für Radioastronomie (MPIfR) on behalf of the Max-PlanckGesellschaft (MPG).

band receiver with two orthogonal linear polarizations operating in the frequency range between 4 and 9.3 GHz. The S20 mm receiver is a dual-beam receiver with two orthogonal linear polarizations covering the frequency range 12–18 GHz.

The focus was adjusted using observations of 3C286 at the beginning of each observing session. 3C286 was also used as flux calibrator, and the flux calibration accuracy is estimated to be within  $\sim 10\%$ . NRAO530 was used as pointing source close to Sgr B2(N). The pointing accuracy was found to be better than  $5''$ , small compared to the half-power beam width (HPBW) of  $99''$  at 7.2 GHz and  $57''$  at 13 GHz. The main beam efficiency is about 69% at 7.2 GHz and 66% at 13 GHz, and the typical system temperature was about 33 K at 7.2 GHz and 41 K at 13 GHz. The spectroscopic backends were Fast Fourier Transform spectrometers (FFTSs, e.g., Klein et al. 2012). The FFTSs cover, in two bands, the frequency ranges 4–8 GHz and 12–17 GHz with a channel width of 38.1 kHz, corresponding to a velocity spacing of  $1.6 \text{ km s}^{-1}$  at 7 GHz and  $0.9 \text{ km s}^{-1}$  at 13 GHz. This is adequate, given the expected line widths of  $\sim 10 \text{ km s}^{-1}$  (see below). The data reduction was performed using the GILDAS/CLASS software<sup>2</sup>.

### 2.5.2 *Nondetection of glycolamide*

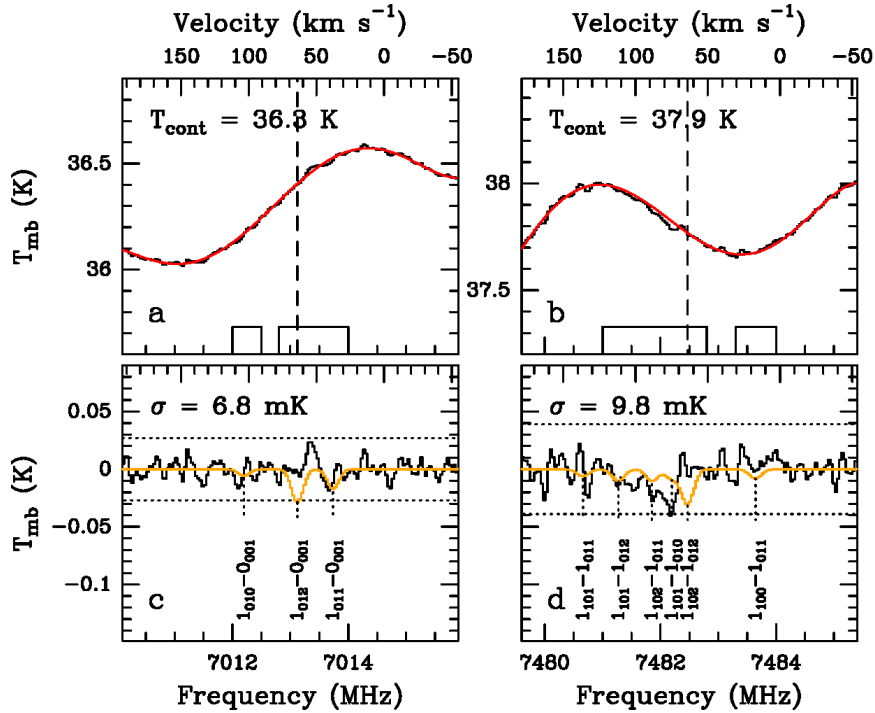
We used the spectroscopic predictions (with hyperfine structure) derived in Section 2.5 to search for glycolamide in the Effelsberg spectra of Sgr B2(N). None of the glycolamide transitions covered by our observations were detected either in absorption or in emission. Figures 13 and A.11 illustrate these nondetections with several portions of the spectra that cover some of the lowest energy transitions of glycolamide included in our observations. The spectra of Figs. 13a–b and A.11a–d were obtained after averaging both polarizations. We then measured the average continuum level in each panel and removed a fifth-order baseline to produce the spectra shown in Figs. 13c–d and A.11e–h.

We used the software Weeds to produce synthetic spectra assuming a beam-filling factor of 1. We made the same assumptions as Hollis et al. (2006) made to analyze their detection of acetamide with the GBT: a systemic velocity of  $64 \text{ km s}^{-1}$  and a rotational temperature of 8 K. We assumed a linewidth (FWHM) of  $10 \text{ km s}^{-1}$ , which is similar to the linewidth of glycolaldehyde detected with the GBT (Hollis et al. 2004b). We assumed

---

<sup>2</sup> A detailed information is given in <http://www.iram.fr/IRAMFR/GILDAS>.

a background temperature equal to the average continuum level determined in Figs. 5a–b and B.11a–d. With these parameters, we obtained a column density upper limit of  $1.2 \times 10^{14} \text{ cm}^{-2}$  for glycolamide at the  $4\text{--}5\sigma$  level for extended emission from this molecule. This value is similar to the column densities derived by Hollis et al. (2004b, 2006) for acetamide and glycolaldehyde with the GBT.



**Fig. 13** (a), (b) Examples of spectra obtained with Effelsberg between 4 and 8 GHz toward Sgr B2(N) resulting from averaging both polarizations. In each panel, the vertical dashed line marks a velocity of  $64 \text{ km s}^{-1}$ . The red curve shows the baseline fitted with a 5th-order polynomial function. The boxes at the bottom show the windows avoided to perform the baseline fit. The average continuum level used to compute the synthetic spectrum is indicated in the top left corner. (c), (d) Synthetic spectra (orange) used to derive the upper limit to the column density of glycolamide, overlaid on observed spectra after baseline subtraction (black). The rms noise level is indicated in the top left corner. The horizontal dotted lines show the  $4\sigma$  and  $-4\sigma$  levels. The components of each hyperfine multiplet are labeled and marked with vertical dotted lines. The velocity axis shown at the top was computed relative to the strongest component of each hyperfine multiplet.

For completeness, Table A.3 lists all glycolamide transitions with a lower level energy in temperature units lower than 10 K covered by the Effelsberg observations. None of these transitions are detected, and the table indicates the noise level and the average continuum level.

### 2.5.3 Discussion of the Effelsberg nondetection

Our astronomical search for centimeter wavelength transitions from glycolamide was motivated by the fact that in the early days of molecular radio astronomy, the early 1970s, radiation from many large molecules was first detected toward Sgr B2 with single dish telescopes (for a summary and a list of early detections, see Menten 2004). These identifications were made via detections of cm-wavelength low- $J$  rotational ground-state or near ground state transitions. More recent observations with the GBT added a significant number of species to this list, namely glycolaldehyde ( $\text{CH}_2(\text{OH})\text{CHO}$ ), ethylene glycol ( $\text{HOCH}_2\text{CH}_2\text{OH}$ ), propenal ( $\text{CH}_2\text{CHCHO}$ ), propanal ( $\text{CH}_3\text{CH}_2\text{CHO}$ ) and acetamide ( $\text{CH}_3\text{C}(\text{O})\text{NH}_2$ ) (Hollis et al. 2002, 2004a,b, 2006).

In many cases, these low- $J$  transitions were exclusively found toward Sgr B2 and its larger environment, the central molecular zone (CMZ) of our Galaxy (Morris & Serabyn 1996). The CMZ is a  $\sim 200$  pc long region stretched around the Galactic center that contains molecular gas that has much higher temperatures and densities than giant molecular clouds (GMCs) in the Galactic disk as well as a peculiar chemistry. The latter two facts cause a variety of molecules to be observable over the whole CMZ, via mm wavelength rotational lines (e.g., Jones et al. 2012; Requena-Torres et al. 2006), but also in the 834.3 MHz  $1_1 - 1_1\text{A}^\mp$  line of  $\text{CH}_3\text{OH}$  and low- $J$  lines from various COMs (Gottlieb et al. 1979; Requena-Torres et al. 2008).

The star formation/GMC complex Sagittarius B2 is a prominent part of the CMZ and the mentioned low- $J$  lines were first (and many of them only) detected toward the extended molecular envelope of this source, which is a prominent source of mm-wavelength emission from very many molecules (e.g., Jones et al. 2007). Important here is the imaging of the 1065 MHz  $1_{11}-1_{10}$  rotational transition of acetaldehyde ( $\text{CH}_3\text{CHO}$ ), which shows its emission to be extended over the whole complex, that is, a  $3 \times 10$  pc<sup>2</sup> sized region (Chengalur & Kanekar 2003) that appears to be part of the several times larger extended envelope of the Sgr B2 complex (Comito et al. 2010; Schmiedeke et al. 2016). The even (much) more extreme conditions in its embedded dense molecular cores, for example, Sgr B2(N2), allow the detection of many rare species in these very compact regions that have sizes of order  $1''$  (or  $\approx 0.04$  pc), for instance, with the EMoCA survey (see Table 5). For the column densities and sizes reported in Table 5, any thermal (LTE) signal in the cm lines for the listed species would be undetectable because of the lines'

low transition probabilities and the large beams they are observed with. In fact, the Weeds LTE model of Sgr B2(N2) (Sect. 2.4) produces a peak temperature of a few  $\mu\text{K}$  for the 7013 MHz lines of glycolamide when computed at the resolution of the Effelsberg telescope.

For various reasons, it is clear that the low frequency, low- $J$ , emission or absorption lines observed from some molecules in the CMZ are only observable because their observed intensities are the result of a population inversion causing weak maser emission in some lines or anti-inversion causing enhanced absorption (over-cooling) in others. The weak (anti-)inversion results in generally small (absolute) optical depths. Nevertheless, the prominent background continuum emission causes the weak maser emission or absorption to be detectable, since the line signal is proportional to the optical depth times the (very high) continuum flux. To give an example: that nonthermal excitation produces the observed intensities of the formamide and acetamide lines reported by Hollis et al. (2006) follows directly from the facts that, first, some of them show absorption, while others show emission. Second the total column densities calculated from the upper level column densities determined for individual lines vary by a factor of six for the acetamide lines and a factor of ten for the formamide lines, indicating non-LTE conditions. To calculate these column densities, Hollis et al. (2006) assume a rotation temperature of 8 K, which suggests a widespread “cold envelope” as the origin of these molecules, which is also implied by GBT observations of other COMs (see references in Hollis et al. 2006). Also Requena-Torres et al. (2006) find mm-wavelength emission from many COMs to be widespread over the CMZ and characterized by a rotation temperature of 8 K, while the same group of authors (Requena-Torres et al. 2008) derive 8-16 K from the cm lines of assorted COMs. They invoke subthermal excitation to explain these low numbers, which are much lower than the gas kinetic temperatures of  $>60$  K found throughout the CMZ by Ginsburg et al. (2016), who find the extended Sgr B2 cloud to have some of the highest values ( $>100$  K). For consistency with other studies of cm lines, we have adopted 8 K to calculate our column density upper limits.

Whether inversion or anti-inversion (over-cooling) of a line can be attained, depends on the interplay of the molecule’s energy level structure and its radiative and collisional (de)excitation rates. Since for most species the latter are unknown, it is in general impossible to predict or explain observed maser (or anti-maser) action. Recently, collisional rate coefficients were calculated for methyl formate ( $\text{CH}_3\text{OCHO}$ ) and



methanimine ( $\text{CH}_2\text{NH}$ ) by Faure et al. (2014) and Faure et al. (2018), respectively. Using their results, these authors' excitation calculations indeed confirm non-LTE excitation and reproduce the inversion observed in several low- $J$  lines of these molecules.

While some low- $J$  lines from the above-mentioned COMs and assorted others show extended non-thermal weak maser emission or enhanced absorption, others do not. Obviously, glycolamide does not and neither does amino acetonitrile, according to Brown et al. (1977) and Belloche et al. (2008), who unsuccessfully searched for this molecule's  $1_{01}-0_{00}$  multiplet at 9071.28 MHz (but did detect its mm-wavelength spectrum in the Sgr B2(N) hot core).

The upper limit we obtained from the cm-wavelength data in the previous section  $1.2 \times 10^{14} \text{ cm}^{-2}$ , is two orders of magnitude lower than the upper limit we calculated for the glycolamide column density in the Sgr B2(N2) hot core,  $2.4 \times 10^{16} \text{ cm}^{-2}$  (Table 5). These two values cannot be compared directly because the  $\text{H}_2$  column densities of the hot core and Sgr B2(N)'s envelope probed with ALMA and Effelsberg, respectively, are widely different. We note that the value for Sgr B2(N2) is calculated from a well-constrained measurement, while the value for the envelope depends on the questionable assumption of LTE with a rotation temperature of 8 K. Would we assume a value of 100 K, similar to the kinetic temperature of the envelope (Ginsburg et al. 2016), the upper limit based on the cm data would increase by a factor of 43 to  $5.3 \times 10^{15} \text{ cm}^{-2}$ . The acetamide column density presented by Hollis et al. (2006) would increase by a similar factor. Assuming 100 K for the rotation temperature is taking the opposite extreme to 8 K and is also an unrealistic assumption. The density in the molecular envelope is estimated to be of order a few times  $10^3 \text{ cm}^{-3}$  (Comito et al. 2010). At such moderate densities, the level population of any molecule with a substantial dipole moment will not be in LTE. Column density and abundance determinations for this environment are very difficult to make on the basis of a limited number of observed transitions, and impossible if their excitation is affected by inversion or antiinversion.

## 2.6 *Astrochemical modelling*

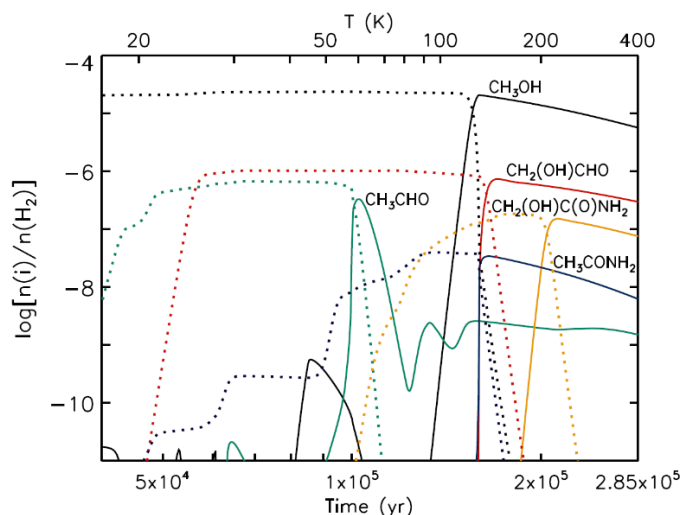
In order to interpret the observational findings, here we present results from the astrochemical kinetics model MAGICKAL (Model for Astrophysical Gas and Ice Chemical Kinetics and Layering) relating to the cold collapse and gradual warm up (8–400 K)

of a hot core with physical conditions appropriate to Sgr B2(N2). Selected model results were originally published by Garrod et al. (2017), who extended their chemical network to simulate the production of butyl cyanide and related species. The results presented here for other molecules are produced by the medium-timescale warm-up model described in detail by those authors. In brief, the physical model involves a freefall collapse from  $n_{\text{H}} = 3 \times 10^3$  to  $2 \times 10^8 \text{ cm}^{-3}$  over a period of  $\sim 1$  Myr. The warm-up stage then proceeds over a period  $2.85 \times 10^5$  yr, reaching a temperature of 200 K after  $2 \times 10^5$  yr. The chemical model includes gas-phase, grain-surface, and ice-mantle chemistry. In the latter two phases in particular, the addition of functional group radicals produced by cosmic ray-induced UV photolysis can result in a range of complex organic molecules, prior to the thermal desorption of the ice mantles into the gas phase.

Figure 14 shows the simulated time-dependent fractional abundances (with respect to  $\text{H}_2$ ) of the same molecules whose observational parameters are shown in Table 12; both the gas phase and grain-surface abundances are indicated for each species. Table 14 shows the peak gas-phase fractional abundances and the associated gas and dust temperatures obtained from the model for each of these species and some related molecules. In order to compare with the observational data, the ratios between the peak fractional abundances of key molecules are considered.

The ratio of glycolamide to glycolaldehyde found in the models (0.21) is consistent with the upper limit of the observational ratio (0.22). In the models, the main precursor to glycolamide is indeed glycolaldehyde; the latter is produced on the grains at around 30 K through the addition of the HCO and  $\text{CH}_2\text{OH}$  radicals, which are methanol photo-products. The abstraction of a hydrogen atom from the aldehyde end of the glycolaldehyde molecule, by various species including the radicals  $\text{NH}_2$  and OH, produces a radical that may combine with  $\text{NH}_2$  in the ice, or with NH followed by atomic hydrogen, resulting in glycolamide.

In the models, the main precursor to acetamide is its smaller homologue, formamide. Similarly to the grain-surface mechanism described above for glycolamide, hydrogen abstraction from the aldehyde end of formamide results in a radical that, with the addition of the methyl radical ( $\text{CH}_3$ ), produces acetamide. The model ratio of acetamide to formamide (0.071) is comparable to the observational ratio (0.054), based on data from Table 12 and from Belloche et al. (2019, their Table 6).



**Fig. 14.** Chemical model results for the molecules whose observational data are shown in Table 5; methanol is included for reference. Solid lines indicate gas-phase abundances with respect to  $\text{H}_2$ . Dashed lines of the same color indicate the solid-phase abundance of the same species, also with respect to gas-phase  $\text{H}_2$ . Data shown correspond to the warmup period following free-fall collapse.

While the ratios between each molecule and its precursor appear to agree fairly well with the observational values for glycolamide and acetamide, the direct ratio between these two species themselves does not; the models suggest a ratio of glycolamide to acetamide of 4.4, compared with an observational upper limit of 0.17, discrepant by a factor  $\sim 26$ . The modeled ratios of the precursors of these molecules show similar behavior, being too large by a factor  $\sim 33$ . This suggests that, while the chemistry of glycolamide and acetamide are well treated within the model, the relative abundances of their precursors require attention. It may be noted that the model abundance of methyl formate ( $\text{CH}_3\text{OCHO}$ ), also shown in Table 14, is around one order of magnitude lower than that of glycolaldehyde. This also defies the observational expectation for Sgr B2(N2) and various other star-forming regions that typically show the inverse relationship; Belloche et al. (2016) found a column density of methyl formate of  $1.2 \times 10^{18} \text{ cm}^{-2}$ , ten times larger than the glycolaldehyde abundance shown in Table 12. The inversion of this ratio in the models was noted by Garrod (2013), and could be related to the choice of dissociation branching ratios of methanol, or to the branching ratios of the radical-addition reaction ( $\text{HCO} + \text{CH}_2\text{OH}$ ) leading to glycolaldehyde on the grains. Whatever the cause, a reduction in the overall production of glycolaldehyde (and, by consequence, of glycolamide) in the models of around 30 times would appear to make the model ratios consistent with the observations for each of the molecules discussed here. Even with this

adjustment, the modeled ratio of glycolamide to glycolaldehyde remains unchanged, suggesting that the observational upper limit obtained for glycolamide may be close to the true column density.

**Table 14.** Peak fractional abundances for selected molecules with respect to H<sub>2</sub> obtained from the chemical kinetics model.

Molecule	$n(i)/n(\text{H}_2)$	$T$ (K)
CH <sub>3</sub> OH	2.1(-5)	130
CH <sub>3</sub> CHO	3.3(-7)	58
CH <sub>3</sub> (OH)CHO	7.4(-7)	148
CH <sub>3</sub> C(O)NH <sub>2</sub>	3.4(-8)	138
CH <sub>3</sub> (OH)C(O)NH <sub>2</sub>	1.5(-7)	226
NH <sub>2</sub> CHO	4.8(-7)	158
CH <sub>3</sub> OCHO	7.8(-8)	125

**Notes.** The associated temperature at which the peak value is achieved is also indicated. x(y) means  $x \times 10^y$ .

### 3. CONCLUSIONS

We have carried out a theoretical study of the possible products obtained from reaction between protonated hydroxylamine and acetic acid. These products have [H<sub>6</sub>C<sub>2</sub>O<sub>2</sub>N]<sup>+</sup> molecular formula and are structural isomers of protonated glycine. In addition, their neutral counterparts, with [H<sub>5</sub>C<sub>2</sub>O<sub>2</sub>N] molecular formula have been analyzed. The relevant spectroscopic parameters to rotational spectroscopy, harmonic and anharmonic frequencies, and IR intensities are reported for neutral and protonated isomers. This information could aid in their laboratory characterization by micro- and millimeter-wave spectroscopy or a preliminar astronomical search by radioastronomy or IR spectroscopy.

All reaction processes that could form [H<sub>6</sub>C<sub>2</sub>O<sub>2</sub>N]<sup>+</sup> isomers from the reaction between protonated hydroxylamine and acetic acid are exothermic except for those initiated by the most stable isomer of protonated hydroxylamine which lead to the P1 (CH<sub>3</sub>CONH<sub>2</sub><sup>+</sup>OH) and P7 (CH<sub>3</sub>COHNOH<sub>2</sub><sup>+</sup>) isomers. The study has been focused on the most abundant products, denoted as P1 (CH<sub>3</sub>CONH<sub>2</sub><sup>+</sup>OH), P2 (CH<sub>3</sub>COONH<sub>3</sub><sup>+</sup>), and P3 (CH<sub>3</sub>C(OH)<sup>+</sup>ONH<sub>2</sub>). Thus, for these isomers a detailed analysis of the corresponding singlet potential energy surface has been performed. Taken together, these result conclude that even if the reactions of formation of P2 and P3 isomers are exothermic processes, significant activation barriers were found in the paths leading to these products. The only exothermic process ( $\Delta E = -23.9$  kcal mol<sup>-1</sup> at the CCSD(T) level) with no net activation

barrier was initiated by the high-energy isomer of protonated hydroxylamine which leads to the P1 ( $\text{CH}_3\text{CONH}_2^+\text{OH}$ ) isomer. Therefore, the formation of  $\text{CH}_3\text{CONH}_2^+\text{OH}$  could be feasible under interstellar conditions from the reaction of the less stable isomer of protonated hydroxylamine and acetic acid. Consequently, from these results, acetoxyhydroxamic acid might be a candidate molecule to be searched for in the interstellar medium.

Moreover, we presented the first microwave study of acetoxyhydroxamic acid ( $\text{CH}_3\text{C}(\text{O})\text{NHOH}$ ), an archetypical hydroxamic acid and also a glycine isomer, using broadband and narrowband rotational spectroscopies equipped with a laser ablation vaporization system (LA-CP-FTMW and LA-MB-FTMW). We have characterized two distinct structures (*Z*-amide and *E*-amide) of acetoxyhydroxamic acid in the supersonic expansion. Accurate experimental values of the  $^{14}\text{N}$  nuclear quadrupole coupling constants together with the rotational constants and also the barrier to the internal rotation ( $V_3$ ) of the methyl group have been provided for both conformers in order to perfectly reproduce the spectrum to enable their distinct possible identification using low-frequency surveys. In conclusion, the precise spectroscopic information provided here should be relevant to guide radioastronomical searches of acetoxyhydroxamic acid in the ISM.

Furthermore, the rotational spectrum of glycolamide, another glycine isomer as well as direct a derivative of acetamide, has been recorded in a large spectral range interesting for astrophysical use (up to  $\sim 460$  GHz). Precise ground state spectroscopic constants up to sextic centrifugal distortion constants are provided for the most stable conformer obtained from the fit of more than 1480 transition lines. In this work, using the heated-cell millimeter-wave spectrometer at the University of Valladolid, we were able not only to transfer this organic molecule from the solid state to the gas phase but also to avoid condensation problems in the walls of the free-space cell.

We used this precise set of spectroscopic parameters to search for glycolamide toward the hot core Sgr B2(N2) and the envelope of Sgr B2(N). Meanwhile, for the astronomical search, we report a nondetection of glycolamide toward Sgr B2(N2) with ALMA with a column density upper limit of  $2.4 \times 10^{16} \text{ cm}^{-2}$ . This implies that glycolamide is at least six and five times less abundant than acetamide and glycolaldehyde, respectively, in this source. The former lower limit is only slightly higher than the column density ratio of acetaldehyde to glycolaldehyde, suggesting that the

actual column density of glycolamide in Sgr B2(N2) may not be far from the derived upper limit. We also report a nondetection of glycolamide toward Sgr B2(N) with the Effelsberg 100 meter radio telescope. We derived an upper limit to its beam-averaged column density of  $1.2 \times 10^{14} \text{ cm}^{-2}$  that is similar to the column densities of glycolaldehyde and acetamide reported earlier toward the same source with the Green Bank Telescope. Our astrochemical models suggest that glycolamide is a product of atomic-H abstraction from glycolaldehyde, followed by radical addition. Relative fractional abundances are close to the observed upper limit for glycolamide in Sgr B2(N2), suggesting that a slightly more sensitive search in this source could indeed yield a detection.

#### 4. EXPERIMENTAL SECTION

##### 4.2. *Microwave spectrum of acetohydroxamic acid: LA-CP-FTMW, LA-MB-FTMW and DR-LA-MB-FTMW spectroscopy*

Aiming to provide rotational signatures of the isolated AHA free from environmental disturbances, we used broadband chirped pulse Fourier transform microwave spectroscopy equipped with a laser ablation vaporization system (LA-CP-FTMW) (Bermúdez et al. 2014). In the experimental procedure, finely powdered AHA was mixed with a small amount of a commercial copolymeric binder as well as a low quantity (~ 5%) of starch (wheat amyllum), which was used as a thickening and gluing agent, and pressed into cylindrical rods, which were ablated by using the third harmonic (355 nm, 15 mJ per pulse) of an Nd-YAG picosecond laser. The vaporized products were seeded in neon at a backing pressure of 10 bar and adiabatically expanded into the vacuum chamber of the spectrometer, and then probed by laser ablation broadband chirped pulse Fourier transform microwave spectroscopy. A high-power excitation pulse of 300 W was used to polarize the molecules at frequencies from 6.8 to 18 GHz (in three different intervals: 6.8 - 14.2, 14 - 16 and 16 - 18 GHz). Up to 60000 individual free induction decays (4 FIDs on each valve cycle at a 2 Hz repetition rate) were averaged in the time domain, and then Fourier transformed to obtain the broadband frequency domain spectrum.

Then, we used our narrowband instrument (LA-CP-FTMW) to improve the quality of the measurements and further resolve the  $^{14}\text{N}$  nuclear quadrupole hyperfine structure. (see Chapter III: Methodology for additional details). Finally, looking forward to extend the measurements of the global minimum conformer (astronomical candidate) up to 40

GHz and enhance the precision of the centrifugal distortion constants, we employed our newly developed double-resonance configuration using as well the laser ablation system (DR-LA-MB-FTMW).

#### 4.3. *Microwave spectrum of glycolamide: CP-FTMW, MB-FTMW and DR-MB-FTMW spectroscopy*

A commercial sample of glycolamide was purchased from Aldrich and used without further purification (98%). It is a white crystalline substance melting at 122–126 °C. Rotational spectra measurements were carried out with three different spectroscopic techniques.

The supersonic-jet rotational spectrum in the 9.2–17.1 GHz frequency range was investigated using a broadband CP-FTMW spectrometer (Mata et al. 2012; Kolesníková et al. 2017a). A pulsed heated nozzle was used to expand a gas mixture containing the gaseous glycolamide (heated from the solid at 130 °C) seeded in neon (backing pressure of 1 bar). Up to 60 000 individual free induction decays (four FIDs on each valve cycle) were averaged in the time domain, and Fourier-transformed using a Kaiser–Bessel window function to obtain the frequency-domain spectrum.

In an attempt to fully resolve the  $^{14}\text{N}$  hyperfine structure of glycolamide, we took advantage of the higher sensitivity and resolution of our molecular beam Fourier transform microwave (MB-FTMW) spectrometer, which operates in the frequency range from 6 to 18 GHz and is described elsewhere (Balle & Flygare 1981; Alonso et al. 1997). Glycolamide was conventionally heated at 130 °C and then seeded in Ne (stagnation pressure 1 bar) and expanded adiabatically to form a supersonic jet into a Fabry–Pérot resonator. The microwave transient FID was again registered in the time domain and Fourier transformed into the frequency domain.

In a next step, we employed our double-resonance configuration (DR-MB-FTMW) to extend the coverage of the microwave spectrum of the most stable conformer of glycolamide up to 40 GHz. As we already mentioned, this new configuration is of great importance because it allows us to measure in the same frequency region covered by astronomical surveys conducted at cm wavelengths with the Effelsberg and Green Bank Telescope, and also serves as a straightforward connection between the microwave and

the millimeter-wave data. For a detailed description of the experimental methodology see Chapter III: Methodology.

#### 4.4. Millimeter- and submillimeter-wave spectrum of glycolamide

Finally, the rotational spectrum was measured from 80 to 460 GHz using the millimeter-wave absorption spectrometer at the University of Valladolid coupled with a new heating cell. A small amount of solid (0.8 g) was introduced in a single neck round bottom flask connected directly to the newly built free space cell of the spectrometer. During the experiment, an optimum gas pressure of about 15  $\mu$ bar was maintained that allowed us to record the spectra in the millimeter and submillimeter-wave region from 80 to 460 GHz. The output of the synthesizer was frequency modulated at  $f = 10.2$  kHz with modulation depth between 20 and 40 kHz. After the second pass through the cell, the signal was detected using solid-state zero-bias detectors and was then sent to a phase-sensitive lock-in amplifier with  $2f$  demodulation (time constant 30 ms), resulting in a second derivative line shape. Rotational transitions were measured using an average of two scans: one recorded in increasing and the other in decreasing frequency. The experimental uncertainty of the isolated symmetric lines is estimated to be around 50 kHz (higher than the usual 30 kHz due to the corresponding temperature broadening). This frequency-modulated spectrometer is also described in detail in Chapter III: Methodology and in Daly et al. 2014.

## 5. REFERENCES

- Alonso, J. L., Lorenzo, F. J., López, J. C., et al. 1997, *Chem. Phys.*, 218, 267
- Alonso, J. L., Lozoya, M. A., Peña, I. et al., *Chem. Sci.* 2014, 5, 515–522.
- Alonso, J. L. and López, J. C. *In Gas-Phase IR Spectroscopy and Structure of Biological Molecules*, ed. A. M. Rijs & J. Oomens (Cham: Springer International Publishing), 2015. 335.
- Altwegg, K.; Balsiger, H.; Bar-Nun, A.; et al. 2016, *Sci. Adv.*, 2, e1600285.
- Balle, T. J., & Flygare, W. H. 1981, *Rev. Sci. Instrum.*, 52, 33
- Belloche, A., Menten, K. M., Comito, C., et al. 2008, *A&A*, 482, 179.
- Belloche, A., Müller, H. S. P., Menten, K. M., Schilke, P., & Comito, C. 2013, *A&A*, 559, A47.
- Belloche, A., Müller, H. S. P., Garrod, R. T., & Menten, K. M. 2016, *A&A*, 587, A91.
- Belloche, A., Meshcheryakov, A. A., Garrod, R. T., et al. 2017, *A&A*, 601, A49 and references therein.
- Belloche, A., Garrod, R. T., Müller, H. S. P., et al. 2019, *A&A*, 628, A10. Brown, R. D



- Bermúdez, C., Mata, S., Cabezas, C., Alonso, J.L., *Angew. Chem. Int. Ed.* 2014, 53, 11015–11018.
- Blagojevic, V.; Petrie, S.; Bohme, D. K. 2003, *Mon. Not. R. Astron. Soc.* 339, L7–L11.
- Bonfand, M., Belloche, A., Garrod, R. T., et al. 2019, *A&A*, 628, A27.
- Bracher, B. H., Small, R. W. H., *Acta Cryst.* 1970, B26, 1705-1709.
- Brown, R. D., Godfrey, P. D., Ottrey, A. L., et al. 1977, *J. Mol. Spectrosc.*, 68, 52 359.
- Burton, A. S.; Stern, J. C.; Elsila, J. E.; Glavin, D. P.; Dworkin, J. P., 2012, *Chem. Soc. Rev.*, 41, 5459-5472.
- Butler R. A. H., De Lucia F. C., Petkie D. T., Møllendal H., Horn A., Herbst E., 2001, *ApJS*, 134, 319.
- Cabezas, C., Alonso, J.L., López, J.C., Mata, S. 2012, *Angew. Chem.* 1404-1407.
- Cabezas, C., Varela, M., Alonso, J. L. 2017, *Angew. Chem. Int. Ed*, 56, 6420.
- Cernicharo, J., Kisiel, Z., Tercero B., et al. 2016, *A&A*, 587, L4, 44.
- Cernicharo, J., Cabezas, C., Angúdez, M., et al. 2021, *A&A*, 648, L3.
- Chengalur, J. N., & Kanekar, N. 2003, *A&A*, 403, L43.
- Codd, R., *Coord. Chem. Rev.* 2008, 252, 1387– 1408.
- Colzi, L., Rivilla, V.M., Beltrán, M.T. et al. *A&A*, 2021, 653, A129.
- Comito, C., Schilke, P., Rolffs, R., et al. 2010, *A&A*, 521, L38.
- Cunningham, M. R., Jones, P. A., Godfrey, P. D. et al. 2007, *Mon. Not. R. Astron. Soc.*, 376, 1201–1210.
- Daly, A., Kolesniková, L. Mata, S., J. Alonso, J. L. 2014, *Journal of Molecular Spectroscopy*, 306, 11.
- Ehrenfreund, P.; Bernstein, M.; Dworkin, J. P.; Sandford, S. A.; Allamandola, L. 2001, *Astrophys. J.*, 550, L95-L99.
- Endres, C. P. Schlemmer, S., Schilke, P., et al. 2016, *J. Mol. Spectrosc.*, 327, 95.
- Farkas, E.; Enyedy, E. A.; Zekany, L.; Deak, G., *J. Inorg. Biochem.* 2001, 83, 107–114.
- Faure, A., Remijan, A. J., Szalewicz, K., et al. 2014, *ApJ*, 783, 72.
- Faure, A., Lique, F., & Remijan, A. J. 2018, *The Journal of Physical Chemistry Letters*, 9, 3199
- Foley, H. M. 1947, *Physical Review*, 71, 747.
- Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16, Revision B.01*; Gaussian, Inc.: Wallingford CT, 2016.
- Garrod, R. T. 2013, *ApJ*, 765, 60.
- Garrod, R. T., Belloche, A., Müller, H. S. P., et al. 2017, *A&A*, 601, A48.
- Ginsburg, A., Henkel, C., Ao, Y., et al. 2016, *A&A*, 586, A50.
- Gordy, W., Cook, R. L. *Microwave Molecular spectra*, Interscience Pub, 1970, New York, Chap. III, 58-70

- Gottlieb, C. A., Ball, J. A., Gottlieb, E. W., et al. 1979, *ApJ*, 227, 422.
- Griffith, D. P.; Gleeson, M. J.; Lee, H.; Longuet, R.; Deman, E.; Earle, N., *Eur. Urol.* 2017, 20, 243–247.
- Gupta, S. P., *Hydroxamic Acids. A Unique Family of Chemicals with Multiple Biological Activities*; Ed. Springer-Verlag: Berlin, Heidelberg, 2013.
- Halfen, D. T., Ilyushin, V. Ziurys, L. L., 2011, *ApJ*, 743, 60.
- Hartwig, H. & Dreizler, H. 1996, *Zeitschrift Naturforschung Teil A*, 51, 923.
- Haykal, I., Motiyenko, R. A., Margulès, L., et al. 2013, *A&A*, 549, A96.
- He, J.; Vidali, G.; Lemaire, J.-M.; Garrod, R. T., 2015. *Astrophys. J.*, 799:49 (9pp).
- Herbst, E. 2001, *Chem. Soc. Rev.*, 30, 168-176.
- Hollis, J. M.; Pedelty, J. A.; Snyder, L. E.; Jewell, P. R.; Lovas, F. J.; Palmer, P.; Liu, S.-Y. 2003, *Astrophys. J.*, 588, 353–359.
- Hollis, J. M., Lovas, F. J., Jewell, P. R., et al. 2002, *ApJ*, 571, L59.
- Hollis, J. M., Pedelty, J. A., Snyder, L. E., et al. 2003, *ApJ*, 588, 353.
- Hollis, J. M., Jewell, P. R., Lovas, F. J., et al. 2004a, *ApJ*, 610, L21
- Hollis, J. M., Jewell, P. R., Lovas, F. J., et al. 2004b, *ApJ*, 613, L45
- Hollis, J. M., Lovas, F. J., Remijan, A. J., et al. 2006, *ApJ*, 643, L25.
- Hollis, J. M., Jewell, P. R., Remijan, A. J., et al. 2007, *ApJ*, 660, L125
- Ilyushin, V. V., Alekseev, E. A., Dyubko, S. F., et al. 2004, *Journal of Molecular Spectroscopy*, 227, 115.
- Irvine, W. 1998, *Orig Life Evol Biosph.*, 28, 365-383.
- Jeanvoine, Y.; Largo, A.; Hase, W. L.; Spezia, R. 2018, *J. Phys. Chem. A.*, 122, 869–877.
- Jobst, K. J.; Ruttink, P.J.A.; Terlouw, J.K. 2008, *Int. J. Mass Spectrom.*, 269, 165-176.
- Johnson, A. P.; Cleaves, H. J.; Dworkin, J. P.; Glavin, D. P.; Lazcano, A.; Bada, J. L. 2008, *Science.*, 322, 404.
- Johnson, T. J., Sams, R. L., Profeta, L. T. M., et al. 2013, *J. Phys. Chem. A*, 117,4096.
- Jones, P. A., Cunningham, M. R., Godfrey, P. D., et al. 2007, *MNRAS*, 374, 579.
- Jones, P. A., Burton, M. G., Cunningham, M. R., et al. 2008, *MNRAS*, 386, 117.
- Jones, P. A., Burton, M. G., Cunningham, M. R., et al. 2012, *MNRAS*, 419, 2961.
- Klein, B., Hochgürtel, S., Krämer, I., et al. 2012, *A&A*, 542, L3.
- Kleiner, I., Hougen, J. T., Suenram, R. D., et al. 1991, *Journal of Molecular Spectroscopy*, 148, 38.
- Kleiner, I., Hougen, J. T., Suenram, R. D., et al. 1992, *Journal of Molecular Spectroscopy*, 153, 578.
- Kleiner, I., Lovas, F. J., & Godefroid, M. 1996, *Journal of Physical and Chemical Reference Data*, 25, 1113.
- Kolesniková, L., Alonso, E. R., Mata, S., et al. 2017a, *ApJS*, 233, 24.
- Largo, L.; Rayon, V. M.; Barrientos, C.; Largo, A.; Redondo, P. 2009, *Chem, Phys. Lett.*, 476, 174-177.
- Lattalais, M.; Pauzat, F.; Pilmé, J.; Ellinger, Y.; Ceccarelli, C. 2011, *Astron. Astrophys.*, 532, A39 (7pp).
- Maret, S., Hily-Blant, P., Pety, J., Bardeau, S., & Reynier, E. 2011, *A&A*, 526, A47.
- Margulès, L., Motiyenko, R. A., Ilyushin, V. V., et al. 2015, *A&A*, 579, A46.
- Margulès, L., Belloche, A., Müller, H. S. P., et al. 2016, *A&A*, 590, A93.

- Maris, A. 2004, *Phys. Chem. Chem. Phys.*, *6*, 2611-2616.
- Marmion, C. J.; Griffith, D.; Nolan, K. B., *Eur. J. Inorg. Chem.*, 2004, 3003–3016.
- Mata, S., Pena, I., Cabezas, C., & Alonso, J. L., 2012, *J. Mol. Spectr.*, *280*, 91
- McCarthy, M. C., & McGuire, B. A. 2021, *JPCA*, *125*, 3231 .
- McCarthy, M. C., Kelvin Lee, K. L., Loomis, R. A., et al. 2021, *NatAs*, *5*, 176
- McGuire, B. A. 2018, *ApJS*, *239*,
- McGuire, B. A., Loomis, R. A., Burkhardt, A. M., et al., 2021, *Science*, *371*,1265.
- 17.
- Mehring, D. M.; Snyder, L. E.; Miao, Y.; Lovas, F. J. 1997, *Astrophys. J.*, *480*, L71–L74.
- Menten, K. M. 2004, *The Dense Interstellar Medium in Galaxies*, *69*.
- Miller, S. L. 1953, *Science*, *117*, 528-529.
- Miller, M. J., *Chem. Rev.* 1989, *89*, 1563– 1579.
- Montgomery Jr., J. A.; Frisch, M. J.; Ochterski, J.W.; Petersson, G.A. 2000, *J. Chem. Phys.*, *112*, 6532-6542.
- Morris, M., & Serabyn, E. 1996, *ARA&A*, *34*, 645
- Motiyenko, R. A., Tercero, B., Cernicharo, J., et al. 2012, *A&A*, *548*, A71 and references therein.
- Müller, H. S. P., Belloche, A., Xu, L.-H., et al. 2016, *A&A*, *587*, A92.
- Muri, E. M.; Nieto, J.; Sindelar, R. D.; Williamson, J. S. *Curr. Med. Chem.* 2002, *9*, 1631–1653
- Nuevo, M., Bredehöft, U. J., Meierhenrich, U. J., D’Hendecourt, L., et al. 2011, *Astrobiology*, *10*, 245.
- Pal, D.; Saha, S., *J. Adv. Pharm. Technol. Res.* 2012, *3*, 92–99.
- Pickett, H. M. 1991, *J. Mol. Spectrosc.*, *148*, 371–377.
- Pickett, H.M., Poynter, R.L., Cohen, E.A. et al.1998, *J. Quant. Spectr. Rad. Transf.*, *60*, 883
- Pizzarello, S. 2006, *Acc. Chem. Res.*, *39*, 231-237.
- Pizzarello, S.; Shock, E. 2010, *Cold Spring Harbor Perspect. Biol.*, *2*, No. a002105 (19pp).
- Remijan, A.; Snyder, L. E.; Liu, S.-Y.; Mehringer, D.; Kuan, Y.-J. 2002, *Astrophys. J.*, *576*, 264–273.
- Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2019, *ApJ*, *885*, 131.
- Requena-Torres, M. A., Martín-Pintado, J., Rodríguez-Franco, A., et al. 2006, *130 A&A*, *455*, 971 131.
- Requena-Torres, M. A., Martín-Pintado, J., Martín, S., et al. 2008, *ApJ*, *672*, 352
- Rivilla, V. M., Martín-Pintado, J., Jiménez-Serra, I., et al. 2020, *ApJ*, *899*, L28.
- Robinson, G. W., & Cornwell, C. D. 1953, *J. Chem. Phys.*, *21*, 1436.
- Rodríguez-Almeida, L.F., Rivilla, V.M., Jiménez-Serra, I. et al., 2021, *A&A*, Forthcoming article, DOI: <https://doi.org/10.1051/0004-6361/202141989> and references therein.
- Rubin, R., Swenson Jr, G., Benson, R., Tigelaar, H., & Flygare, W. 1971, *ApJ*, *169*, L39.
- Ruiz-Mirazo, K.; Briones, C.; de la Escosura, 2014, *Chem. Rev.*, *114*, 285–366
- Saldyka, M.; Mielke, Z., 2003, *Polish J. Chem.*, *77*, 1587–1598.
- Saldyka, M.; Mielke, Z., 2007, *Vib. Spectrosc.*, *45*, 46–54.

Sałyda, M.; Mielke, Z., 2018, *J. Phys. Chem. A.*, 122, 1, 60-71

Schmiedeke, A., Schilke, P., Möller, T., et al. 2016, *A&A*, 588, A143.

Snyder, L. E.; Lovas, F. J.; Hollis, J. M.; Friedel, D. N.; Jewell, P. R.; Remijan, A. 2005, *Astrophys. J.*, 619, 914–930.

Stanton, J.F.; Gauss, J.; Harding, M.E.; Szalay, P.G. *CFOUR, A Quantum Chemical Program Package*, 2013.

Szóri, M., Jójárt, B., Izsák, R., et al. 2011, *Physical Chemistry Chemical Physics* (Incorporating Faraday Transactions), 13, 7449.

Thiel, V., Belloche, A., Menten, K. M., et al. 2019, *A&A*, 623, A68.

Turner, B. E. 1991, *ApJS*, 76, 617.

Watson, J. K. G. 1977, *J. Mol. Spectrosc.*, 66, 500.

Widicus Weaver, S. L., Butler, R. A. H., Drouin, B. J., et al. 2005, *ApJS*, 158, 188

## CHAPTER V. HUNTING COMPLEX CYANIDES IN THE INTERSTELLAR MEDIUM

---

Adapted from Sanz-Novo et al. 2020, *Astronomy & Astrophysics*, 639, A135 and Sanz-Novo et al. 2021, *The Astrophysical Journal*, 915:76 (8pp).



Cyanoacetamide ( $\text{CH}_2(\text{CN})\text{C}(\text{O})\text{NH}_2$ ) and cyanoacetic acid ( $\text{CH}_2(\text{CN})\text{C}(\text{O})\text{OH}$ ) are both complex -CN bearing molecules from an astronomical point of view, which appear as very appealing interstellar candidates. In this chapter, our investigation aims to provide direct experimental frequencies of the ground vibrational state of both molecular systems in the microwave region, in order to enable their eventual identification in the ISM using low-frequency surveys. We employed a battery of state-of-the-art high-resolution spectroscopic techniques in the time domain to measure the frequencies of both species. We report the first rotational characterization of cyanoacetamide and cyanoacetic acid, whose rotational features have remained unknown until now. A precise set of the relevant rotational spectroscopic constants have been respectively determined, including the  $^{14}\text{N}$  nuclear quadrupole coupling constants, as a first step to identifying these molecules in the interstellar medium. In addition, we fully explored the potential energy surface to study a gas-phase reaction on the formation process of protonated cyanoacetamide. We found an exothermic process with no net activation barrier, feasible under interstellar conditions, leading to protonated cyanoacetamide.

## 1. INTRODUCTION

In the early 40's, the diatomic cyano radical, CN, was the second molecule to be identified in the ISM (McKellar 1940). CN's first rotational transition ( $J = 1 \leftarrow 0$ ) was detected 30 years later in the Orion Nebula and W<sub>51</sub> (Jefferts et al. 1970). Since then, more than 30 molecules containing the cyano ( $-C\equiv N$ ) functional group have been observed to date in diverse types of astronomical environments. Among these are, for instance, the well-known family of cyanopolyynes,  $H-(C\equiv C)_n-C\equiv N$ , which are commonly found in cold interstellar clouds (Broten et al. 1978; Bell et al. 1997), in exoplanetary envelopes, such as the carbon- and nitrogen-rich atmosphere of Titan (Kunde et al. 1981; Coustenis et al. 1991), and in circumstellar envelopes around the stars (Winnewisser & Walmsley 1978; Pardo et al. 2005). In addition, dicyanopolyynes, defined as molecules that contain two cyano groups,  $N\equiv C-(C\equiv C)_n-C\equiv N$ , could also be abundant in interstellar and circumstellar clouds (Kołos & Grabowski 2000; Petrie et al. 2003). Very recently, isocyanogen (CNCN), the simplest member of this series, has been detected in the interstellar medium (ISM) (Agúndez et al. 2018). However, as we mentioned in Chapter II, many simple nitriles have yet to be discovered.

The symbiotic relationship between astronomical searches and laboratory experiments using rotational spectroscopy has allowed successfully identifying many interstellar compounds (Herbst & van Dishoeck 2009, Jorgensen 2020). In 2008, aminoacetonitrile ( $NH_2CH_2CN$ ) was discovered in Sgr B2(N) (Belloche et al. 2008) by means of rotational spectroscopy data (Bogey et al. 1990). Kolesníková et al. (2017) carried out a detailed analysis of its rotational spectra up to 480 GHz. Furthermore, acetamide ( $CH_3C(O)NH_2$ ), and N-methylformamide ( $HC(O)NHCH_3$ ), two N-bearing molecules containing a peptide-type bond ( $-NH-C(O)-$ ), have also been detected in space (Hollis et al. 2006; Belloche et al. 2017). These molecules are very interesting from a prebiotic point of view because the amide linkage is an essential feature in proteins. In current astronomical research, great efforts have been made to study more complex members of these molecular systems. This suggests that cyanoacetamide ( $CH_2(CN)C(O)NH_2$ ), an amide derivative, and also a -CN containing molecule, may exist in the ISM and motivates astronomers to search for it.

On the other hand, organic acids constitute much of the required material a planet needs to support life. These molecules have received increasing attention in the last

decade since they are essential to unveil how chemical complexity builds up. The simplest forms of these molecular classifications, formic acid ( $\text{HC(O)OH}$ ) and acetic acid ( $\text{CH}_3\text{C(O)OH}$ ) are well-known interstellar molecules (Liu et al. 2002; Lefloch et al. 2017; Favre et al. 2018; Remijan et al. 2003 and Jørgensen et al. 2016). In this context, we proposed the study of cyanoacetic acid, a relevant organic acid as well as a  $-\text{CN}$  bearing species. So far, cyanoacetic acid has been studied by IR and FTIR-matrix studies (Binev et al. 1998 and Reva et al. 2003 (a, b)) and, to the best of our knowledge, no rotational spectroscopic data is available for this molecule.

Both cyanoacetamide and cyanoacetic acid belong to an influential group of molecules entitled Complex Organic Molecules (COMs, Herbst & van Dishoeck 2009). First, cyanoacetamide is of great importance in organic chemistry and biology; it assumes a special role in heterocyclic chemistry (El Bialy et al. 2011). It is a solid in standard conditions for temperature and pressure, with a melting point of 119-121 °C. So far, no detailed experimental structural information of the isolated molecule has been reported. Only an assignment of its IR and Raman spectra has been reported quite a long time ago by Haverbeke & Herman (1975), but its conformational panorama remains unknown. Fourier transform (FT) microwave spectroscopy is extremely sensitive to molecular geometry and has been proved a powerful tool in the conformational analysis (Schermann 2008). Coupled with laser-ablation techniques, it enables an ultra-fast vaporization of organic and biological molecules and allows them to be probed in the isolated environment of a supersonic expansion (Alonso et al. 2014). With this outstanding strategy, several relevant biomolecules and most of the amino acids (Alonso et al. 2015), as well as species of astrophysical interest, such as hydantoin (Alonso et al. 2017) and glycineamide (Alonso et al. 2018), have recently been revealed.

Moreover, new gas-phase theoretical and experimental studies exploring different nitrile species should represent an essential source of reference data in constraining models of nitrile formation in the ISM and in predicting which other nitrile species might be present and detectable. The knowledge of their formation processes is essential to understand the evolution of more complex molecules. In this context, it was concluded in several experimental studies (He et al. 2015; Zheng & Kaiser 2010; Congiu et al. 2012) that hydroxylamine, which is expected to be one of the main precursors of amino acids in space, could be generated in the solid phase and then desorbed through a temperature-induced process into the gas phase. Thus, we have carried out an accompanying theoretical

study, consisting of a full exploration of the potential energy surface (PES), to analyze the feasibility of a plausible gas-phase reaction as a source of protonated cyanoacetamide in the ISM and thereby disclose the corresponding neutral form as a candidate molecule to be searched for in the ISM.

Following our previous work (see [Chapter IV: Glycolamide](#)), we propose the study of cyanoacetamide and cyanoacetic acid, as candidate molecular systems for interstellar search. Our current study aims to provide accurate rotational information, which is the initial and prerequisite step for its radio astronomical search, to correctly reproduce the spectrum and interpret astronomical data in low-frequency regions. In this work, the experimental values of the  $^{14}\text{N}$  nuclear quadrupole coupling constants together with the rotational constants are provided for both species. For observations in low-frequency regions such as those available for the Green Bank Telescope (GBT) or the Effelsberg 100 m Radio Telescope, the hyperfine structure could be a drawback for interpreting the astronomical spectra. The Effelsberg observatory, for instance, has the capability to search for transitions in the microwave region such as the lines we aim to provide in this work (see [Chapter IV: Glycolamide](#)). For some interstellar molecules such as cyanoallene and methanimine, the interpretation of the hyperfine pattern was moreover crucial for its conclusive identification in the ISM (Lovas et al. 2006; Godfrey et al. 1973).

This motivated us to perform high-resolution state-of-the-art experiments to achieve the first rotational characterization of cyanoacetamide and cyanoacetic acid. In addition, we have carried out a theoretical study of the gas-phase formation process of protonated cyanoacetamide from the reaction of protonated hydroxylamine and cyanoacetaldehyde. To the best of our knowledge, no theoretical studies have been devoted to proposing potential mechanisms of cyanoacetamide formation in the ISM. The astrochemical interpretation of these might pave the route to the elucidation of models of nitrile-derivatives formation in space and may act as a source of reference data to reveal the evolution of the more complex molecules.



## 2. RESULTS AND DISCUSSION

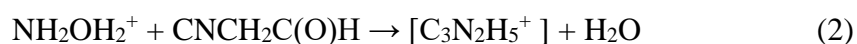
### 2.1. *Rotational and computational study of cyanoacetamide*

#### 2.1.1. *Formation process of protonated cyanoacetamide*

In a first step of the investigation, we studied the gas-phase formation process of protonated cyanoacetamide  $[\text{C}_3\text{N}_2\text{H}_5]^+$  from the reaction of protonated hydroxylamine and cyanoacetaldehyde ( $\text{CNCH}_2\text{C}(\text{O})\text{H}$ ). We therefore carried out a full exploration of the singlet potential energy surfaces because all the reactants are in their singlet electronic state ( $^1\text{A}'$ ). Our main goal in this section is to analyze whether the reactions we studied might be a source of protonated cyanoacetamide, and we discuss the results in the context of the ISM. In conditions of the ISM, basically low temperatures and low densities, interstellar viable reactions should be exothermic and barrier-free processes. Only a few protonated complex organic molecules (COMs) have been detected in the ISM so far, although it seems that they could evolve to the corresponding neutral species by dissociative recombination. According to Snow et al. (2007), the two isomers of protonated hydroxylamine ( $\text{NH}_3\text{OH}^+$  and  $\text{NH}_2\text{OH}_2^+$ ) might be obtained in the protonation process of hydroxylamine using  $\text{CH}_5^+$  as protonating agent. Neither of them reacts with  $\text{H}_2$  under interstellar conditions (Largo et al. 2009), and if they exist in the ISM, they may therefore be able to react with other molecules, for instance, with cyanoacetaldehyde. Despite cyanoacetaldehyde have not been detected in space to date, very recently Rivilla et al. (2020) reported the first detection of hydroxylamine ( $\text{NH}_2\text{OH}$ ) in the ISM toward the quiescent molecular cloud G+0.693-0.027 located in the Galactic Center. Moreover, hydroxylamine has been produced by ultraviolet laser irradiation of ice mixtures simulating interstellar grains (Nishi et al. 1984). In addition, formation of hydroxylamine was observed in electron-irradiated ammonia-water ices at temperatures of 10 and 50 K (Zheng & Kaiser 2010) and also in electron-irradiated ammonia-oxygen ices at astrophysically relevant temperatures of 5.5 K (Tsegaw et al. 2017).

The rotational spectrum and conformational composition of cyanoacetaldehyde has previously been reported by Møllendal et al. (2012). Its existence in interstellar space and in planetary atmospheres is quite likely because of the facile formation of cyanoacetylene in liquid water (Ferris et al. 1968), presumably formed in a catalytic process (Horn et al. 2008). It is known that cyanoacetylene, a compound formed in a spark-discharge reaction

in a mixture of methane and nitrogen (Sánchez et al. 1967), is widespread in the ISM, and found in comets (BockeléeMorvan et al. 2000) and in the atmosphere of Titan (Coustenis et al. 1999). However, several gas-phase models of an uncatalyzed addition reaction between the two neutral molecules cyanoacetylene and water as well as the addition of water to protonated cyanoacetylene, which also exists in the ISM (Kawaguchi et al. 1994), followed by an electron recombination reaction have been predicted to be inefficient by quantum-chemical calculations (Horn et al. 2008). Furthermore, we studied a plausible formation route showing that cyanoacetaldehyde could be formed by the addition of the O (<sup>3</sup>P) radical to vinyl cyanide, which is a well-known interstellar molecule (for detailed information, see Fig. B.2 in Section B.2 of Appendix B). The reactions between protonated hydroxylamine (NH<sub>3</sub>OH<sup>+</sup> /NH<sub>2</sub>OH<sub>2</sub><sup>+</sup>) and cyanoacetaldehyde to give protonated cyanoacetamide can be summarized as follows:

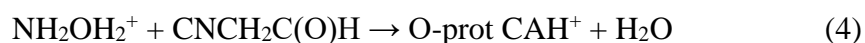
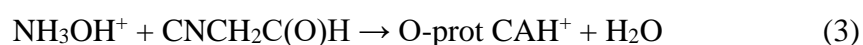


Protonation of cyanoacetamide gives two isomers with the [C<sub>3</sub>N<sub>2</sub>H<sub>5</sub><sup>+</sup>] molecular formula: O-protonated cyanoacetamide, (O-prot) CNCH<sub>2</sub>COHNH<sub>2</sub><sup>+</sup>, and N-protonated cyanoacetamide, (N-prot) CNCH<sub>2</sub>C(O)NH<sub>2</sub><sup>+</sup>. Protonation on the oxygen atom is the most favorable process because (O-prot) CNCH<sub>2</sub>COHNH<sub>2</sub><sup>+</sup> is located 12.2 kcal/mol lower in energy, at the CCSD(T) level, than the corresponding N-protonated isomer.

In Table 1 we collect the relative energies with respect to reactants of the possible products that can be formed in the reaction between either NH<sub>3</sub>OH<sup>+</sup> or NH<sub>2</sub>OH<sub>2</sub><sup>+</sup> and CNCH<sub>2</sub>C(O)H computed at different levels of theory. At the CCSD(T) level of theory, hydroxylamine protonated in its oxygen atom, NH<sub>2</sub>OH<sub>2</sub><sup>+</sup>, was found to lie 25.5 kcal/mol higher in energy than the most stable isomer, NH<sub>3</sub>OH<sup>+</sup>. This value therefore is the difference of the reaction energies of paths 1 (first entry of Table 1) and 2 (second entry) computed at this level of theory.

Table 1 shows that all reactions we studied are exothermic processes. Regardless of the level of theory, the most favorable process from thermodynamic arguments is the formation of O-protonated cyanoacetamide. This might have been expected, and it mainly

arises from the different chemical natures of the isomers. The multiple topological dispositions of the hydrogen atom allow us to generate different arrangements of the -NH and -OH groups. This leads to the formation of specific intramolecular hydrogen bonds that further stabilize this molecular system. In particular, the N-H $\cdots\pi$  weak hydrogen bond is expected to be one of the leading structural motifs stabilizing the structure, as pointed out in previous studies (Cabezas et al. 2012). We focused on the reactions of protonated hydroxylamine with cyanoacetaldehyde to give the most stable isomer of protonated cyanoacetamide, (O-prot) CNCH<sub>2</sub>COHNH<sub>2</sub><sup>+</sup>, from now on O-prot CAH<sup>+</sup>:



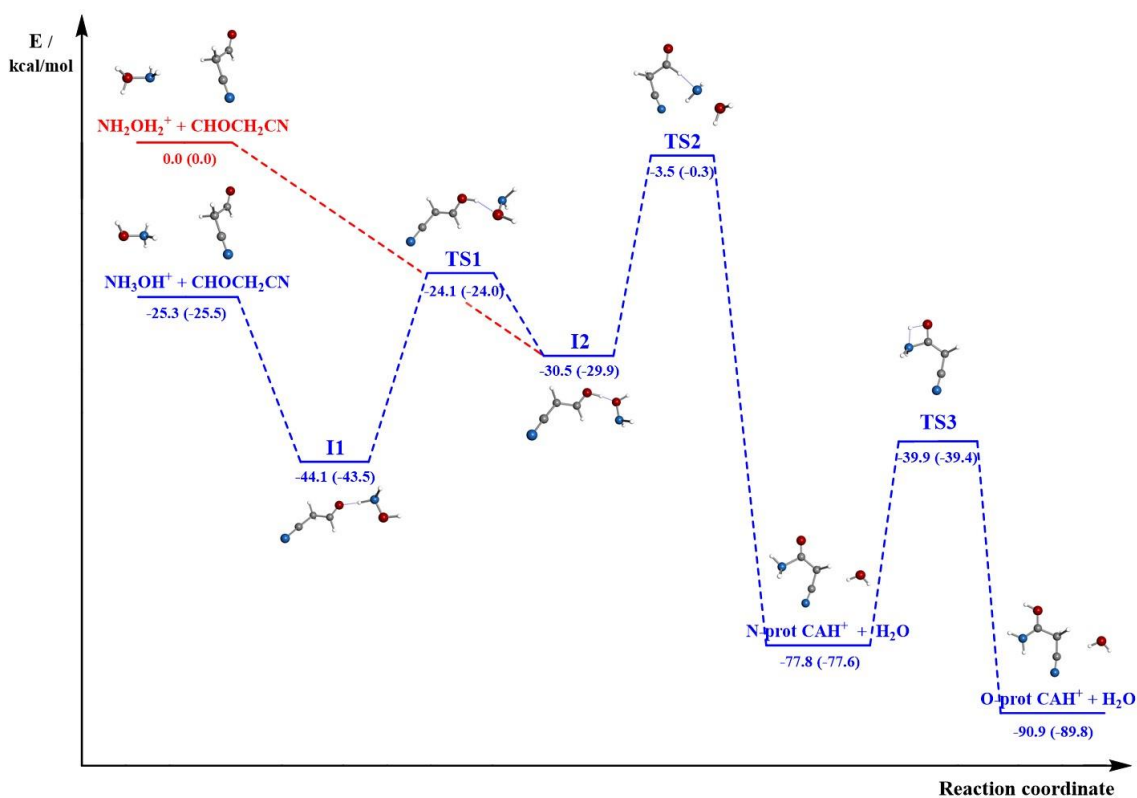
The energy profiles for both reactions were computed at B2PLYPD3 and CCSD(T) levels of theory, and they are depicted in Fig. 1. This figure shows that the results obtained with the B2PLYPD3 and CCSD(T) methods agree well. For reaction of the most stable isomer of protonated hydroxylamine (NH<sub>3</sub>OH<sup>+</sup>), the process starts with the formation of the I1 intermediate, which is the result of the interaction between the carbonylic oxygen of cyanoacetaldehyde with one of the hydrogen atoms of the NH<sub>3</sub> group of protonated hydroxylamine.

For reaction of the most stable isomer of protonated hydroxylamine (NH<sub>3</sub>OH<sup>+</sup>), the process starts with the formation of the I1 intermediate, which is the result of the interaction between the carbonylic oxygen of cyanoacetaldehyde with one of the hydrogen atoms of the NH<sub>3</sub> group of protonated hydroxylamine.

**Table 1.** Relative energies (referred to reactants), including zero-point corrections, in kcal/mol, obtained at different levels of theory for the reaction between protonated hydroxylamine and cyanoacetaldehyde yielding protonated cyanoacetamide.

Product <sup>(a)</sup>	B3LYP <sup>(b)</sup>	MP2 <sup>(c)</sup>	B2PLYPD3 <sup>(d)</sup>	CCSDT <sup>(e)</sup>
(O-prot) CNCH <sub>2</sub> COHNH <sub>2</sub> <sup>+</sup> + H <sub>2</sub> O	-90.0	-93.6	-90.9	-89.8
(O-prot) CNCH <sub>2</sub> COHNH <sub>2</sub> <sup>+</sup> + H <sub>2</sub> O	-65.0	-67.3	-65.6	-64.3
(N-prot) CNCH <sub>2</sub> CONH <sub>3</sub> <sup>+</sup> + H <sub>2</sub> O	-76.2	-82.1	-77.8	-77.6
(N-prot) CNCH <sub>2</sub> CONH <sub>3</sub> <sup>+</sup> + H <sub>2</sub> O	-51.2	-55.8	-52.5	-52.1

**Notes.** <sup>(a)</sup> The first entry refers to the NH<sub>2</sub>OH<sub>2</sub><sup>+</sup> + CNCH<sub>2</sub>C(O)H reaction and the second to the NH<sub>3</sub>OH<sup>+</sup> + CNCH<sub>2</sub>C(O)H reaction. <sup>(b)</sup> Electronic energy calculated at the B3LYP/aug-cc-pVTZ level with the Grimme D3 dispersion factor. <sup>(c)</sup> Electronic energy calculated at the MP2/aug-cc-pVTZ level. <sup>(d)</sup> Electronic energy calculated at the B2PLYPD3/aug-cc-pVTZ level. <sup>(e)</sup> Electronic Energy calculated at the CCSD/aug-cc-pVTZ//B2PLPD3/aug-cc-pVTZ level.



**Fig. 1.** Energy profile in kcal/mol for the reactions of protonated hydroxylamine with cyanoacetaldehyde producing O-protonated cyanoacetamide computed at the B2PLYPD3/aug-cc-pVTZ and CCSD(T)/aug-cc-pVTZ (in parentheses) levels of theory. The zero-point vibrational energy computed at the B2PLYPD3/aug-cc-pVTZ level is included.

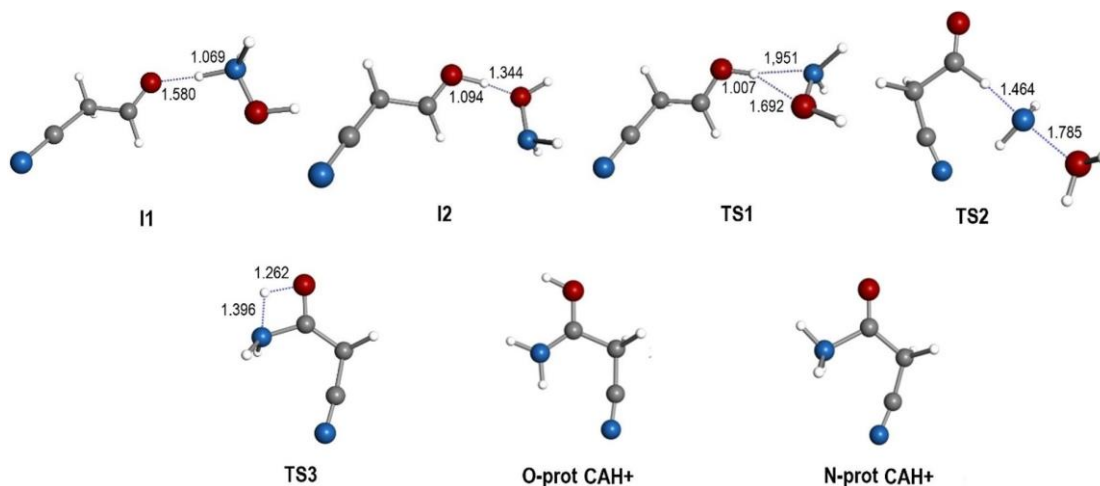
The exothermic formation of the I1 intermediate ( $\Delta E = -18.0$  kcal/mol at the CCSD(T) level) produces an energy reservoir that is used as the reaction proceeds toward the products. After I1 is formed, a hydrogen atom migrates from nitrogen to oxygen through the TS1 transition state, which is located 1.5 kcal/mol (755 K) above reactants (at the CCSD(T) level), giving the I2 intermediate. This intermediate evolves to the I3 intermediate through the TS2 transition state, which is located 25.3 kcal/mol (12681 K), at the CCSD(T) level, above reactants. In the TS2 transition state, a water molecule is coordinated to the NH<sub>2</sub> group. Thus, when the NH<sub>2</sub> group is properly oriented, an N-C bond is formed, evolving to the formation of N-protonated cyanoacetamide and releasing a water molecule. Finally, an isomerization of the N-protonated form to the most stable isomer, kcal/mol, at the CCSD(T) level, below reactants. In the TS3 transition state, a hydrogen atom of the NH<sub>3</sub> group migrates to the carbonylic oxygen. The process can be summarized as follows:



As shown in Fig. 1, the global process is exothermic ( $\Delta E = -64.3$  kcal/mol at the CCSD(T) level), and it has a non-negligible activation barrier of 25.3 kcal/mol (12732 K) at the CCSD(T) level of theory. This barrier precludes this mechanism from taking place in the ISM. The first step in the mechanism starting from the less stable isomer of protonated hydroxylamine is the direct interaction between one of the hydrogen atoms bonded to oxygen of NH<sub>2</sub>OH<sub>2</sub><sup>+</sup> and the carbonylic oxygen of CNCH<sub>2</sub>C(O)H, giving the I2 intermediate. This complex, as in path 1, finally evolves through TS2 to produce N-protonated cyanoacetamide and H<sub>2</sub>O, as is schematized in path 2. Finally, the N-protonated form can isomerize through the TS3 transition state to give CAH<sup>+</sup> as the final product, which is located -39.4 kcal/mol, at the CCSD(T) level, below reactants:



This reaction is a clear exothermic process ( $\Delta E = -89.8$  kcal/mol at the CCSD(T) level) without a net activation barrier because now the TS2 transition state lies 0.3 kcal/mol (at the CCSD(T) level) below reactants. Thus, the formation of CAH<sup>+</sup> from the reaction of the less stable isomer of protonated hydroxylamine and cyanoacetaldehyde might be feasible under interstellar conditions. As we pointed out, although protonated COMs have not yet been detected in the ISM, it seems that they could evolve to neutral species by dissociative recombination. The neutral form, CH<sub>2</sub>(CN)C(O)NH<sub>2</sub>, might therefore be a candidate molecule to be searched for in the ISM. The optimized geometries at the B2PLYPD3/aug-cc-pVTZ level of theory for the intermediate species (I1 and I2) and transition states (TS1, TS2, and TS3) involved in the formation process of CAH<sup>+</sup> are depicted in Fig. 2.

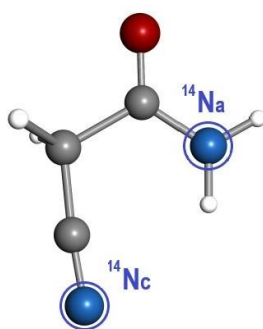


**Fig. 2.** Intermediates, transition states, and products involved on the reaction of formation of protonated cyanoacetamide. Geometries are optimized at the B3PLYPD3/aug-cc-pVTZ level of theory. Bond lengths are given in Angstroms for the intermediates and TS.

### 2.1.2. Microwave rotational spectrum of cyanoacetamide

After we found a feasible process to form protonated cyanoacetamide under interstellar conditions, we recorded the rotational spectrum of the neutral form, (CH<sub>2</sub>(CN)C(O)NH<sub>2</sub>), as a first step for its eventual radio-astronomical detection. The anticipated equilibrium structures, *trans* and *gauche* cyanoacetamide, are near prolate asymmetric tops. Recognizable patterns of  $\mu_a$ -type *R*-branch transitions separated by approximately B + C frequency intervals are expected in the rotational spectrum. However, with the aid of

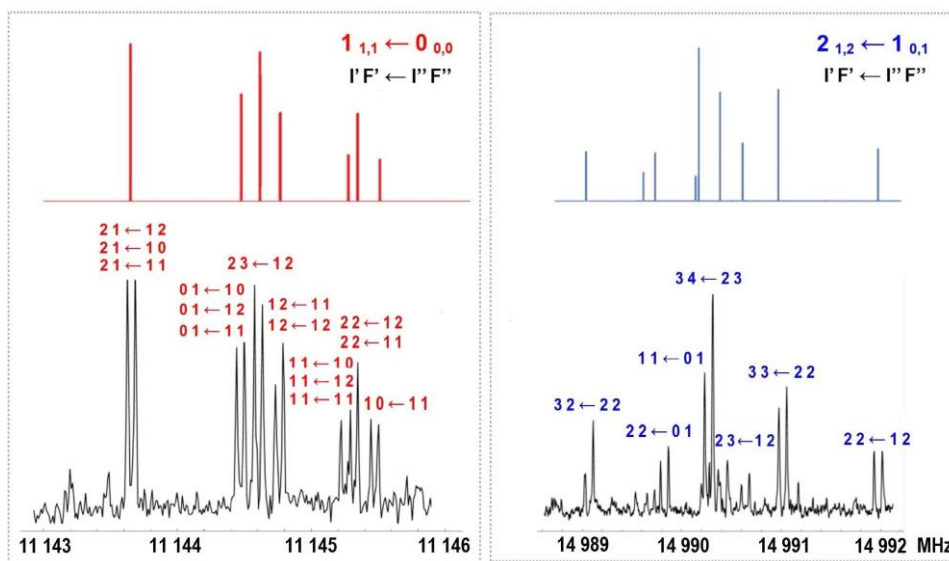
high-level B2PLYPD3 computations, it was found that only the *trans* conformer was a real equilibrium structure on the PES. A vibrational analysis for the gauche-conformer showed an imaginary frequency associated with the bending mode. We then focused on the analysis of *trans* cyanoacetamide (see Fig. 3). Cyanoacetamide is a challenging problem for high-resolution rotational studies. Its weak dipole moment ( $\mu_a = 0.3$  D and  $\mu_b = 0.5$  D) predicts very weak rotational transitions; additionally, the two nitrogen atoms divide the intensity of each rotational transition into several components due to the hyperfine structure. Therefore the experimental study of cyanoacetamide is difficult.



**Fig. 3.** Structure of *trans* cyanoacetamide, showing the two different  $^{14}\text{N}$  nuclei: an amide ( $\text{N}_a$ ) and a cyano ( $\text{N}_c$ ) nitrogen atom.

At the first stage of the line assignment, we scanned long frequency intervals of about 100 MHz around the predicted  $\mu_a$ -type *R*-branch rotational transitions (with  $J' \leftarrow J'' = 2 \leftarrow 1$  and  $3 \leftarrow 2$ ). Then, the polarization power was optimized according to the really low  $\mu_a$  electric dipole-moment component ( $\mu_a = 0.3$  D). Finally, the  $2_{02} \leftarrow 1_{01}$  and  $3_{03} \leftarrow 2_{02}$  transitions were identified. Furthermore, all the observed rotational transitions display a complex hyperfine structure as a result of the two different  $^{14}\text{N}$  nuclei (see Figure 3). This structure arises from the interaction of the electric quadrupole moment of the two  $^{14}\text{N}$  ( $I = 1$ ) nuclei, with the electric field gradient created at the site of the quadrupolar nucleus by the remaining electronic and nuclear charges of the molecule. This interaction results in a coupling between the  $^{14}\text{N}$  nuclear spin and the overall angular momentum, which is shown as a characteristic hyperfine pattern in the rotational spectra (Gordy & Cook 1970). This coupling moreover leads to a splitting in the rotational energy levels that decreases the overall intensity of each rotational transition, which are already very weak because of the low dipole moment component. The initial assignment was then extended to other

intense  $\mu_b$ -type transitions. After iterative fitting and prediction, a total of 28 hyperfine components were collected and fit (Pickett et al. 1998) to a rigid rotor Hamiltonian,  $H_R$ , supplemented with a term to take into account the quadrupole interaction  $H_Q$  (Foley 1947; Robinson & Cornwell 1953), namely  $H = H_R + H_Q$ . The quadrupole coupling Hamiltonian was set up in the coupled basis set ( $I_1, I_2, I, J, K$  and  $F$ ), where  $I_1 + I_2 = I$ , and  $I + J = F$ . The energy levels involved in each transition were thus labeled with the quantum numbers,  $K_a, K_c, I$ , and  $F$ . The resolution attained with our broadband LA-MB-FTMW technique was sufficient to fully resolve the hyperfine patterns. However, only the diagonal elements ( $\chi_{aa}, \chi_{bb}$ , and  $\chi_{cc}$ ) of the quadrupole coupling tensor  $\chi$  were determined in the fit of the observed transitions within the estimated accuracy of the frequency measurements. The derived experimental and theoretical rotational and quadrupole coupling constants for *trans* cyanoacetamide are listed in Table 2, and the measured rotational transitions are collected in Table 3.



**Fig. 4.** Rotational transitions  $1_{11} \leftarrow 0_{00}$  and  $2_{12} \leftarrow 1_{01}$  of *trans* cyanoacetamide observed during the LA-MB-FTMW experiment along with the B2PLYPD3/aug-cc-pVTZ predictions, displaying the high resolution of our spectrometer as well as the great agreement between the experimental and theoretically predicted frequencies. The resonance frequency is determined by the arithmetic mean of two Doppler components. The LA-MB-FTMW spectrum was obtained by averaging 200 experimental cycles (six free induction decay signals per cycle).

The first piece of evidence in the identification of the observed structure comes from the values of the rotational constants, which critically depend on the mass distribution of the molecule. Table 2 shows that the spectroscopic parameters predicted for the *trans*



cyanoacetamide at the B2PLYPD3/aug-cc-pVTZ level are very close to those experimentally determined. In this case, another interesting spectroscopic tool is taken into account: the two  $^{14}\text{N}$  nuclei in our molecule, which are also helpful to identify structures. In our predictions, we included the values of the quadrupole coupling constants for both  $^{14}\text{N}_c$  (cyano) and  $^{14}\text{N}_a$  (amide) nuclei. It is known that the nuclear quadrupole-coupling interactions strongly depend on the electronic environment, position, and orientation of the  $^{14}\text{N}$  nuclei. The  $^{14}\text{N}$  nuclei introduce hyperfine rotational probes at defined sites of cyanoacetamide and act as a probe of the chemical environment of both  $\text{N}_c$  and  $\text{N}_a$  quadrupolar nuclei. When we compare the experimental and theoretical rotational constants, factors ranging from 0.9999 to 1.0066 bring the values of predicted rotational constants of Table 2 to those experimentally obtained. Thus, we can infer that its actual geometry should be close to that calculated at the B2PLYPD3 level. The excellent match between the experimental and theoretical values shown in Table 2 enables the unambiguous identification of the observed form as *trans* cyanoacetamide. In conclusion, the microwave spectrum of the most stable form of cyanoacetamide has been studied up to 18 GHz to provide accurate laboratory reference spectra that can be used with confidence in future searches for cyanoacetamide in several low-frequency surveys.

**Table 2.** Theoretical and experimental spectroscopic parameters for *trans* cyanoacetamide.

Parameters	Experiment	Theory <sup>(a)</sup>
$A^{(b)}$ (MHz)	9221.8945 (33) <sup>(h)</sup>	9283.3
$B$ (MHz)	2390.8717 (17)	2389.5
$C$ (MHz)	1922.8523 (13)	1922.7
$P_c^{(c)}$ (MHz)	1.6764 (62)	1.545
$ \mu_a ,  \mu_b ,  \mu_c ^{(d)}$ (D)	Yes / Yes / No	0.3 / 0.6 / 0.0
$^{14}\text{N}_c \chi_{aa}^{(e)}$ (MHz)	-2.754 (19)	-2.872
$^{14}\text{N}_c \chi_{bb} - \chi_{cc}$ (MHz)	-1.209 (12)	-1.260
$^{14}\text{N}_a \chi_{aa}$ (MHz)	1.587 (25)	1.683
$^{14}\text{N}_a \chi_{bb} - \chi_{cc}$	5.815 (67)	6.007
$N^{(f)}$	28	-
$\sigma^{(g)}$ (kHz)	6.4	-

**Notes.** <sup>(a)</sup> Theoretical calculations at B2PLYPD3/aug-cc-pVTZ level of theory. <sup>(b)</sup> A, B, and C represent the rotational constants. <sup>(c)</sup>  $P_c$  is the planar inertial moment (in  $\text{u}\text{\AA}^2$ ), conversion factor: 505379.1 MHz– $\text{u}\text{\AA}^2$ . <sup>(d)</sup>  $|\mu_a|, |\mu_b|, |\mu_c|$  are the absolute values of the electric dipole moment components (in D). <sup>(e)</sup>  $\chi_{aa}, \chi_{bb}, \chi_{cc}$  are the diagonal elements of the  $^{14}\text{N}$  nuclear quadrupole coupling tensor. <sup>(f)</sup> N is the number of measured hyperfine components. <sup>(g)</sup>  $\sigma$  is the root mean square (rms) deviation of the fit. <sup>(h)</sup> Standard error in parentheses in units of the last digit.

Finally, we provide the rotational ( $Q_r$ ) and vibrational ( $Q_v$ ) partition functions of cyanoacetamide, which are shown in Table 4. We calculated the values of  $Q_r$  from first principles at different temperatures, using the Picket program (Pickett et al. 1998). We estimated the vibrational part,  $Q_v$ , using an anharmonic approximation and a simple formula that corresponds to Eq. 3.60 of Gordy & Cook (1970), where only the ten lowest vibrational modes were taken into account. We obtained the frequencies of the normal modes from double-hybrid calculations of both harmonic and anharmonic approximations, respectively, using the B2PLYPD3 method and the aug-cc-pVTZ basis set (see Table B.3 of Appendix B). The full partition function,  $Q_{tot}$ , is therefore the product of  $Q_r$  and  $Q_v$ . Finally, the complete list of transitions of the calculated reference spectra is available at the CDS via anonymous ftp to [cdsarc.u-strasbg.fr](ftp://cdsarc.u-strasbg.fr).

**Table 3.** Complete list of the measured transition frequencies for the ground-state of cyanoacetamide.

J'	K <sub>a</sub> '	K <sub>c</sub> '	I'	F'	J''	K <sub>a</sub> ''	K <sub>c</sub> ''	I''	F''	$\nu_{obs}^a$ (MHz)	$\nu_{calc}^b$ (MHz)	$\nu_{obs}-\nu_{calc}^c$ (MHz)
1	1	1	2	3	0	0	0	1	2	11144.605	11144.602	0.003
1	1	1	2	1	0	0	0	1	2	11143.657	11143.653	0.004
1	1	1	0	1	0	0	0	1	0	11144.473	11144.473	0.000
1	1	1	1	2	0	0	0	1	1	11144.761	11144.757	0.004
1	1	1	1	1	0	0	0	1	0	11145.250	11145.248	0.002
1	1	1	2	2	0	0	0	1	1	11145.318	11145.313	0.005
1	1	1	2	0	0	0	0	1	1	11145.474	11145.468	0.006
2	1	2	3	2	1	0	1	2	2	14989.067	14989.068	-0.001
2	1	2	3	2	1	0	1	0	1	14989.825	14989.827	-0.002
2	1	2	3	4	1	0	1	2	3	14990.270	14990.270	0.000
2	1	2	3	3	1	0	1	1	2	14990.421	14990.426	-0.005
2	1	2	3	3	1	0	1	2	2	14991.015	14991.017	-0.002
2	1	2	3	2	1	0	1	1	2	14991.974	14991.974	0.000
2	0	2	2	2	1	0	1	2	2	8603.114	8603.111	0.003
2	0	2	1	2	1	0	1	0	1	8603.475	8603.466	0.009
2	0	2	1	1	1	0	1	0	1	8603.757	8603.751	0.006
2	0	2	2	3	1	0	1	1	2	8604.126	8604.117	0.009
2	0	2	3	4	1	0	1	2	3	8604.236	8604.228	0.008
2	0	2	3	3	1	0	1	2	2	8604.372	8604.364	0.008
2	0	2	2	2	1	0	1	1	1	8604.685	8604.681	0.004
2	0	2	1	2	1	0	1	1	2	8605.620	8605.614	0.006
2	1	1	2	3	1	1	0	1	2	9094.888	9094.888	0.000
2	1	1	3	4	1	1	0	2	3	9095.587	9095.588	-0.001
3	0	3	4	5	2	0	2	3	4	12848.592	12848.609	-0.017
3	0	3	4	4	2	0	2	3	3	12848.723	12848.739	-0.016
2	1	1	3	4	2	0	2	3	4	7790.422	7790.429	-0.007
1	1	0	2	3	1	0	1	2	3	7299.064	7299.070	-0.006
3	1	3	4	5	2	1	2	3	4	12224.963	12224.965	-0.002

**Notes.** Upper and lower state quantum numbers are indicated by a single and double quote, respectively.<sup>(a)</sup> Observed frequency.<sup>(b)</sup> Calculated frequency.<sup>(c)</sup> Observed minus calculated frequency.

**Table 4.** Rotational and vibrational partition functions of cyanoacetamide.

Temperature (K)	$Q_r^{(a)}$	$Q_v^{(b)}$
9.38	745.7004	1.0002
18.75	2105.8326	1.0160
37.50	5952.0755	1.1454
75.00	16832.4575	1.6178
150.00	47614.2903	3.3948
225.00	87530.9193	7.6126
300.00	134828.227	17.7697

**Notes.**<sup>(a)</sup>  $Q_r$  is the rotational partition function (without taking the hyperfine splitting into account). We took  $J=250$  as the maximum value, where the centrifugal distortion (up to the sextic terms) was taken into account to calculate the partition function within a semirigid rotor approximation. A complete set of the calculated B2PLYPD3/aug-cc-pVTZ centrifugal distortion constants (Watson's  $S$ -reduction in the  $I'$  representation) is given in Table B.1 of Appendix B.<sup>(b)</sup>  $Q_v$  is the vibrational partition function. The total partition function of the molecule without hyperfine splitting is  $Q_r \times Q_v$ .

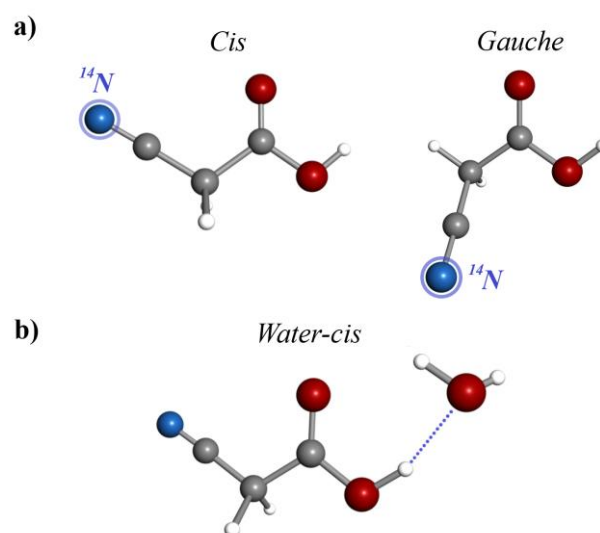
## 2.2. Rotational study of cyanoacetic acid

### 2.2.1 Jet-cooled rotational spectrum of cyanoacetic acid

Concurrent to the laboratory work and to guide the spectral search, we performed high-level computations on the conformational panorama of cyanoacetic acid complementing those of Reva et al. 2003 (a, b). We employed a double-hybrid density functional (B2PLYPD3), and the augmented correlation-consistent polarized valence triple- $\zeta$ , (aug-cc-pVTZ) basis set (Dunning 1989; Woon & Dunning 1993). Two predominant conformers (labeled as *cis* and *gauche*) were found, differing in the values of the CCC=O dihedral angle of the molecule (see Figure 5). We theoretically explored the plausible relaxation pathways from *cis*- to *gauche*- cyanoacetic acid and found a barrier of  $275 \text{ cm}^{-1}$  ( $3.3 \text{ kJ/mol}$ ) at the B2PLYPD3/aug-cc-pVTZ level. The predicted spectroscopic constants are collected in Table 1.

Briefly, we have transferred neutral molecules of cyanoacetic acid to the gas phase using a multi-nozzle heated system and probed them with a chirped pulse Fourier transform microwave spectrometer (CP-FTMW) (Mata et al. 2012). In this experiment, cyanoacetic acid was conventionally heated at  $70 \text{ }^\circ\text{C}$ . It was seeded in Helium to minimize plausible interconversion between the *gauche*- and *cis*- forms (Ruoff et al. 1990; Godfrey et al. 1996) (backing pressure of 1 bar) and adiabatically expanded into the vacuum

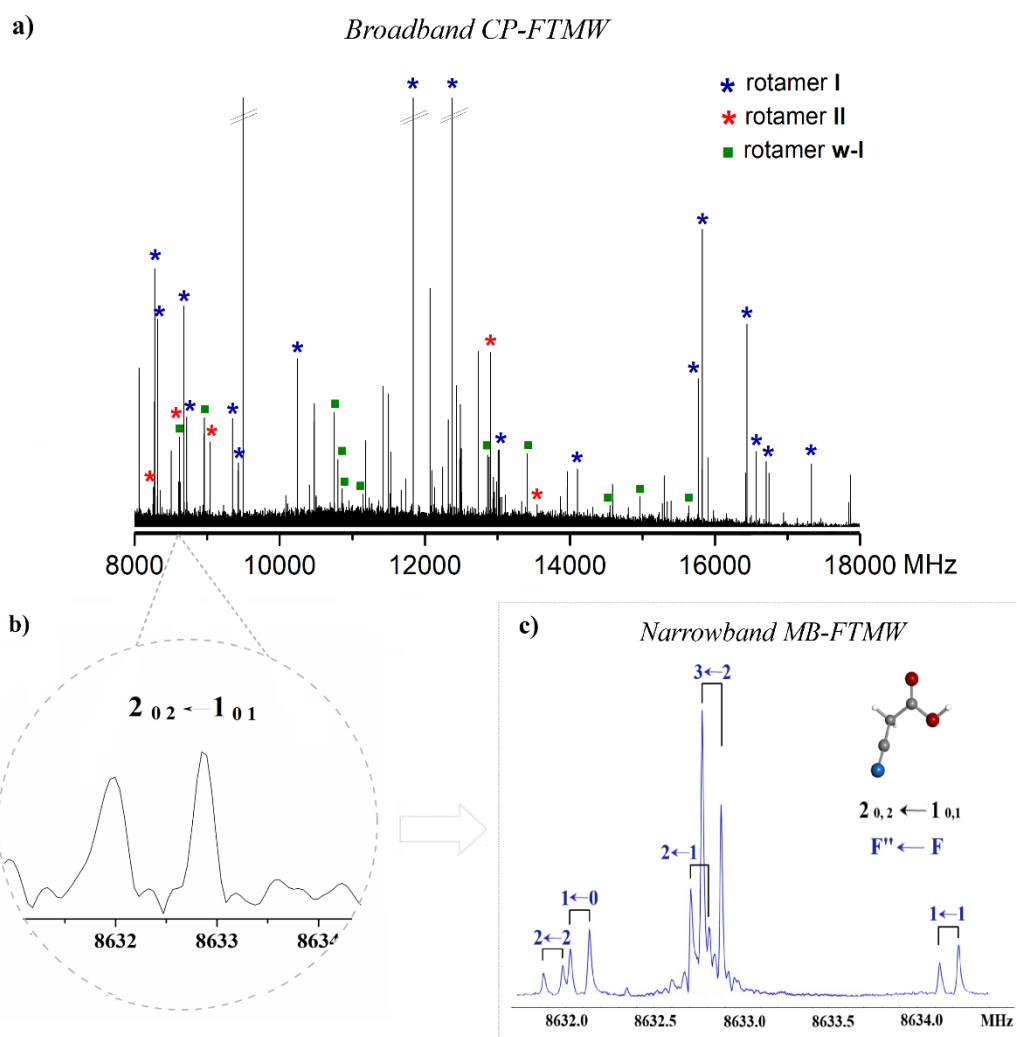
chamber of the spectrometer to generate a supersonic-jet. We employed a 300 W microwave excitation pulse to polarize the molecules macroscopically and collected 57,000 individual free induction decay signals (four FIDs per cycle) in the time domain. These were subsequently Fourier-transformed to obtain the broadband spectrum in the frequency domain.



**Figure 5.** a) Structures of the cyanoacetic acid monomers: *cis*- (left) and *gauche*-cyanoacetic acid (right) showing the  $^{14}\text{N}$  nucleus; b) Structure of the water cluster: *w-cis* cyanoacetic acid.

The broadband rotational spectrum measured for cyanoacetic acid from 8 to 18 GHz is presented in Figure 6 (a). Strong water dimer ( $\text{H}_2\text{O}$ )<sub>2</sub> signals and other lines belonging to decomposition products were quickly spotted and rejected from the analysis in a first step. At first glance, we quickly recognized an intense  $^3R$ -branch  $(J+1)_{0,J+1} \leftarrow J_{0,J}$  progression ascribable to a first rotamer, labeled as I. We further extended the initial assignment to *b*-type rotational transitions (with *J* ranging from 0 to 5). All transitions were observed split into various hyperfine components due to the  $^{14}\text{N}$  nuclear quadrupole coupling interactions. However, the resolution reached with our broadband CP-FTMW technique was not enough to completely resolve the hyperfine patterns (see Figure 6 (b)). In a first step, no attempt was performed to analyze the  $^{14}\text{N}$  nuclear hyperfine structure, but rather the intensity-weighted mean of the hyperfine line cluster was measured and fitted (Pickett 1991) to a rigid rotor Hamiltonian to give the first set of rotational constants ( $A = 10191.3$ ,  $B = 2267.2$  and  $C = 1877.3$ ) (in MHz) for rotamer I. As shown in Table 5, they perfectly match those theoretically predicted for the *cis*- form.

Once all the lines corresponding to *cis*- form were discarded from the analysis, we discovered a new weaker progression of *a*-*R*-branch transitions, attributed to another rotamer, labeled as II. We further confirmed the assignment by predicting and measuring additional *a*-type lines. However, *b*- and *c*-type lines were predicted but not observed. We obtained an initial set of rotational constants for rotamer II ( $A = 9278.6$ ,  $B = 2357.8$  and  $C = 1966.7$ ) (in MHz), which nicely agrees with those calculated for the *gauche*-form (see Table 5).



**Figure 6.** a) Jet-cooled broadband rotational spectrum measured for cyanoacetic acid (8 to 18 GHz). b) Zoom-in view of the spectrum showing the  $2_{0,2} \leftarrow 1_{0,1}$  transition of rotamer II (*gauche*-cyanoacetic acid) with unresolved  $^{14}\text{N}$  hyperfine structure; c) Completely resolved hyperfine pattern of the same transition attained with the cavity-based MB-FTMW technique. The intensity is given in arbitrary units.

**Table 5.** Theoretical and experimental spectroscopic parameters of cyanoacetic acid.

Parameters	Rotamer I	Rotamer II	<i>cis</i> <sup>[a]</sup>	<i>gauche</i>
$A$ <sup>[b]</sup> / MHz	10191.2722 (19)	9255.1 (12)	10199.5	9200.8
$B$ / MHz	2267.21502 (58)	2357.7873 (10)	2262.7	2352.4
$C$ / MHz	1877.26749 (44)	1966.7410 (10)	1873.0	1967.6
$\Delta J$ / kHz	0.344 (18)	-3.771 (77)	-	-
$\Delta JK$ / kHz	2.184 (86)	23.9 (13)	-	-
$\Delta K$ / kHz	9.57 (64)	-	-	-
$d_J$ / kHz	0.0776 (45)	-	-	-
$P_c$ <sup>[c]</sup>	1.6435 (10)	5.983 (13)	1.539	6.456
$ \mu_a ,  \mu_b ,  \mu_c $ <sup>[d]</sup>	Yes / Yes / No	Yes / No / No	3.9 / 2.9 / 0.0	2.1 / 0.5 / 1.3
$^{14}\text{N } X_{aa}$ <sup>[e]</sup>	-3.2108 (32)	-2.8203 (46)	-3.379	-2.937
$^{14}\text{N } X_{bb} - X_{cc}$	-0.9058 (76)	-0.692 (14)	-0.951	-0.637
$^{14}\text{N }  X_{ab} $	2.271 (91)	-	2.564	2.474
$N$ <sup>[f]</sup>	90	25	-	-
$\sigma$ <sup>[g]</sup>	13.0	3.3	-	-
$\Delta E_{\text{TOTAL}}$ <sup>[h]</sup>	-	-	0.0	49.3

[a] Theoretical calculations at B2PLYPD3/aug-cc-pVTZ level of theory. [b] A, B, and C represent the rotational constants (in MHz); [c]  $P_c$  is the planar inertial moment (in  $\text{u}\text{\AA}^2$ ), conversion factor:  $505379.1 \text{ MHz}\cdot\text{u}\text{\AA}^2$ . [d]  $\mu_a$ ,  $\mu_b$ , and  $\mu_c$  are the electric dipole moment components (in D). [e]  $\chi_{aa}$ ,  $\chi_{bb}$  and  $\chi_{cc}$  are the diagonal elements of the  $^{14}\text{N}$  nuclear quadrupole coupling tensor (in MHz); [f] N is the number of measured transitions. [g]  $\sigma$  is the root mean square (rms) deviation of the fit (in kHz). [h]  $\Delta E_{\text{TOTAL}}$  is the electronic and Gibbs energies (in  $\text{cm}^{-1}$ ) at 298 K relative to the global minimum calculated at the B2PLYPD3/aug-cc-pVTZ level of theory, taking into account the zero-point vibrational energy (ZPE) for the electronic energy ( $E_{\text{TOTAL}} = E + E_{\text{ZPE}}$ ) calculated at the same level. [i] Standard error in parentheses in units of the last digit.

Surprisingly, after excluding all the spectral signatures corresponding to both *cis*- and *gauche*- conformers, a small number of rotational lines remained unassigned. These lines most probably belong to the cyanoacetic acid monohydrate founded on the observation of strong water dimer ( $\text{H}_2\text{O}$ )<sub>2</sub> rotational transitions due to residual water (for detailed information see Section B.3 of Appendix B). In this context, we managed to recognize a set of *a*-type *R*-branch transitions (spaced approximately by  $B + C$ ) corresponding to a cyanoacetic acid-water rotamer [see Fig. 6 (a)]. We performed detailed searches for *b*- and *c*- type transitions, but they were not observed. Rotational constants and planar moments of inertia are collected in Table 6, while the complete list of measured transitions is provided in Table B.4 in Section B.4 of Appendix B. The structure of the cyanoacetic acid-water cluster is unambiguously confirmed by a comparison of the experimental and B2PLYPD3 predicted rotational parameters (see Table 6), showing that the observed rotational constants are in good agreement with those

calculated for the global minimum (*w-cis* cyanoacetic acid in Figure 5 (b)). Moreover, the value of the planar moment (1.9162 uÅ<sup>2</sup>) is very similar to that of the *cis*- monomer (1.6434 uÅ<sup>2</sup>). It suggests a nearly planar configuration, with the skeleton framework lying on the *ab* inertial plane and only the water hydrogen's out-of-plane contributions, which further corroborates the assignment. Although the monohydrated form itself is not expected to be a candidate for interstellar detection, this precise spectroscopic information should be helpful in future theoretical studies as a reference point in the characterization of plausible gas and gas-grain mechanisms involving cyanoacetic acid and water in the context of the ISM.

**Table 6.** Predicted and experimental spectroscopic parameters for the cyanoacetic acid-water cluster.

Parameters	rotamer w-I	<i>w-cis</i> <sup>[a]</sup>
$A$ <sup>[b]</sup> / MHz	7074.50 (33) <sup>[h]</sup>	7116.6
$B$ / MHz	1158.6515 (11)	1169.8
$C$ / MHz	1003.1681 (13)	1012.7
$\Delta_J$ / kHz	0.204 (9)	-
$\Delta_{JK}$ / kHz	-4.68 (21)	-
$P_c$ <sup>[c]</sup>	1.9162 (11)	1.997
$ \mu_a ,  \mu_b ,  \mu_c $ <sup>[d]</sup>	Yes / No / No	3.5 / 2.4 / 1.0
$N$ <sup>[e]</sup>	14	-
$\sigma$ <sup>[f]</sup>	15.2	-
$\Delta E_{TOTAL}$ <sup>[g]</sup>	-	0.0

[a] Theoretical computations at the B2PLYPD3/aug-cc-pVTZ level of theory. [b]  $A$ ,  $B$ , and  $C$  are the rotational constants (in MHz); [c]  $P_c$  is the planar inertial moment (in uÅ<sup>2</sup>), conversion factor: 505379.1 MHz–uÅ<sup>2</sup>. [d]  $\mu_a$ ,  $\mu_b$ , and  $\mu_c$  are the electric dipole moment components (in D). [e]  $N$  is the number of measured transitions. [f]  $\sigma$  is the root mean square (rms) deviation of the fit (in kHz). [g]  $\Delta E_{TOTAL}$  is the electronic energy (in cm<sup>-1</sup>) relative to the global minimum calculated at the B2PLYPD3/aug-c-pVTZ level, considering the zero-point vibrational energy (ZPE) for the electronic energy ( $E_{TOTAL} = E + E_{ZPE}$ ). [h] Standard error in parentheses in units of the last digit.

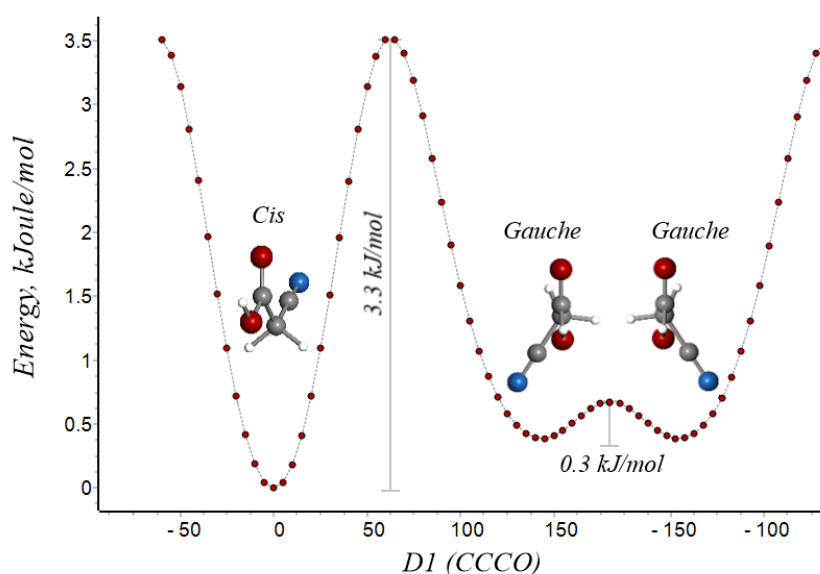
### 2.2.2 $^{14}\text{N}$ Hyperfine nuclear quadrupole coupling: narrowband rotational spectrum of cyanoacetic acid

A more challenging endeavor was the resolution of the nuclear quadrupole hyperfine structure. For observations in cold and quiescent molecular clouds such as TMC-1, the hyperfine structure could be a key issue for decoding the radioastronomical spectra (McCarthy & McGuire 2021). For instance, the interpretation of the hyperfine patterns of several interstellar molecules, such as cyanoallene,  $\text{CH}_2\text{CCHCN}$ , (Lovas et al. 2006) and benzonitrile,  $\text{C}_6\text{H}_5\text{CN}$ , (McGuire et al. 2018), was crucial for their conclusive identification in the ISM. Hence, we profit from the sub-doppler resolution of our narrowband MB-FTMW spectrometer (Balle & Flygare 1981; Alonso et al. 1997) to resolve the hyperfine structure of cyanoacetic acid. The analysis began with measuring five hyperfine components of the  $2_{0,2} \leftarrow 1_{0,1}$  transition belonging to *cis*-cyanoacetic acid and was further extended to other *a*- and *b*- type *R*-branch transitions. Once the analysis was completed, new  $K=1 \leftarrow 0$  *a*-*R*-branch transitions attributable to the *gauche*-form were identified. In this case, the MB-FTMW spectrometer's resolution was mandatory to unravel the  $^{14}\text{N}$  nuclear quadrupole structure (see Figure 6 (c)).

In a further step, we employed our double-resonance configuration (described in Chapter III), with which hyperfine components can be resolved, seeking an extension of our measurements to higher frequencies. We focused our attention on the most stable form, *cis*- cyanoacetic acid. We measured a total of 90 hyperfine components (together with the previous CP-FTMW and MB-FTMW measurements correctly weighted). We fitted them (Pickett 1991) using the *A*-reduction Hamiltonian in the *F* representation ( $H_R^{(A)}$ ) (Watson 1977), complemented with an additional term that accounts for the nuclear quadrupole coupling effects ( $H_Q$ ) (Foley 1947; Robinson & Cornwell 1953), ( $H = H_A^{(A)} + H_Q$ ). The rotational Hamiltonian was assembled in the coupled basis set  $F = J + I$  to describe the energy levels of the hyperfine structure in terms of the quantum numbers *J*, *K<sub>a</sub>*, *K<sub>c</sub>*, and *F*. We list the final set of spectroscopic parameters for *cis*- and *gauche*-cyanoacetic acid in the first and third column of Table 1, respectively. A sample list of the measured transitions for the *cis*- form is reported in Table 7. All measured lines for both rotamers are provided in Tables B.4–B.6 in Appendix B. We detected an unusual nonrigid behavior for the *gauche*- form reflected in large values of the  $\Delta_J$  and  $\Delta_{JK}$  centrifugal distortion constants compared to those obtained for the *cis*-form. This behavior can be ascribed to a large amplitude motion regarding the CCCO bond torsion,



resulting in a double minimum potential function with a low-energy barrier at the planar skeleton framework. The potential energy surface scan of the  $\angle\text{CCCO}$  torsion confirmed the presence of such a double minimum, separated by a barrier of  $25\text{ cm}^{-1}$  ( $0.3\text{ kJ/mol}$ ), which is similar to that observed for glycineamide (Alonso et al. 2018). As shown in Figure 7, two isoenergetic forms are connected by a torsion barrier that is low enough to split the ground state into two close-lying states ( $0^+$  and  $0^-$ ). Only the lowest-lying  $0^+$  state is observed due to vibrational cooling. This fact helps us to rationalize the requirement of rather large centrifugal distortional parameters for low  $J$  values and, therefore, explain the conformer's unexpected nonrigid nature.



**Figure 7.** Relaxed potential energy surface (PES) scan computed at the B2PLYPD3/aug-cc-pVTZ level, choosing the  $\angle\text{CCCO}$  torsion as the driving coordinate. A torsion barrier of  $0.3\text{ kJ/mol}$  ( $25\text{ cm}^{-1}$ ) connects the two isoenergetic *gauche*-forms.

All in all, the microwave spectrum of cyanoacetic acid has been explored up to 40 GHz to provide precise laboratory reference spectra that are needed to compare directly against low-frequency observational data in different regions of the ISM. With our measurements, the rotational transitions of cyanoacetic acid can be predicted with linewidths (FWHM) of about  $0.5\text{ kms}^{-1}$  up to 40 GHz in its equivalent radial velocity. This is enough for a radioastronomical search in the dark, cold molecular clouds, where lines with very narrow linewidths are expected ( $\sim 0.4\text{ kms}^{-1}$ ) (McGuire et al. 2018, McCarthy et al. 2020). In conclusion, this high-resolution spectroscopic information can be confidently used in future searches for cyanoacetic acid in various low-frequency surveys, such as the Green Bank Observatory (GBT) surveys (McCarthy et al. 2021) and

the Effelsberg 100-m (see Chapter IV) and Yebes 40-m observations (Cernicharo et al. 2021).

**Table 7.** Sample list of the measured transition frequencies for the ground-state of cyanoacetic acid.

Technique	J'	K <sub>a</sub> '	K <sub>c</sub> '	F'	J''	K <sub>a</sub> ''	K <sub>c</sub> ''	F''	$\nu_{obs}^a$ (MHz)	$\nu_{calc}^b$ (MHz)	$\nu_{obs}-\nu_{calc}^c$ (MHz)
MB-FTMW	2	1	2	2	1	1	1	1	7898.1967	7898.1967	0.0001
	2	1	2	2	1	1	1	2	7898.5439	7898.5424	0.0015
	2	1	2	3	1	1	1	2	7899.2045	7899.2039	0.0006
	2	1	2	1	1	1	1	2	7899.5725	7899.5717	0.0007
	2	1	2	1	1	1	1	0	7900.0909	7900.0901	0.0008
	3	1	2	3	2	1	1	2	13009.1663	13009.1648	0.0015
	3	1	2	4	2	1	1	3	13009.4489	13009.4480	0.0009
	3	1	2	2	2	1	1	1	13009.4668	13009.4709	-0.0041
CP-FTMW	2	1	2	1	1	0	1	0	15821.9328	15821.9533	-0.0205
	2	1	2	2	1	0	1	2	15822.3606	15822.3688	-0.0081
	2	1	2	3	1	0	1	2	15823.0169	15823.0303	0.0133
	2	1	2	2	1	0	1	1	15823.3103	15823.3319	-0.0216
	2	1	2	1	1	0	1	1	15824.3454	15824.3613	-0.0159
DR-MB-FTMW	5	1	5	5	4	0	4	4	26032.0800	26032.0715	0.0085
	5	1	5	6	4	0	4	5	26032.1200	26032.1139	0.0061
	2	2	0	1	1	1	1	0	32854.6280	32854.6368	-0.0088
	2	2	0	3	1	1	1	2	32854.6940	32854.6951	0.0011
	2	2	0	2	1	1	1	1	32854.7600	32863.7728	0.0138

**Notes.** Upper and lower state quantum numbers are indicated by ' and'', respectively.<sup>(a)</sup> Observed frequency.<sup>(b)</sup> Calculated frequency.<sup>(c)</sup> Observed minus calculated frequency. Table A3 containing the complete list of measured transitions.

Finally, in Table 8 we list the rotational ( $Q_r$ ) and vibrational ( $Q_v$ ) partition functions of cyanoacetic acid. We calculated  $Q_r$ 's values from first principles across the standard temperatures by the SPCAT program (Pickett 1991) as implemented in the JPL database (Pickett et al. 1998). We performed double-hybrid calculations at the B2PLYPD3/aug-cc-pVTZ level of theory to obtain the expected normal frequency modes (see Table A3). Afterward, we predicted the vibrational part,  $Q_v$ , using a harmonic approximation and a simple formula that coincides with Eq. (3.60) of Gordy & Cook (1970). The vibrational contributions were calculated by considering the lowest vibrational modes up to  $1000\text{ cm}^{-1}$ . The total partition function,  $Q_{tot}$ , is, therefore, the product of  $Q_r$  and  $Q_v$ , and does not include the hyperfine structure.

**Table 8.** Rotational and vibrational partition functions of cyanoacetic acid.

Temperature (K)	$Q_r^{(a)}$	$Q_v^{(b)}$
3.00	134.3515	1.0001
9.38	737.1679	1.0008
18.37	2081.8097	1.0293
37.50	5884.2420	1.2065
75.00	16640.5107	1.7985
150.00	47079.8270	3.9574
225.00	86551.9930	8.7615
300.00	133184.4558	19.5677

**Notes.** (a)  $Q_r$  is the rotational partition function. It does not take the hyperfine splitting into account and we took  $J = 150$  as the maximum  $J$  value. (b)  $Q_v$  is the vibrational partition function. The total partition function of the molecule (without hyperfine splitting) is  $Q_r \times Q_v$ .

### 3. CONCLUSIONS

We have carried out a theoretical study of the formation process of protonated cyanoacetamide as a plausible cyanoacetamide precursor, from protonated hydroxylamine and cyanoacetaldehyde. A detailed analysis of the corresponding singlet potential energy surface has been performed. Thus, we found an exothermic process ( $\Delta E = -89.8$  kcal/mol at the CCSD(T) level) with no net activation barrier initiated by the high-energy isomer of protonated hydroxylamine, which leads to protonated cyanoacetamide. Our results suggest that the neutral form,  $(\text{CH}_2(\text{CN})\text{C}(\text{O})\text{NH}_2)$ , might be a candidate molecule to be searched for in the ISM.

Furthermore, we have reported the first high-resolution rotational characterization of *trans* cyanoacetamide. The capability provided by the laser ablation coupled with time domain Fourier transform microwave techniques to obtain very accurate spectroscopic constants, in combination with high-level theoretical computations, enables the unequivocal identification of this elusive molecule. The two  $^{14}\text{N}_c$  and  $^{14}\text{N}_a$  nuclei of cyanoacetamide act as hyperfine rotational signatures of molecular conformation that increase the usefulness of this spectroscopic technique even more in molecular astrophysics. The experimental rotational constants together with the values of the  $^{14}\text{N}$  nuclear quadrupole coupling constants presented in this work are provided in order to reproduce the spectrum perfectly, and they constitute a first step to identify this molecule in the ISM.

We have also investigated the rotational spectrum of cyanoacetic acid in the microwave

region using time-domain Fourier transform microwave spectroscopy. For both *cis*- and *gauche*- forms, precise experimental values of the rotational constants, the  $^{14}\text{N}$  nuclear quadrupole coupling constants, and the quartic centrifugal distortion constants are determined to reproduce the spectrum. The  $^{14}\text{N}$  hyperfine structure acts again as a “fingerprint” probe for the molecular identification in low-frequency surveys. There is good reason to expect many astronomical discoveries given improvements in laboratory instrumentation and techniques and the construction of radio telescopes with unprecedented sensitivity regimes, remarking those directed towards low-frequency regions. As we already mentioned, thoroughly coordinated laboratory, modeling and observational effort will be essential to understand the origin and evolution of more complex –CN bearing systems.

We hope this work will benefit the astrophysics community by allowing a possible detection in the near future of cyanoacetamide and cyanoacetic acid, whose rotational signatures have remained unknown until now. In this regard, very recently, an interstellar search for both cyanides have been carried out toward Sgr B2(N) (A. Belloche priv. comm.) and toward the molecular cloud G+0.693–0.027 (V.M. Rivilla priv. comm.) using the ALMA REMoCA Survey and based on IRAM 30-m and Yebes 40-m observations, respectively. Nevertheless, both molecules remain undetected in these chemically rich astronomical sources.

## 4. EXPERIMENTAL SECTION

### 4.1. *Rotational spectrum of cyanoacetamide: LA-MB-FTMW spectroscopy*

A commercial sample of cyanoacetamide was purchased from Aldrich and used without further purification (99%). It is a white crystalline substance with a melting point of 119–121 °C. The rotational spectrum was recorded using a LA-MB-FTMW spectrometer, which works in the frequency range 8–18 GHz and has been described elsewhere (Alonso et al. 2009; Bermudez et al. 2014). The laser-ablated products were seeded in Ne (stagnation pressure 10 bar) and adiabatically expanded to form a supersonic jet into a Fabry–Pérot resonator. A short microwave radiation pulse (0.3  $\mu\text{s}$ ) was then applied to polarize the vaporized molecules macroscopically. We have recently implemented a microwave radiation pulse sequence that allows for multiple free induction decays (FIDs) collection to enhance further the instrument's sensitivity (León et al., 2017). The

microwave FID was recorded in the time domain and Fourier transformed to the frequency domain.

#### 4.2. *Rotational spectrum of cyanoacetic acid: Broadband CP-FTMW spectroscopy*

Cyanoacetic acid is commercially available and was used without further purification (98%). It is a white crystalline and hygroscopic substance with a melting point of 64-70 °C. Rotational spectra measurements were performed with two different spectroscopic techniques and using three different carrier gases: Ar, Ne, and He.

We investigated the jet-cooled rotational spectrum of cyanoacetic acid using a broadband CP-FTMW experimental setup. The spectrometer, based on a previous design (Mata et al. 2012), has been modified and improved with a multi-nozzle system. The use of multiple pulsed valve nozzles, where one can accommodate two or three nozzles in the interaction region of the spectrometer, leads to a decrease of the measurement time, which is directly related to increasing the spectrometer sensitivity, as previously reported in the literature (Brown et al. 2008). In this experiment, a two nozzle system was used to adiabatically expand a gas mixture containing the gaseous cyanoacetic acid. The solid heated at 70 °C was seeded in argon, neon, or helium, at stagnation pressure of 1 bar. A 300 W microwave excitation pulse was used to macroscopically polarize the molecules in the 8.0–18.0 GHz frequency range. In conclusion, up to 57,000 individual FIDs (four FIDs on each valve cycle) were averaged in the time domain, and Fourier-transformed into the frequency-domain. The uncertainty of the line measurements was estimated to be smaller than 20 kHz.

#### 4.3. *Jet-cooled narrowband MB-FTMW and DR-MB-FTMW spectroscopy of cyanoacetic acid*

To improve the quality of our measurements and enhance the precision of the  $^{14}\text{N}$  quadrupole coupling constants, we exploited the higher sensitivity and resolution of our molecular beam Fourier transform microwave (MB-FTMW) spectrometer, which operates between the frequency range from 4 to 14 GHz (Balle & Flygare 1981; Alonso et al. 1997). Cyanoacetic acid was conventionally heated at 70 °C and then seeded in Ar, Ne, or He (at a backing pressure 1 bar) and adiabatically expanded to generate a supersonic-jet into the Fabry-Pérot cavity. The microwave transient FID was again recorded in the time domain, and Fourier transformed to the frequency domain. The pulsed molecular beam was introduced parallel to the resonator's axis, so each observed transition emerges as a Doppler doublet. Finally, the resonance frequency is determined

by the arithmetic mean of the two Doppler components.

In the following step, we used our double-resonance (DR-MB-FTMW) technique to extend the frequency coverage of the most stable conformer's microwave data in two different frequency intervals, 18-26 GHz and 26-40 GHz, respectively (see Chapter III: Methodology). This new configuration is of great relevance since it serves as a straightforward connection between the microwave and millimeter-wave data, and more importantly, it allows us to measure in the same frequency region covered by several astronomical surveys conducted at cm wavelengths.

## 5. REFERENCES

- Agúndez, M., Marcelino, N., & Cernicharo, J. 2018, *ApJ*, 861, L22
- Alonso, J. L., Lorenzo, F. J., López, J. C., et al. 1997, *Chemical Physics*, 218, 267.
- Alonso, J. L., Pérez, C., Eugenia Sanz, M., et al. 2009, *Physical Chemistry Chemical Physics (Incorporating Faraday Transactions)*, 11, 617.
- Alonso, J. L. Lozoya, M. A. I. Peña, López, J. C. et al. 2014, *Chem. Sci.* 5, 515– 522
- Alonso, J. L. López, J. C. In *Gas-Phase IR Spectroscopy and Structure of Biological Molecules*, Springer International Publishing, Cham, 2015, p. 335.
- Alonso, E. R., Kolesniková, L., & Alonso, J. L. 2017, *J. Chem. Phys.*, 147, 124312 and references therein.
- Alonso, E. R., Kolesniková, L., Białkowska-Jaworska, E., et al. 2018, *ApJ*, 861, 70.
- Balle, T. J., & Flygare, W. H. 1981, *Review of Scientific Instruments*, 52, 33
- Becke, A. D. 1993, *J. Chem. Phys.*, 98, 5648
- Bell, M. B., Feldman, P. A., Travers, M. J., et al. 1997, *ApJ*, 483, L61 Belloche, A., Menten, K. M., Comito, C., et al. 2008, *A&A*, 482, 179. Belloche, A., Garrod, R. T., Müller, H. S. P., et al. 2014, *Science*, 345, 1584
- Belloche, A., Meshcheryakov, A. A., Garrod, R. T., et al. 2017, *A&A*, 601, A49 and references therein.
- Bermudez, C., Mata, S., Cabezas, C., Alonso, J. L. 2014, *Angew Chem Int Ed Engl.* 53(41):11015-11018.
- Binev, I. G., Stamboliyska, B. A., Binev, Y.I., 1997 *Journal of Molecular Structure*, 444, 235-245.
- Bockelée-Morvan, D., Lis, D. C., Wink, J. E., et al. 2000, *A&A*, 353, 1101 Bodo E., Bovolenta G., Simha C., Spezia R., 2019, *ThChA*, 138, 97.
- Bogey, M., Dubus, H., & Guillemin, J. C. 1990, *Journal of Molecular Spectroscopy*, 143, 180
- Brotten, N. W., Oka, T., Avery, L. W., et al. 1978, *ApJ*, 223, L105.
- Cabezas, C., Varela, M., Peña, I., et al. 2012, *Physical Chemistry Chemical Physics (Incorporating Faraday Transactions)*, 14, 13618.
- Cernicharo, J., Agúndez, M., Cabezas, C., et al. **2021**, *A&A*, 649, L15.
- Congiu, E., Fedoseev, G., Ioppolo, S., et al. 2012, *ApJ*, 750, L12 Coustenis, A., Bezard, B., Gautier, D., et al. 1991, *Icarus*, 89, 152

- Coustenis, A., Schmitt, B., Khanna, R. K., et al. 1999, *Planet. Space Sci.*, 47, 1305
- Dunning, T. H. 1989, *J. Chem. Phys.*, 90, 1007
- El Bialy, S.A.A. & Gouda, M.A. 2011, *J. Heterocyclic Chem.*, 48: 1280-1286.
- Favre, C., Fedele, D. Semenov, D., et al. 2018, *APJL*, 862:L2 (6pp).
- Ferris, J.P., Sánchez, R.A., Orgel, L.E. 2006, *J Mol Biol.* 1968;33(3):693-704
- Foley, H. M. 1947, *Physical Review*, 71, 747.
- Gaussian 16, Revision A.03, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B. et al., Gaussian, Inc., Wallingford CT, 2016.
- Godfrey, P. D., Brown, R. D., and Rodgers, F. M., 1996, *J. Mol. Struct.*, 376, 65.
- Godfrey, P. D., Brown, R. D., Robinson, B. J., et al. 1973, *Astrophys. Lett.*, 13, 119
- Gonzalez, C., & Schlegel, H. B. 1991, *J. Chem. Phys.*, 95, 5853
- Gordy, W., Cook, R. L. *Microwave Molecular spectra*, Interscience Pub, 1970, New York.
- Grimme, S., Antony, J., Ehrlich, S., et al. 2010, *J. Chem. Phys.*, 132, 154104
- Van Haverbeke, L. & Herman, M. A. 1975, *M.A.Spectrochim. Acta Part A*31, 959.
- He, J., Vidali, G., Lemaire, J.-L., et al. 2015, *ApJ*, 799, 49
- Herbst, E., & van Dishoeck, E. F. 2009, *ARA&A*, 47, 427
- Hollis, J. M., Lovas, F. J., Remijan, A. J., et al. 2006, *ApJ*, 643, L25.
- Horn, A., Møllendal, H., & Guillemin, J.-C. 2008, *Journal of Physical Chemistry A*, 112, 11009
- Hoy, A. R., Mills, I. M., & Strey, G. 1972, *Molecular Physics*, 24, 1265.
- Jefferts, K. B., Penzias, A. A., Wilson, R. W., 1970, *ApJ*, 161, p.L87.
- Jørgensen, J. K., Van der Wiel, M. H. D., Coutens, A., et al. 2016, *A&A*, 595, A117.
- Jørgensen, J. K., Belloche, A., & Garrod, R. T. 2020, *ARA&A*, 58, 727.
- Kawaguchi, K., Kasai, Y., Ishikawa, S.-I., et al. 1994, *ApJ*, 420, L95
- Kolesníková, L., Alonso, E. R., Mata, S., et al. 2017, *ApJS*, 233, 24.
- Kołos, R., & Grabowski, Z. R. 2000, *Ap&SS*, 271, 65
- Kunde, V. G., Aikin, A. C., Hanel, R. A., et al. 1981, *Nature*, 292, 686
- Largo, L., Rayón, V. M., Barrientos, C., et al. 2009, *Chemical Physics Letters*, 476, 174
- Lee, C., Yang, W., & Parr, R. G. 1988, *Phys. Rev. B*, 37, 785
- Lefloch, B., Ceccarelli, C., Codella, C., et al. 2017, *MNRAS*, 469, L73
- León, I., Alonso, E. R., Mata, S., et al. 2017, *Physical Chemistry Chemical Physics (Incorporating Faraday Transactions)*, 19, 24985.
- Liu, S., Girart, J. M., Remijan, A., & Snyder, L. E. 2002, *ApJ*, 576, 255
- Lovas, F. J., Remijan, A. J., Hollis, J. M., et al. 2006, *ApJ*, 637, L37.
- Mata, S., Peña, I., Cabezas, C., J. Alonso, J. L. 2012, *Journal of Molecular Spectroscopy*, 280, 91.
- McCarthy, M. C., Kelvin Lee, K. L., Loomis, R. A. et al., 2020, *Nature Astronomy*, 5, 176-180.
- McCarthy, M. C. & McGuire B. A. 2021, *J. Phys. Chem. A Article ASAP* DOI:

- McGuire, B. A., Burkhardt, A. M., Kalenskii, S. et al., 2018, *Science*, 359 (6372): 202-205.
- McKellar, A. 1940, *PASP*, 52, 187.
- Møllendal, H., Margulès, L., Motiyenko, R. A., et al. 2012, *Journal of Physical Chemistry A*, 116, 4047
- Møller, C., & Plesset, M. S. 1934, *Physical Review*, 46, 618. Nishi, N., Shinohara, H., & Okuyama, T. 1984, *J. Chem. Phys.*, 80, 3898
- Pardo, J. R., Cernicharo, J., Goicoechea, J. R., et al. 2005, *Astrochemistry: Recent Successes and Current Challenges*, 40
- Petrie, S., Millar, T. J., & Markwick, A. J. 2003, *MNRAS*, 341, 609.
- Pickett, H. M. 1991, *Journal of Molecular Spectroscopy*, 148, 371.
- Pickett, H. M., Poynter, R. L., Cohen, E. A., Delitsky, M. L., Pearson, J. C., & Müller, H. S. P. 1998, *J. Quant. Spectr. Rad. Transf.*, 60, 883
- Raghavachari, K., Trucks, G. W., Pople, J. A., et al. 1989, *Chemical Physics Letters*, 157, 479.
- Remijan, A., Snyder, L. E., Friedel, D. N., et al. 2003, *APJ*, 590, 1, pp. 314-332.
- Reva, I. D. (a), Stepanian, S. G., Adamowicz, L. Fausto R., 2003, *Chemical Phys. Letters* 107, 6351-6359.
- Reva, I. D. (b), Stepanian, S. G., Adamowicz, L. Fausto R., 2003, *J Phys. Chem A.*, 374, 631-638.
- Rivilla, V. M., Martín-Pintado, J., Jiménez-Serra, I., et al. 2020, *ApJ*, 899, L28
- Robinson, G. W., & Cornwell, C. D. 1953, *J. Chem. Phys.*, 21, 1436.
- Ruoff, R. S., Klots, T. D., Emilson T. and Gutowski, H. S. 1990, *J. Chem. Phys.*, **93**, 3142
- Sánchez, R.A., Ferris, J.P., Orgel, L.E. 1967, *J Mol Biol.* 1968;33(3):693-704.
- Schwabe, T., & Grimme, S. 2007, *Physical Chemistry Chemical Physics (Incorporating Faraday Transactions)*, 9, 3397
- Schermann, J.-P. 2008 *Spectroscopy and Modeling of Biomolecular Building Blocks*, Elsevier, Amsterdam.
- Schwabe, T., & Grimme, S. 2007, *Physical Chemistry Chemical Physics (Incorporating Faraday Transactions)*, 9, 3397.
- Snow, J.L., Orlova, G., Blagojevic, V., Bohme, D.K.. 2007 *J. Am. Chem. Soc.*, 129(32):9910-9917.
- Tsegaw, Y. A., Góbi, S., Förstel, M., et al. 2017, *Journal of Physical Chemistry A*, 121, 7477.
- Watson, J. K. G. 1977, *Journal of Molecular Spectroscopy*, 66, 500.
- Winnewisser, G., & Walmsley, C. M. 1978, *A&A*, 70, L37.
- Woon, D. E., & Dunning, T. H. 1993, *J. Chem. Phys.*, 98, 1358.
- Zheng, W., & Kaiser, R. I. 2010, *Journal of Physical Chemistry A*, 114, 5251.



# CHAPTER VI. TOWARD THE LIMITS OF COMPLEXITY OF INTERSTELLAR CHEMISTRY: ROTATIONAL SPECTROSCOPY AND ASTRONOMICAL SEARCH OF *N*- AND *I*- BUTYRALDEHYDE

Adapted from Sanz-Novo et al. 2021 under review.



In recent times, large organic molecules of exceptional complexity have been found in diverse regions of the interstellar medium (ISM). In this context, we aim to provide accurate frequencies of the ground vibrational state of two key aliphatic aldehydes, *n*-butyraldehyde and its branched-chain isomer, *i*-butyraldehyde, to enable their eventual detection in space. We want to test the level of complexity that interstellar chemistry can reach in regions of star formation. We employed a frequency modulated millimeter- and submillimeter-wave absorption spectrometer to measure the rotational features of *n*- and *i*-butyraldehyde. Several thousands of transitions belonging to the lower-energy conformers of two distinct linear and branched isomers have been assigned up to 325 GHz. A precise set of the relevant rotational spectroscopic constants have been determined for each structure as a first step to identifying both molecules in the ISM. We then used the spectral line survey named Re-Exploring Molecular Complexity with ALMA (REMoCA), performed toward the star forming region Sgr B2(N) with ALMA to search for *n*- and *i*-butyraldehyde. We also searched for both aldehydes toward the molecular cloud G+0.693-0.027 with IRAM 30m and Yebes 40m observations. The observational results are compared with computational results from a recent gas-grain astrochemical model. We report the nondetection of these isomers toward both astronomical sources, highlighting a leap around one order of magnitude in the aldehyde's abundance while increasing the level of complexity (acetaldehyde > propionaldehyde > butyraldehyde)

## 1. INTRODUCTION

Our understanding of the chemistry of star-forming regions is constantly growing in view of the construction of radio telescopes with unparalleled regimes of sensitivity and the advances in laboratory astrophysics. In current astrochemical research, significant endeavors have been made to study the so-called complex organic molecules, (COMs, Jorgensen et al. 2020), which are carbon-based molecules containing more than six atoms (see McGuire 2018 for a census). Among them, aldehydes are some of the most widespread species in nature, being precursors of many biologically relevant molecules. Apart from the already identified interstellar aldehydes (See Section 5.3 of Chapter II), more complex aldehydes such as amino acetaldehyde (see Appendix E) and lactaldehyde ( $\text{CH}_3\text{CH}(\text{OH})\text{CH}(\text{O})$ , Alonso et al. 2019), which is the rational step up in complexity from glycolaldehyde, are yet to be detected. In this context, we propose the study of two key aliphatic aldehydes of exceptional complexity, *n*-butyraldehyde ( $n\text{-C}_4\text{H}_8\text{O}$ ;  $\text{CH}_3\text{CH}_2\text{CH}_2\text{CHO}$ ) and the branched *i*-butyraldehyde ( $i\text{-C}_4\text{H}_8\text{O}$ ;  $(\text{CH}_3)_2\text{CHCHO}$ ), which have been already identified in several chondritic meteorites (Aponte et al. 2019), as targets for interstellar detection.

Further increasing the interest of *n*- and *i*-butyraldehyde (hereafter *n*-PrCHO and *i*-PrCHO, respectively) as astronomical candidates, Abplanalp et al. (2018) revealed that both isomers, together with distinct  $\text{C}_3\text{H}_8\text{O}$  species (propanol, methyl ethyl ether), could be generated by exposing interstellar analog ices to ionizing radiation. Therefore, they should be detectable in the gas phase after desorption from icy grains in regions of star formation (Abplanalp & Kaiser 2019). This further suggests that both aldehydes are expected to be formed in the ISM and motivate the astrophysical community even more to search for them.

Both *n*- and *i*-PrCHO are colorless liquids with a boiling point of 74.8 °C and 63.0 °C, respectively. They have been previously investigated in the condensed phases by NMR, IR and Raman spectroscopy (Sbrana et al. 1970, Piart-Goyperon et al. 1991). Also, their conformational panoramas have already been explored in the gas phase using microwave and far infrared spectroscopy (Lee 1971, Stiefvater 1986; Durig et al. 1989

and Hotopp et al 2012). However, to our knowledge, their spectra have remained uncharted in the mm/submm-wave regions until now. Accurate spectroscopic data at high frequency is usually mandatory for the eventual astronomical detection of many interstellar molecules in several spectral line surveys, such as the EMoCA and REMoCA survey conducted with ALMA towards SgrB2(N) (Belloche 2017 and Belloche et al. 2019) and the IRAM 30m survey of the G+0.693-0.027 giant molecular cloud (Rivilla et al. 2020, Jiménez-Serra et al. 2020)

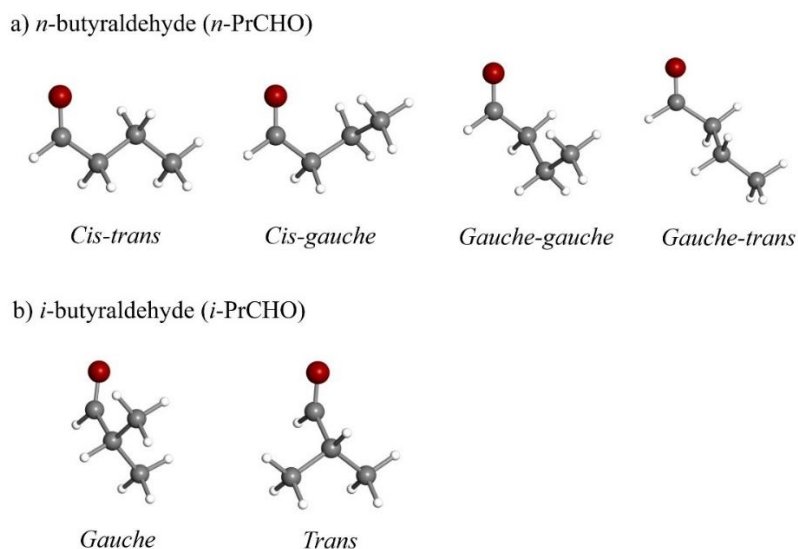
The lack of accurate millimeter-wave data for *n*- and *i*-PrCHO thus prompted new laboratory spectroscopic measurements over the frequency range from 75 to 325 GHz. In this work, we aim to provide an extensive rotational characterization of both aldehydes in the millimeter- and submillimeter-wave regions. The analysis of the rotational spectra at these wavelengths will enable us to confidently search for both species toward different regions of the ISM. However, a future harmonized observational and laboratory effort will be essential to unveil the pathways of chemical evolution of more complex systems as well as to conclusively disclose the ratio between aliphatic and cyclic molecules in the ISM (Loomis et al. 2016; McCarthy & McGuire 2021).

In Section 2.1 we show the complementary quantum chemical computations. Sections 2.2 and 2.3 present the results and discussion of our analysis of the experimental data. The astronomical results of a search for *n*- and *i*-PrCHO toward the high-mass star-forming region Sgr B2(N) as well as the G+0.693-0.027 giant molecular cloud in the Galactic Center are reported in Sections 2.4 and 2.5, respectively. Afterward, in Section 2.6 the astronomical part is discussed in a broader astrochemical framework. Then, we present in Section 3 the conclusions of this work. Finally, we describe in Section 4 the experimental setup that was employed to record the spectra for both aldehydes.

## 2. RESULTS AND DISCUSSION

### 2.1 Molecular modeling

In addition to the previous computational work on *n*-PrCHO (Klimkowski et al. 1984; Dwivedi & Rai 2008; Hotopp et al. 2012), we performed high level calculations on both *n*- and *i*-PrCHO conformers to accompany the experimental study. The geometries were optimized at the density functional theory (DFT) level, employing the double-hybrid functional B2PLYP (Grimme 2006) combined with Grimme's D3BJ dispersion (Grimme et al. 2011) (B2PLYPD3), in conjunction with the correlation-consistent basis set of Dunning aug-cc-pVTZ (correlation-consistent polarized valence triple-zeta including diffuse functions) (Dunning 1989; Woon & Dunning 1993). Additional optimizations are also carried out using the ab initio coupled-cluster with single and double excitations (CCSD) method (Raghavachari et al. 1989). In this case the Dunning's cc-pVTZ (correlation-consistent polarized valence triple-zeta) basis set was employed. On the optimized geometries harmonic vibrational frequencies were calculated at both levels of theory. This allows to estimate the zero-point vibrational energy correction (ZPE), confirm the nature of the stationary points located on the potential energy surface and compute the vibrational contribution to the partition function.



**Figure 1.** Structures of the lowest in energy conformers of *n*- and *i*-PrCHO. Geometries are optimized at the CCSD/aug-cc-pVTZ level of theory.

The structures of the low energy conformers of both isomers are presented in Fig.1, and their harmonic vibrational frequencies calculated at the CCSD/cc-pVTZ level are collected in Table C.1 of Appendix C. To obtain more accurate stability order single point calculation were carried out on the CCSD/cc-pVTZ optimized geometries. Energies are calculated at the CCSD(T) (coupled-cluster with single and double excitations including triple excitations through a perturbative treatment) level extrapolated to the Complete Basis Set limit, denoted as CCSD(T)/CBS. It is computed from the  $n^{-3}$  extrapolation equation (Helgaker et al. 1997) applied to the CCSD(T) energies calculated with the triple- ( $n=3$  aug-cc-pVTZ) and quadruple-zeta ( $n=4$ , aug-cc-pVQZ) basis sets. In Table 1 the calculated spectroscopic parameters and relative energies of the conformer of *n*-PrCHO are reported, the corresponding values for the *i*-PrCHO conformers are shown in Table 2.

**Table 1.** Theoretical ground-state spectroscopic constants for the low-energy *n*-butyraldehyde conformers (A-Reduction, I'-Representation).

Parameters	<i>cis-gauche</i>	<i>cis-trans</i>	<i>gauche-gauche</i>	<i>gauche-trans</i>
A <sup>(a)</sup> (MHz)	8535.664	15184.914	9888.586	20370.216
B (MHz)	3610.610	2565.358	3050.824	2155.419
C (MHz)	2940.554	2286.380	2610.863	2114.933
$ \mu_a ,  \mu_b ,  \mu_c $ <sup>(b)</sup> (D)	0.7 / 2.3 / 0.7	1.5 / 2.0 / 0.0	2.3 / 1.5 / 0.7	2.6 / 0.4 / 1.2
$\Delta_J$ (kHz)	3.54433	0.602729	6.00434	0.899217
$\Delta_K$ (kHz)	21.9814	36.6933	162.286	635.956
$\Delta_{JK}$ (kHz)	-11.5161	-4.179	-51.5314	-34.5039
$\delta_J$ (kHz)	1.0782	0.098764	1.81436	-0.134512
$\delta_K$ (kHz)	5.35409	-0.0926365	9.70505	13.7124
$V_3$ <sup>(c)</sup>	963	1067	997	1069
$\Delta E$ <sup>(d)</sup>	0.87	0.00	3.12	3.20

**Notes.** <sup>(a)</sup> A, B, and C represent the rotational constants computed at the CCSD/cc-pVTZ level. <sup>(b)</sup>  $|\mu_a|, |\mu_b|, |\mu_c|$  are the absolute values of the electric dipole moment components (in D). <sup>(c)</sup>  $V_3$  is the barrier height to the methyl internal rotation (in  $\text{cm}^{-1}$ ). <sup>(d)</sup>  $\Delta E$  is the electronic energy calculated at the CCSD(T)/CBS level, taking into account the zero-point vibrational energy (ZPE) at the CCSD/cc-pVTZ level (in  $\text{kJ mol}^{-1}$ ).

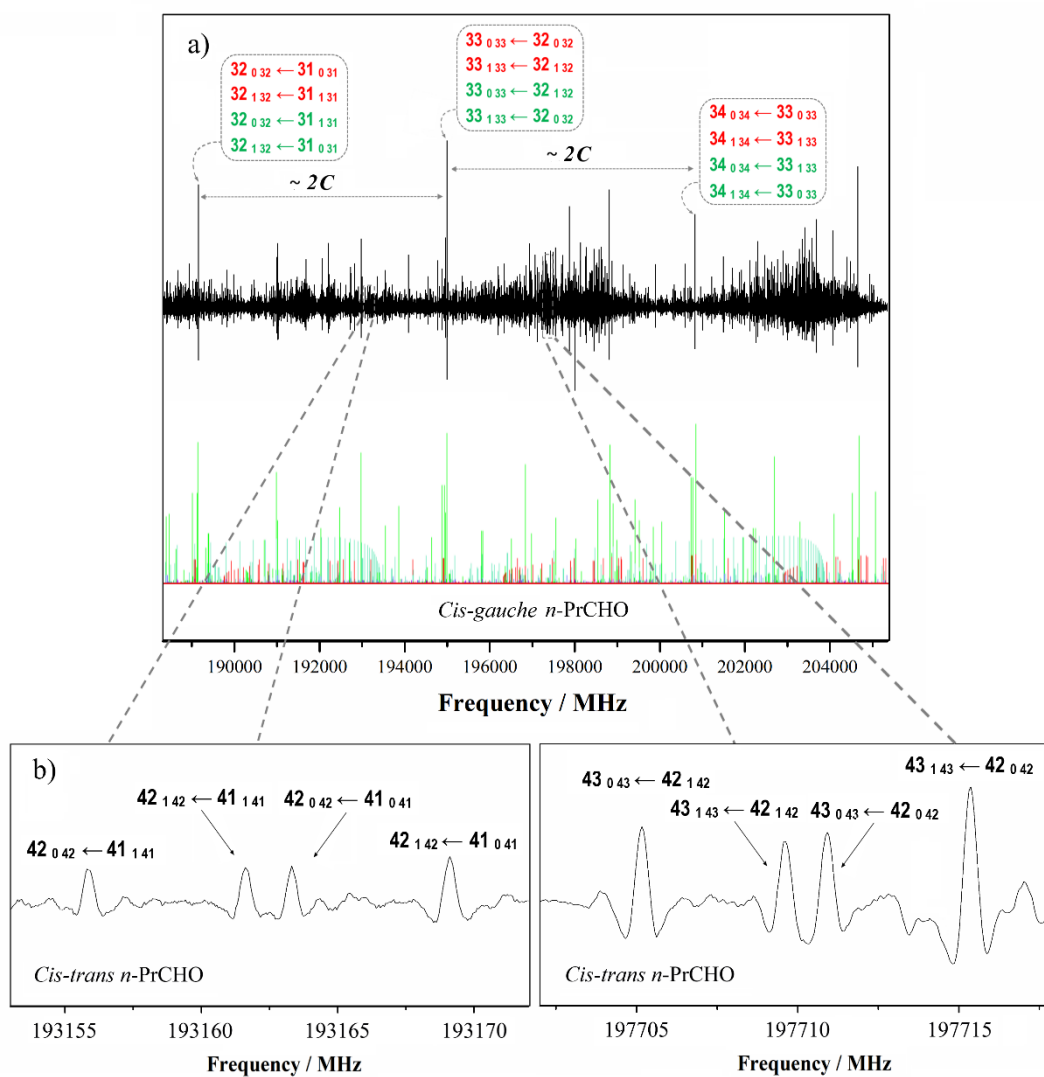
**Table 2.** Theoretical ground-state spectroscopic constants for the two low-energy conformers of *i*-butyraldehyde (*A*-Reduction, *I'*-Representation).

Parameters	<i>gauche</i>	<i>trans</i>
$A^{(a)}$ (MHz)	7534.289	7757.132
$B$ (MHz)	4125.416	3737.809
$C$ (MHz)	3000.209	2827.664
$ \mu_a ,  \mu_b ,  \mu_c ^{(b)}$ (D)	2.4 / 0.8 / 0.9	2.7 / 0.0 / 0.7
$\Delta_J$ (kHz)	1.856450	0.702570
$\Delta_K$ (kHz)	10.484600	-45.357200
$\Delta_{JK}$ (kHz)	0.274349	51.546800
$\delta_J$ (kHz)	0.259161	0.191034
$\delta_K$ (kHz)	3.764370	25.127900
$\Delta E^{(c)}$	0.00	1.86

**Notes.**<sup>(a)</sup>  $A$ ,  $B$ , and  $C$  represent the rotational constants computed at the CCSD/cc-pVTZ level.<sup>(b)</sup>  $|\mu_a|$ ,  $|\mu_b|$ ,  $|\mu_c|$  are the absolute values of the electric dipole moment components (in D).<sup>(c)</sup>  $\Delta E$  is the electronic energy calculated at the CCSD(T)/CBS level, taking into account the zero-point vibrational energy (ZPE) at the CCSD/cc-pVTZ level (in  $\text{kJ mol}^{-1}$ ).

## 2.2 Millimeter and submillimeter-wave spectra of *n*-PrCHO

In Figure 2(a) we present a fragment of the room-temperature millimeter-wave spectrum. The spectrum is governed by sets of strong *b*-type *R*-branch  $(J+1)_{0J+1} \leftarrow J_{1J}$  and  $(J+1)_{1J+1} \leftarrow J_{0J}$  transitions, which are separated about 6 GHz and outshine the rest of spectral features. Starting from the previous microwave assignments (Hotopp et al. 2012), we found that this patterns approximately coincides with the 2C predicted value of the *cis-gauche* conformer. We noted that at these high frequencies the rotational energy levels corresponding to the lowest  $K_a$  values become quasi degenerate (depicted in Figure 3(a) for  $K_a = 0, 1$ ). Thus, a quadruple-degenerated blended line arises from the coalescence of pairs of *b*-type rotational transitions with corresponding *a*-type lines involving these energy levels. The early assignments were further extended to other weaker *b*-type transitions. Subsequently, we assigned higher  $K_a$  *R*-branch *a*- and *b*-type transitions together with several *Q*-branch transitions, spanning  $J$  and  $K_a$  values up to 70 and 31, respectively.



**Figure 2.** a) Section of the millimeter-wave spectrum of *n*-PrCHO showing the dominant quadruple degenerated lines of the *cis-gauche* conformer, containing pairs of *a*- and *b*-type *R*-branch transitions. The final predicted spectrum of this conformer, computed for 300 K, is given for comparison (*a*-type lines are depicted in red and *b*-type lines are depicted in green). b) Zoomed view of the millimeter-wave spectrum showing unblended quadruplets of *a*- and *b*-type *R*-branch transitions of *cis-trans* conformer with  $K_a = 0$ , 1 at  $J'' = 41$  (*left panel*) and  $J'' = 42$  (*right panel*). Intensity is given in arbitrary units.

In a second step, a closer inspection of the millimeter-wave spectrum was directed to search for the *cis-trans* conformer, which shows a practically prolate behavior. Guided by the previous experimental data (Hotopp et al 2012), we managed to identify several

progressions of *a*- and *b*- type R-branch transitions belonging to *cis-trans n*-PrCHO, of relatively lower intensity (see Figure 3(b)). Afterward, new rotational lines were assigned and after iterative fittings and predictions, we completed the rotational quantum number's coverage up to  $J = 69$  and  $K_a = 20$ .

In parallel to the laboratory work, the barriers of the methyl group internal rotation were predicted by high-level B2PLYPD3 quantum chemical calculations to be higher than 11.5 and 12.5 kJ/mol for the *cis-gauche* and *cis-trans* conformers, respectively. Nevertheless, we employed the XIAM program (Hartwig & Dreizler 1996) as well as the Loomis–Wood type plot technique (Loomis & Wood 1928; Kisiel et al. 2012) to have a thorough look at the spectra. No splitting due to internal rotation motion could be observed for neither of the conformers given the ~50 kHz experimental resolution of our spectrometer, which is in perfect agreement with the computed barrier heights.

All in all, we measured more than 2800 rotational transitions for the *cis-gauche n*-PrCHO along with 1375 transitions for the *cis-trans* conformer, including in the data set the microwave lines previously observed by Hotopp et al. (2012). The Pickett's SPFIT/SPCAT program package (Pickett 1991) was used to perform the corresponding fits using a Watson's *A*-reduced Hamiltonian in the *I'*-representation. In Table 2 we present the final set of spectroscopic constants for both conformers, which conclusively enhance the previous microwave experimental data (Hotopp et al 2012). It is worth noticing the frequency coverage (75 to 325 GHz) extends well beyond the absorption maximum at the typical temperatures of hot interstellar clouds, which enables a confident search for these species in space.

Finally, we present the rotational ( $Q_r$ ) and vibrational ( $Q_v$ ) partition functions of *cis-gauche* and *cis-trans n*-PrCHO separately, which are shown in Tables 4 and Tables 5. We used SPCAT (Pickett 1991) to compute  $Q_r$ 's values from first principles at the typical temperatures as applied in the JPL database (Pickett et al. 1998). We estimated the vibrational part,  $Q_v$ , using a harmonic approximation and Eq. 3.60 of Gordy & Cook (1970), where only the vibrational modes below 1000 cm<sup>-1</sup> were taken into consideration. We obtained the frequencies of the normal modes from CCSD calculations (see Table C.1 of Appendix C). The full partition function,  $Q_{tot}$ , is therefore the product of  $Q_r$  and



$Q_v$ . Note that an additional conformational correction factor should be introduced in order to derive a proper estimation of the total column density (upper limit) of the molecule.

**Table 3.** Experimental ground-state spectroscopic constants of n-PrCHO (A-Reduction, I'-Representation).

Parameter	<i>Cis-gauche</i>	<i>Cis-trans</i>
A <sup>[a]</sup> / MHz	8508.53795 (16) <sup>[e]</sup>	15069.38645 (92)
B / MHz	3588.81565 (14)	2555.992540 (85)
C / MHz	2928.57776 (10)	2278.607628 (88)
$ \mu_a ,  \mu_b ,  \mu_c $ <sup>[b]</sup> / D	Yes / Yes / Yes	Yes / Yes / No
$\Delta_J$ / kHz	3.648341 (87)	0.611752 (30)
$\Delta_K$ / kHz	23.5595 (33)	37.8014 (89)
$\Delta_{JK}$ / kHz	-12.199378 (19)	-4.23054 (30)
$\delta_J$ / kHz	1.110506 (38)	0.0994029 (45)
$\delta_K$ / kHz	5.60931 (81)	-0.1183 (10)
$\Phi_J$ / mHz	-12.147 (23)	0.0344(35)
$\Phi_{JK}$ / Hz	0.12797 (67)	0.004310 (50)
$\Phi_K$ / Hz	0.71765 (40)	0.492 (27)
$\Phi_{KJ}$ / Hz	-0.49693 (20)	-0.2641 (11)
$\phi_J$ / mHz	-5.603 (11)	-
$\phi_{JK}$ / Hz	-0.06452 (33)	-
N <sup>[c]</sup>	2839	1375
$\sigma$ <sup>[d]</sup>	29	28

[a]  $A$ ,  $B$ , and  $C$  represent the rotational constants; [b]  $\mu_a$ ,  $\mu_b$ , and  $\mu_c$  are the electric dipole moment components. [c]  $N$  is the number of measured rotational transitions. [d]  $\sigma$  is the root mean square (rms) deviation of the fit (in kHz). [e] Standard error in parentheses in units of the last digit.

**Table 4.** Rotational and vibrational partition functions of *cis-gauche n*-PrCHO

Temperature (K)	$Q_r^{(a)}$	$Q_v^{(b)}$
9.38	514.3036	1.0000
18.75	1451.3130	1.0003
37.50	4101.0172	1.0206
75.00	11598.5581	1.2410
150.00	32831.0285	2.6021
225.00	60373.8787	6.0802
300.00	93047.2540	14.1692

**Notes.** <sup>(a)</sup>  $Q_r$  is the rotational partition function. It does not take the hyperfine splitting into account. <sup>(b)</sup>  $Q_v$  is the vibrational partition function. The total partition function of the molecule (without hyperfine splitting) is  $Q_r \times Q_v$ .

**Table 5.** Rotational and vibrational partition functions of *cis-trans n*-PrCHO

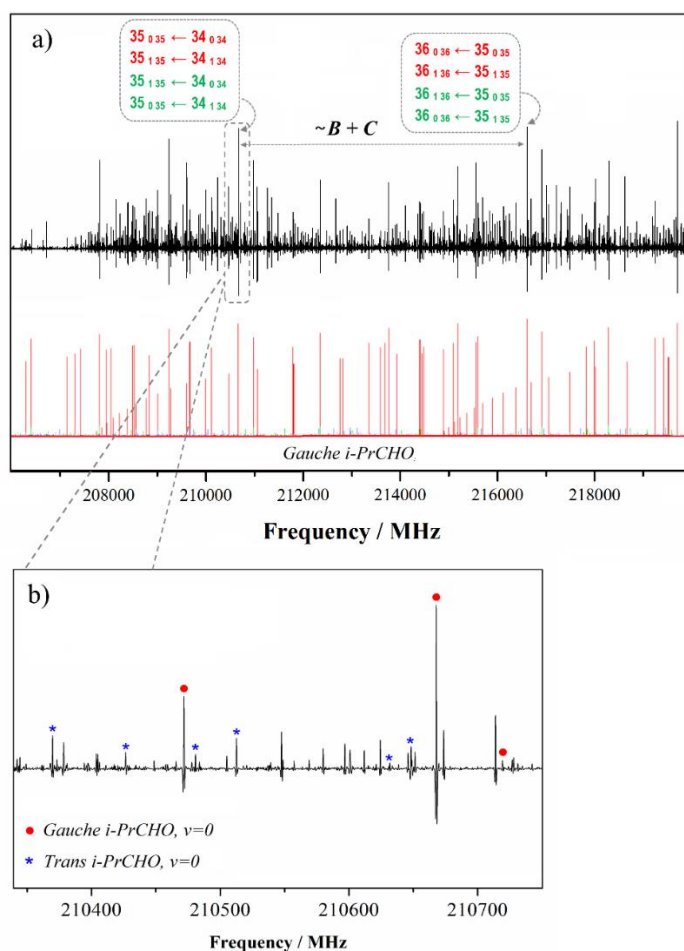
Temperature (K)	$Q_r^{(a)}$	$Q_v^{(b)}$
9.38	518.5813	1.0000
18.75	1464.0056	1.0020
37.50	4137.2687	1.0486
75.00	11698.8301	1.3624
150.00	33095.6561	3.1151
225.00	60821.5883	7.6557
300.00	93661.9204	18.3964

**Notes.** <sup>(a)</sup>  $Q_r$  is the rotational partition function. It does not take the hyperfine splitting into account. <sup>(b)</sup>  $Q_v$  is the vibrational partition function. The total partition function of the molecule (without hyperfine splitting) is  $Q_r \times Q_v$ .

### 2.3 Millimeter and submillimeter-wave spectra of *i*-PrCHO

At the first stage of the spectral analysis, we employed predictions based on the spectroscopic parameters from Stiefvater (1986), including the experimental values of the electric dipole-moment components. Given that the two most stable rotameric species of *i*-PrCHO are near prolate asymmetric tops with non-zero  $\mu_a$  dipole-moment component (*gauche* conformer:  $|\mu_a| = 2.43$  (2),  $|\mu_b| = 0.80$  (3),  $|\mu_c| = 0.83$  (2), and *trans* conformer:  $|\mu_a| = 2.82$  (2),  $|\mu_b| = 0.0$  by  $C_s$  symmetry,  $|\mu_c| = 0.46$  (3), we first explore the rich millimeter-wave spectrum to search for intense sets of *a*-type *R*-branch transitions, typically separated by  $\sim B+C$  (see Figure 3(a)). Also in this case, as the quantum number

$J$  increases, the rotational energy levels which involves the lowest  $K_a$  quantum numbers become near-degenerate, and pairs of  $a$ - and  $b$ -type lines form quartets, which are already blended into one quadruple-degenerated line at the high frequencies shown in Fig. 3(a). After we assigned low  $K_a$   $a$ -type transitions, we were able to locate many  $b$ -type  $R$ - and  $Q$ -branch lines, despite the relatively lower value of the  $\mu_b$  dipole moment component. Then, several  $c$ -type lines were successfully identified as well. This analysis was further extended to higher  $K_a$   $R$ - and  $Q$ -branch transitions.



**Figure 3.** a) Portion of the room-temperature millimeter-wave spectrum of *i*-PrCHO showing intense  $a$ -type  $R$ -branch transitions of the *gauche* conformer (approximately separated by  $B+C$ ). Also, the final predicted spectrum of this conformer, computed for 300 K, is given for comparison ( $a$ -type lines are depicted in red,  $b$ -type lines are depicted in green and  $c$ -type lines are depicted in purple). b) Zoom-in view

of the millimeter-wave spectrum showing several *a*-type *R*-branch lines belonging to *trans i*-PrCHO (in blue). Intensity is given in arbitrary units.

Once we finished with the analysis of *gauche i*-PrCHO, we searched for rotational signatures attributable to the *trans* conformer. Guided by the previous experimental data (Stiefvater 1986), we managed to identify a much weaker progression of *a*-type *R*-branch transitions belonging to *trans i*-PrCHO. Also, keeping in mind this conformation presents a  $C_s$  symmetry, no *b*-type spectrum was neither expected nor observed for *trans i*-PrCHO. Afterward, we completed the analysis up to  $K_a = 12$ . Also, *c*-type lines were predicted but not observed.

On the whole, almost 3400 rotational transitions were measured for the *gauche* form of *i*-PrCHO up to  $J = 63$  and  $K_a = 24$ , while only 500 transitions were recorded for the *trans* form, which is reasonable given that it is a higher in energy conformer (see Table 2). In this case, we did not encompass the previous microwave data (Stiefvater 1986) because its high standard deviation (97 kHz) significantly deteriorated the quality of our final fit. The corresponding fits and predictions were made in terms of Watson's *A*-reduced Hamiltonian in *I'*-representation with the Pickett's SPFIT/SPCAT program package (Pickett 1991). Hence, the presently experimental data enabled the determination of accurate rotational constants and the full set of quartic and sextic centrifugal distortion constants, showing a remarkable RMS deviation of less than 25 kHz for *gauche i*-PrCHO. The final spectroscopic set, which is listed in Table 6, definitively improves the previous microwave experimental data by up to even four orders of magnitude (Stiefvater 1986). Also, from a comparison between the experimental and theoretically predicted parameters, we can infer that the rotational constants as well as the quartic centrifugal distortion constants are in very good agreement with those CCSD predicted. The line catalogs derived from the accurate spectroscopic parameters should be confident enough for all astronomical observations.

**Table 6.** Experimental ground-state spectroscopic constants of *i*-PrCHO (A-Reduction, F-Representation).

Parameter	<i>Gauche</i>	<i>Trans</i>
A <sup>[a]</sup> / MHz	7494.61375 (19) <sup>[e]</sup>	7707.8665 (48)
B / MHz	4107.511659 (79)	3736.66464 (78)
C / MHz	2980.699617 (78)	2815.08351 (21)
$ \mu_a ,  \mu_b ,  \mu_c $ <sup>[b]</sup> / D	Yes / Yes / Yes	Yes / No / No
$\Delta_J$ / kHz	1.96051 (56)	0.82703 (76)
$\Delta_K$ / kHz	11.39044 (50)	-47.382 (84)
$\Delta_{JK}$ / kHz	-0.29850 (19)	54.5353 (95)
$\delta_J$ / kHz	0.279096 (22)	0.24937 (40)
$\delta_K$ / kHz	3.70440 (20)	27.0688 (53)
$\Phi_J$ / mHz	5.472 (14)	-1.83 (26)
$\Phi_{JK}$ / Hz	0.062991 (94)	0.01070 (34)
$\Phi_K$ / Hz	0.34139 (58)	45.61 (10)
$\Phi_{KJ}$ / Hz	-0.39979 (32)	-4.102 (56)
$\phi_J$ / mHz	0.1924 (53)	-0.883 (135)
$\phi_{JK}$ / Hz	-0.047982 (75)	0.0341 (21)
$\phi_K$ / Hz	-0.093282 (27)	-
N <sup>[c]</sup>	3398	500
$\sigma$ <sup>[d]</sup>	24	28

[a] *A*, *B*, and *C* represent the rotational constants; [b]  $\mu_a$ ,  $\mu_b$ , and  $\mu_c$  are the electric dipole moment components. [c] *N* is the number of measured rotational transitions. [d]  $\sigma$  is the root mean square (rms) deviation of the fit (in kHz). [e] Standard error in parentheses in units of the last digit.

As reported in Section 2.2 of the present Chapter for the linear isomer, in Table 7 we provide the values of the rotational ( $Q_r$ ) and vibrational ( $Q_v$ ) partition functions of *gauche*-*i*-PrCHO, the global minimum and a simple formula that accords to Eq. (3.60) of Gordy & Cook (1970). We computed the values of  $Q_r$  from first principles up to  $J = 150$  at different temperatures, using the Pickett program (Pickett et al. 1998). We performed

theoretical computations at the CCSD/cc-pVTZ level of theory to obtain the frequency of the normal modes (see Table C1 of Appendix C). Afterward, we predicted the vibrational part,  $Q_v$ , using a harmonic approximation and a simple formula that accords to Eq. (3.60) of Gordy & Cook (1970). Finally, we computed the vibrational contribution to the total partition function by taking into account the lowest vibrational modes up to  $1000\text{ cm}^{-1}$ . The complete partition function,  $Q_{tot}$ , is then the product of  $Q_r$  and  $Q_v$ .

**Table 7.** Rotational and vibrational partition functions of *gauche*- i-PrCHO

Temperature (K)	$Q_r^{(a)}$	$Q_v^{(b)}$
9.38	507.7789	1.0000
18.75	1432.7594	1.0007
37.50	4048.2084	1.0277
75.00	11337.6188	1.2408
150.00	32395.4138	2.5019
225.00	59557.2187	5.8979
300.00	91764.5821	14.1391

**Notes.** <sup>(a)</sup>  $Q_r$  is the rotational partition function. <sup>(b)</sup>  $Q_v$  is the vibrational partition function. The total partition function of the molecule (without hyperfine splitting) is  $Q_r \times Q_v$ .

## 2.4 Search for n- and i- butyraldehyde toward Sgr B2 (N1)

### 2.4.1 Observations

The imaging spectral line survey ReMoCA toward Sgr B2(N) was carried out with ALMA during its cycle 4. Belloche et al. (2019) reported the details about the observations and data re-duction of the survey. We summarize here the main features. The full frequency range from 84.1 GHz to 114.4 GHz was covered at a spectral resolution of 488 kHz ( $1.7$  to  $1.3\text{ km s}^{-1}$ ) with five frequency tunings. The achieved sensitivity per spectral channel varies between  $0.35\text{ mJy beam}^{-1}$  and  $1.1\text{ mJy beam}^{-1}$ (rms) de-pending on the setup, with a median value of  $0.8\text{ mJy beam}^{-1}$ . The phase center is located at the equatorial position  $(\alpha, \delta)_{J2000} = (17^{\text{h}}47^{\text{m}}19.87^{\text{s}}, -28^{\circ}22'16.0'')$  which is half-way between the two hot molecular cores Sgr B2(N1) and Sgr B2(N2). The angular resolution (HPBW) ranges from  $\sim 0.3''$  to  $\sim 0.8''$  with a median value of  $0.6''$  that corresponds to  $\sim 4900\text{ au}$  at the

distance of Sgr B2 (8.2 kpc, Reid et al. 2019). We used here an improved version of the data reduction, as described in Melosso et al. (2020).

For this work we analyzed the spectrum obtained toward the position Sgr B2(N1S) at  $(\alpha, \delta)_{J2000} = (17^{\text{h}}47^{\text{m}}19.87^{\text{s}}, -28^{\circ}22'19.48'')$ . This position is offset by about 1'' to the south of the main hot core Sgr B2(N1) and was chosen by Belloche et al. (2019) because of its lower continuum opacity compared to the peak of the hot core. To analyze the spectrum of Sgr B2(N1S), we produced synthetic spectra under the assumption of local thermodynamic equilibrium (LTE) using the software Weeds (Maret et al. 2011). The high densities of the regions where hot core emission is detected in Sgr B2(N) (more than  $1 \times 10^7 \text{ cm}^{-3}$ , see Bonfand et al. 2019) justify this assumption. We derived a best-fit synthetic spectrum for each molecule separately, and then added the contributions of all identified molecules together. Each species was modeled with a set of five parameters: size of the emitting region ( $\theta_s$ ), column density ( $N$ ), temperature ( $T_{rot}$ ), linewidth ( $\Delta V$ ), and velocity offset ( $V_{off}$ ) with respect to the assumed systemic velocity of the source,  $V_{sys} = 62 \text{ km s}^{-1}$ .

#### 2.4.2 *Nondetection of butyraldehyde*

In order to search for *n*- and *i*-butyraldehyde toward Sgr B2(N1S), we relied on the LTE parameters derived for acetaldehyde. The latter is clearly detected in its torsional ground state and its first two torsionally excited states (see Figs. C.1–C.3). We used the spectroscopic predictions of acetaldehyde (tag 44003, version 3) available in the JPL catalog (Pickett et al. 1998), which are based on the work of Kleiner et al. (1996). We selected the transitions of acetaldehyde that are not too much contaminated by emission or absorption from other species to build a population diagram (Fig. C.4). Due to the small energy range covered by the transitions of the torsional ground state reported in this diagram and the residual contamination from unidentified species, the rotational temperature within the torsional ground state is not constrained. Only a fit taking into account the three torsional states allowed us to derive a temperature, which is found to be high, as reported in Table 8. We assumed that the torsional states are populated according

to LTE, which is that the torsional temperature is equal to the rotational temperature, and we adopted a temperature of 250 K to compute the LTE synthetic spectra shown in red in Figs. C.1–C.3. The LTE parameters derived for acetaldehyde are reported in Table 9.

**Table 8.** Rotational temperature of acetaldehyde derived from its population diagram toward Sgr B2(N1S).

Molecule	States <sup>a</sup>	$T_{\text{fit}}^b$ (K)
CH <sub>3</sub> CHO	$v_{mt} = 0, v_t = 1, v_t = 2$	290 (20)

**Notes.** <sup>(a)</sup> Vibrational states that were taken into account to fit the population diagram. <sup>(b)</sup> The standard deviation of the fit is given in parentheses. As explained in Sect. 3 of Belloche et al. (2016) and in Sect. 4.4 of Belloche et al. (2019), this uncertainty is purely statistical and should be viewed with caution. It may be underestimated.

Assuming that the more complex aldehydes propionaldehyde and butyraldehyde trace the same region as acetaldehyde, we produced LTE synthetic spectra for these species adopting the same parameters as acetaldehyde with only the column density left as a free parameter. We employed the spectroscopic predictions derived for butyraldehyde the previous sections and the predictions available in the CDMS catalog (Müller et al. 2005) for *syn* and *gauche* propionaldehyde (tag58505 version 2 and tag 58519 version 1, respectively), which are based on Zingsheim et al. (2017). The LTE synthetic spectra of propionaldehyde and butyraldehyde were used to search for emission of these species in the ReMoCA survey toward Sgr B2(N1S). None of them is detected, as illustrated in Figs. C.5–C.9. The upper limits on their column densities are reported in Table 9.

Table 9 indicates that propionaldehyde is at least  $\sim 5$  times less abundant than acetaldehyde in Sgr B2(N1S). As indicated in Table 9, we get the most stringent upper limit on the column density of *normal*-butyraldehyde from its *ct* conformer. We find that *normal*- butyraldehyde is at least  $\sim 2$  times less abundant than acetaldehyde. Finally, we find that *iso*- butyraldehyde is at least  $\sim 6$  times less abundant than acetaldehyde.



**Table 9.** Parameters of our best-fit LTE model of acetadehyde toward Sgr B2(N1S) and upper limits for propional and butyraldehyde

Molecule	Status <sup>a</sup>	$N_{\text{det}}^b$	Size <sup>c</sup> (")	$T_{\text{rot}}^d$ (K)	$N^e$ ( $\text{cm}^{-2}$ )	$F_{\text{vib}}^f$	$F_{\text{conf}}^g$	$\Delta V^h$ ( $\text{km s}^{-1}$ )	$V_{\text{off}}^i$ ( $\text{km s}^{-1}$ )	$\frac{N_{\text{ref}}^j}{N}$
CH <sub>3</sub> CHO, $v_t = 0^*$	d	15	2.0	250	7.3 (17)	1.32	-	5.0	0.0	1
	$v_t = 1$	d	11	2.0	250	7.3 (17)	1.32	5.0	0.0	1
	$v_t = 2$	d	5	2.0	250	7.3 (17)	1.32	5.0	0.0	1
Upper limits assuming 250 K										
<i>s</i> -C <sub>2</sub> H <sub>5</sub> CHO, $v = 0$	n	0	2.0	250	< 1.3 (17)	4.46	-	5.0	0.0	> 5.0
<i>g</i> -C <sub>2</sub> H <sub>5</sub> CHO, $v = 0$	n	0	2.0	250	< 4.5 (17)	4.46	-	5.0	0.0	> 1.5
<i>cg-n</i> -C <sub>3</sub> H <sub>7</sub> CHO, $v = 0$	n	0	2.0	250	< 5.1 (17)	8.11	2.52	5.0	0.0	> 1.3
<i>ct-n</i> -C <sub>3</sub> H <sub>7</sub> CHO, $v = 0$	n	0	2.0	250	< 3.4 (17)	10.3	1.66	5.0	0.0	> 2.0
<i>g-i</i> -C <sub>3</sub> H <sub>7</sub> CHO, $v = 0$	n	0	2.0	250	< 1.2 (17)	7.97	-	5.0	0.0	> 5.6
Upper limits assuming 170 K										
<i>s</i> -C <sub>2</sub> H <sub>5</sub> CHO, $v = 0$	n	0	2.0	170	< 6.2 (16)	2.07	-	5.0	0.0	> 11
<i>g</i> -C <sub>2</sub> H <sub>5</sub> CHO, $v = 0$	n	0	2.0	170	< 6.2 (17)	2.07	-	5.0	0.0	> 1.1
<i>cg-n</i> -C <sub>3</sub> H <sub>7</sub> CHO, $v = 0$	n	0	2.0	170	< 2.3 (17)	3.26	2.85	5.0	0.0	> 2.9
<i>ct-n</i> -C <sub>3</sub> H <sub>7</sub> CHO, $v = 0$	n	0	2.0	170	< 1.2 (17)	3.96	1.54	5.0	0.0	> 5.5
<i>g-i</i> -C <sub>3</sub> H <sub>7</sub> CHO, $v = 0$	n	0	2.0	170	< 3.8 (16)	3.13	-	5.0	0.0	> 19

**Notes.** <sup>(a)</sup> d: detection. <sup>(b)</sup> Number of detected lines (conservative estimate, see Sect. 3 of Belloche et al. 2016). One line of a given species may mean a group of transitions of that species that are blended together. <sup>(c)</sup> Source diameter (*FWHM*). <sup>(d)</sup> Rotational temperature. <sup>(e)</sup> Total column density of the molecule.  $x$  ( $y$ ) means  $x \times 10^y$ . An identical value for all listed vibrational states of a molecule means that LTE is an adequate description of the vibrational excitation. For propional, the two conformers are considered as parts of the same species and each column density corresponds to the total column density of the molecule. For butyraldehyde, the three isomers and conformers are considered as independent species and each column density corresponds to the column density of the respective isomer or conformer only. <sup>(f)</sup> Correction factor that was applied to the column density to account for the contribution of vibrationally excited states, in the cases where this contribution was not included in the partitionfunction of the spectroscopic predictions. <sup>(g)</sup> Correction factor that was applied to the column density to account for the contribution of other conformers in the cases where this contribution could be estimated but was not included in the partition function of the spectroscopic predictions <sup>(h)</sup> Linewidth (*FWHM*). <sup>(i)</sup> Velocity offset with respect to the assumed systemic velocity of Sgr B2(N1S),  $V_{\text{sys}} = 62 \text{ km s}^{-1}$ . <sup>(j)</sup> Column density ratio, with  $N_{\text{ref}}$  the column density of the previous reference species marked with a \*.

Finally, given that our chemical model suggests that propionaldehyde and butyraldehyde may trace lower temperatures than acetaldehyde (see Sect. 7), we also computed column density upper limits assuming a temperature of 170 K instead of 250 K. This has a significant impact on the upper limits because the vibrational partition functions of propionaldehyde and butyraldehyde are a steep function of temperature in this temperature range. Under this assumption, Table 9 indicates that propionaldehyde, *normal*-butyraldehyde, and *iso*- butyraldehyde are at least 11, 6, and 18 times less abundant than acetaldehyde in Sgr B2(N1S).

## 2.5 *Search for n- and i- butyraldehyde toward the G+0.693-0.027 molecular cloud*

### 2.5.1 *Observations*

The high-sensitivity spectral survey towards the G+0.693-0.027 molecular cloud was carried using the Yebes 40m telescope (Guadalajara, Spain) and the IRAM 30m telescope (Granada, Spain). Detailed information of the observational survey is presented in Rivilla et al. (2021). The observations used the position switching mode, and were centered at  $\alpha(\text{J2000.0}) = 17^{\text{h}}47^{\text{m}}22^{\text{s}}$ ,  $\delta(\text{J2000.0}) = -28^{\circ}21' 27''$ , with the off position at  $(-885'', +290'')$ .

### 2.5.2 *Nondetection of butyraldehyde*

We implemented the spectroscopy presented in this work into the MADCUBA package<sup>1</sup> (version 26/07/2021, Martín et al. 2019), and used the SLIM (Spectral Line Identification and Modeling) tool to generate synthetic spectra of the different conformers of butanal under the assumption of Local Thermodynamic Equilibrium (LTE), which were compared with the observed spectrum. To evaluate if the transitions are blended with emission from other species, we have also considered the LTE model that predicts the total contribution of the more than 120 species identified so far towards G+0.693-0.027 (e.g., Requena-Torres et al. 2008; Zeng et al. 2018; Rivilla et al. 2019, 2020; Jiménez-Serra et al. 2020; Rivilla et al. 2021a,b).

None of the butyraldehyde conformers were clearly detected in the observed data. In most cases the transitions appear strongly blended with brighter lines of abundant species. To compute the upper limits of the column density for the different conformers, we have searched for the brightest predicted spectral features that are not blended with emission from other molecules. The transitions used for *cg-n*-C<sub>3</sub>H<sub>7</sub>CHO, *ct-n*-C<sub>3</sub>H<sub>7</sub>CHO, and *g-i*-C<sub>3</sub>H<sub>7</sub>CHO are  $7_{5,2} \leftarrow 6_{4,3}$  (at 92.905774 GHz),  $10_{0,10} \leftarrow 9_{0,9}$  (at 47.6364885 GHz), and  $6_{0,6} \leftarrow 5_{0,5}$  (at 38.1517752 GHz), respectively. Using MADCUBA, we determined the

---

<sup>1</sup> Madrid Data Cube Analysis on ImageJ is a software developed at the Center of Astrobiology (CAB) in Madrid; <http://cab.intacsic.es/madcuba/Portada.html>.

upper limit of the column density using the  $3\sigma$  value of the integrated intensity (see details in Martín et al. 2019), and the same physical parameters used for *s*-C<sub>2</sub>H<sub>5</sub>CHO (Table 10):  $T_{\text{rot}} = 12$  K,  $V_{\text{LSR}} = 69$  km s<sup>-1</sup> and  $\Delta V = 21$  km s<sup>-1</sup>. Since the excitation temperature of the molecules in G+0.693 are low ( $T_{\text{ex}} = 5\text{--}20$  K), we have not used in this case the vibrational contribution of the partition function ( $Q_{\text{v}}$ ). The two conformers of *n*-butyraldehyde were modeled as independent species, without taking into account their zero-point energy. Given the low excitation temperatures in G+0.693, we do not assume here that the two conformers are populated thermally, and hence we do not apply a conformer correction. The derived upper limits for the two conformers of *normal*-butyraldehyde and for the *g* conformer of *iso*-butyraldehyde are presented in Table 10. The upper limit of the lowest energy conformer, the *ct*-conformer of *normal*-butyraldehyde, is  $< 0.8 \times 10^{13}$  cm<sup>-2</sup>.

As we have done for Sgr B2(N1S), we have also fitted the acetaldehyde (CH<sub>3</sub>CHO) emission towards G+0.693. The details of the analysis is presented in Table C.2 of [Appendix C \(Section C.1\)](#) and the derived parameters are shown in Table 10.

We show in Table 11 the molecular ratios of the different aldehydes, using also the LTE results of acetaldehyde (CH<sub>3</sub>CHO) and *s*-C<sub>2</sub>H<sub>5</sub>CHO from Requena-Torres et al. (2008) and Rivilla et al. (2020), respectively. The acetaldehyde / *s*-propionaldehyde ratio is  $\sim 7$ , while the *s*-propionaldehyde / *ct-n*-butanal ratio is  $> 9$  and the acetaldehyde / *ct-n*-butyraldehyde ratio is  $> 65$ . These values are consistent with those found in Sgr B2(N1S), but the constraints obtained for G+0.693 are more stringent than for Sgr B2(N1S) (Table 11). The relative molecular ratios found in G+0.693 indicate that each jump in complexity in the aldehyde family implies an abundance drop of at least one order of magnitude, similarly to other chemical families such as alcohols, thiols, and isocyanates (Rodríguez-Almeida et al. 2021a,b).

**Table 10.** Derived physical parameters for acetaldehyde, propionaldehyde, and butyraldehyde towards the G+0.693-0.027 molecular cloud.

Molecule	$N$ ( $\times 10^{13}$ cm $^{-2}$ )	$T_{\text{rot}}$ (K)	$V_{\text{LSR}}$ (Km s $^{-1}$ )	$\Delta V$ (Km s $^{-1}$ )	Abundance <sup>a</sup> ( $\times 10^{-10}$ cm $^{-2}$ )	Ref. <sup>b</sup>
CH <sub>3</sub> CHO	52 ± 1	9.2 ± 0.1	69.3 ± 0.1	21.0 ± 0.2	39	(1)
<i>s</i> -C <sub>2</sub> H <sub>5</sub> CHO	7.4 ± 1.5	12 ± 0.8	69	21	5 ± 2	(2)
<i>ct-n</i> -C <sub>3</sub> H <sub>7</sub> CHO <sup>c</sup>	< 0.8	12	69	21	< 0.6	(3)
<i>cg-n</i> -C <sub>3</sub> H <sub>7</sub> CHO <sup>c</sup>	< 0.5	12	69	21	< 0.4	(3)
<i>g-i</i> -C <sub>3</sub> H <sub>7</sub> CHO <sup>c</sup>	< 1.6	12	69	21	< 1.2	(3)

**Notes.** <sup>(a)</sup>We adopted  $N_{\text{H}_2} = (1.4 \pm 0.3) 10^{23}$  cm $^{-2}$ , from Martín et al. (2008). <sup>(b)</sup>References: (1) Requena-Torres et al. (2008); (2) Rivilla et al.(2020); (3) This work. <sup>(c)</sup>For *n*-butyraldehyde, the two conformers were modeled as independent species, without applying a conformer correction. For *i*-butyraldehyde, the *gauche* conformer is considered as an independent species and its column density corresponds to the column density of this conformer.

**Table 11.** Molecular ratios of aldehydes towards Sgr B2(N1S) and G+0.693-0.027.

Ratio	Sgr B2(N1S)	G+0.693-0.027
CH <sub>3</sub> CHO / C <sub>2</sub> H <sub>5</sub> CHO <sup>(a)</sup>	> 5-11	> 7
C <sub>2</sub> H <sub>5</sub> CHO <sup>(a)</sup> / C <sub>3</sub> H <sub>7</sub> CHO <sup>(b)</sup>	-	> 9
CH <sub>3</sub> CHO / C <sub>3</sub> H <sub>7</sub> CHO <sup>(b)</sup>	> 2-6	> 65

**Notes.** <sup>(a)</sup>Using *s*-conformer; <sup>(b)</sup>Using the *ct-n* conformer.

## 2.6 Discussion: Interstellar chemistry of acetaldehyde, propionaldehyde and butyraldehyde

Of the three aldehydes for which column densities or upper limits are presented in Tables 9 and 10, typical astrochemical kinetics models consider only acetaldehyde. However, a treatment for the interstellar formation and destruction of propionaldehyde was included in the network of Garrod (2013), as part of a broader effort to simulate the chemistry of large organic molecules including glycine. Production of propionaldehyde could proceed through the addition of various functional-group radicals on dust-grain surfaces during the star-formation process. The network was developed further in various publications, including Garrod et al. (2017), who added the butyl cyanides, pentanes, and

some associated species. The larger aldehydes were not included in that network, however. The most recent usage of this chemical network was presented by Garrod et al. (2021). In that work, new treatments of so-called nondiffusive reaction mechanisms on grain surfaces and within the dust-grain ice mantles (Jin & Garrod 2020) were applied to a full hot-core chemical simulation. This more complete treatment of grain chemistry enabled the possible production of complex organics at very cold, early times in the evolution of a hot core, as well as during the warm-up of the dust, and – where permitted by the viability of appropriate reactions – in the gas phase, following the thermal desorption of the icy dust-grain mantles.

In spite of the absence of butylaldehyde from the chemical networks, the results of the Garrod et al. (2021) astrochemical kinetics models are instructive, especially in comparison with the observational results toward Sgr B2(N1S). The models showed their best agreement with other molecular abundance data from the ReMoCA survey when a long warm-up timescale (i.e. the slow model setup) was used; the same model is compared here with the present data. The peak gas-phase fractional abundance of acetaldehyde with respect to total hydrogen in the model is around  $2.5 \times 10^{-7}$ , providing a peak ratio with gas-phase methanol of 1:30. This value coincides exactly with the observational ratio, using the methanol column density of  $2 \times 10^{19} \text{ cm}^{-2}$  provided by Motiyenko et al. (2020) (see their Table 3).

Around 30% of acetaldehyde is formed on the grains prior to the desorption of the ices, of which around one third is formed at temperatures below 20 K, mainly through surface chemistry as the ices build up. Reactions between  $\text{CH}_2$  with CO are important in producing ketene,  $\text{CH}_2\text{CO}$ , which is further hydrogenated to acetaldehyde on the grain surfaces by mobile atomic H. We would also expect this cold, grain-surface mechanism to be relevant to acetaldehyde production on grains in G+0.693.

In the hot-core models, some  $\text{CH}_3\text{CHO}$  is also formed at intermediate temperatures (20 - 100 K), as the result of photo processing of the ices by cosmic ray-induced UV photons. However, approximately two thirds of total acetaldehyde production during the entire hot core evolution is calculated to occur in the gas phase, through the barrierless reaction  $\text{O} + \text{C}_2\text{H}_5 \rightarrow \text{CH}_3\text{CHO} + \text{H}$  (Tsang & Hampson 1986). This reaction is fed by

the release of trapped solid-phase  $C_2H_5$  into the gas when the ices sublimate at temperatures greater than 100 K, while atomic O is a product of the gas-phase destruction of more stable species such as  $CO_2$  that are released at around the same time. The dominance of the gas-phase reaction in producing acetaldehyde results in a delay between the sublimation of the dust-grain ice mantles and the attainment of the peak  $CH_3CHO$  abundance. Furthermore, due to the gradually increasing temperature in the model, this delay in peak abundance occurs at an elevated temperature of 257 K, which is in good agreement with the rotational temperature used in the spectral model of Sgr B2(N1S).

The chemical model indicates that propionaldehyde should reach a peak gas-phase abundance around 60 times lower than that of acetaldehyde (although the fast warm-up timescale model suggests a value closer to 10). As shown in Table 11, this is consistent with the observational upper limits using a rotational temperature of 250 K as well as with the ratio reported toward G+0.693. (Note that the chemical models do not explicitly distinguish conformers). The excitation temperature used to determine the propionaldehyde upper limit is chosen to be consistent with acetaldehyde.

Unlike the treatment for acetaldehyde, the chemical model includes only grain-surface mechanisms for the production of propionaldehyde; this solid-phase material is preserved on the grains at low temperatures, and then released into the gas phase when temperatures are sufficient to allow it and other grain-surface species to desorb thermally, or due to a non-thermal shock-induced desorption process, such as that present in G+0.693. Due to its direct release into the gas phase, the abundance of propionaldehyde reaches its peak at an ambient gas temperature of 170 K, substantially lower than the 250 K assumed here for its excitation temperature, and lower than the 257 K at which acetaldehyde reaches its peak in the model. In order to test the influence on observational results of this lower temperature of release, which is predicated on a uniquely grain-surface origin for propionaldehyde, an excitation temperature of 170 K was also tested in the spectral analysis (see bottom part of Table 9). The model results nevertheless remain consistent with observations.

However, a gas-phase formation route for propionaldehyde may in fact be plausible; Tsang (1988) recommends a rate of  $1.6 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}$  for the  $O + n\text{-}C_3H_7$  reaction,

identical to the recommended rate of the equivalent process for acetaldehyde production. Formation of an aldehyde is the preferred outcome in either reaction. The same rate was also proposed by Tsang (1988) for the reaction of the related radical *i*-C<sub>3</sub>H<sub>7</sub>, leading to acetone production

With such a reaction for propionaldehyde included in the astrochemical model, some enhancement in peak abundance would be expected. By comparison with the behavior of acetaldehyde and the other species in the model, it is possible to make a crude prediction as to the strength of influence of this mechanism. The models predict *n*-C<sub>3</sub>H<sub>7</sub> in the ice to be around five times less abundant than C<sub>2</sub>H<sub>5</sub> immediately prior to ice sublimation, suggesting a five times lower base rate of gas-phase propionaldehyde production versus acetaldehyde. To gauge the importance of this lower rate, it is also noteworthy that only some of the gas-phase acetaldehyde is formed through the gas-phase reaction, with the rest coming directly from the grain surfaces. On this basis it might be reasonable to expect that, if gas-phase production of propionaldehyde were to dominate over any grain-surface mechanism, it would reach a maximum abundance of 10% that of acetaldehyde. This crude estimate also is approximately in keeping with the upper limits.

Could such a gas-phase mechanism also produce butyraldehyde? On the assumption that it could – through the reaction of *n*-C<sub>4</sub>H<sub>9</sub> with atomic oxygen – then the abundance of butyraldehyde could be comparable to that of propionaldehyde. All four of the C<sub>4</sub>H<sub>9</sub> radicals in the network (i.e. the straight-chain and branched forms, with the radical site on either the primary or secondary carbon) achieve ice abundances similar to the C<sub>3</sub>H<sub>7</sub> radicals. Two of these might be expected to form ketones rather than aldehydes. The other two, which would produce either *n*- or *i*-butyraldehyde, reach very similar abundances, indicating no obvious distinction between the two forms.

If aldehyde production in G+0.693 instead occurs mainly on dust grains, followed by a rapid shock-induced release, then the cold, grain-surface abundances derived from the early stages of the hot core models may be comparable to the observed gas-phase abundances. The models suggest a purely solid-phase ratio of acetaldehyde to propionaldehyde of  $\sim 7.7$ . This is in excellent agreement with the observational ratio of  $\sim 7$ . However, it is unclear how the shock processing might further affect chemical

abundances, so this solid-phase ratio might not be accurately reflected in the gas-phase abundances of the two aldehydes.

Thus, based on estimates derived from the astrochemical models, the observational upper limits on both propionaldehyde and butyraldehyde could be very close to their true abundances toward Sgr B2(N1S) and toward G+0.693. In Sgr B2(N1S) this assumes that gas-phase production is strong. Weak gas-phase production could produce values rather lower than the upper limits. In G+0.693, the observed ratios are well reproduced without the need for gas-phase chemistry.

### 3. CONCLUSIONS

We are finally starting to fill the gap between observational and laboratory data that existed for several decades. In this context, we have investigated and analyze the millimeter- and submillimeter-wave spectrum of *n*- and *i*-PrCHO from 75 to 325 GHz. With our experimental data set we can predict the rotational transitions for both isomers with narrow linewidths (FWHM) of less than  $0.1 \text{ km s}^{-1}$  at 100 GHz, expressed in its equivalent radial velocity, which is more than enough to perform a confident astronomical search for these species. Afterward, we employed the newly determined rotational parameters to prepare accurate line catalogs to be used in the subsequent search for the molecule toward Sgr B2(N1) and toward the G+0.693-0.027 molecular cloud. In addition, we performed a comprehensive study including the search for shorter aldehydes toward both astronomical sources. We present the following conclusions for the astronomical results:

1. We report a nondetection of *n*- and *i*-butyraldehyde toward Sgr B2(N1S) with ALMA. Depending on the assumed temperature, we find that propionaldehyde, *n*-butyraldehyde, and *i*-butyraldehyde are at least 5–11, 2–6, and 6–18 times less abundant than acetaldehyde, respectively.
2. We also report a nondetection of both isomers toward G+0.693 based on Yebes 40m and IRAM 30m observations. In this case, propionaldehyde is  $\sim 7$  less



abundant than acetaldehyde, while the propionaldehyde / *n*-butyraldehyde ratio is  $> 9$  and the acetaldehyde / *n*-butyraldehyde ratio is  $> 65$ .

3. Finally, our astrochemical models suggest the observational upper limits on both propionaldehyde and butyraldehyde could be close to their true abundances toward Sgr B2(N1S) and G+0.693. This assumes that gas-phase production of aldehydes is dominant in Sgr B2(N1S), while in G+0.693 aldehyde production on dust grains is sufficient to reproduce observed molecular ratios. If gas-phase mechanisms are not effective in Sgr B2(N1S) then the true values could be substantially less than the upper limits observed here.

The most stringent constraints on the relative abundances of the three aldehydes in the interstellar medium were obtained for G+0.693-0.027. Our results show that with each incremental increase in the size of the alkyl group in the aldehyde molecule there is a drop of approximately one order of magnitude in abundance.

#### 4. EXPERIMENTAL SECTION

The room-temperature rotational spectra were recorded separately from 75 to 325 GHz using the millimeter-wave absorption spectrometer at the University of Valladolid (for a detailed description of the instrument see Daly et al. 2014, Alonso et al. 2016 and Chapter III: Methodology). In both experiments, we filled our double-pass free-space glass cell with an optimum gas pressure of about 20  $\mu$ bar (no external heating was needed for any of the molecules), which was maintained during the whole course of the experiment. The millimeter-wave radiation was generated by multiplying the fundamental signal of an Agilent E8257D microwave synthesizer (up to 20 GHz) by a set of active and passive amplifier-multiplier chains (multiplication factors of 6, 9, 12 and 18 in this case). Also, the output of the synthesizer was frequency modulated at frequency of 10.2 kHz, with modulation depths varying between 20 to 35 kHz. We measured the rotational lines employing an average of two up and down frequency scans, also using a Gaussian profile function within the AABS package (Kisiel et al. 2005). The experimental uncertainty of the unblended symmetric lines is estimated to be better than 30 kHz.

## 5. REFERENCES

- Abplanalp, M. J., Góbi, S., Bergantini, A., Turner, A. M., Kaiser, R. I., *ChemPhysChem* 2018, 19, 556
- Abplanalp, M. J. & Kaiser, R. I. 2019, *Physical Chemistry Chemical Physics (Incorporating Faraday Transactions)*, 21, 16949.
- Alonso, E. R., Kolesníková, L., Tercero, B., et al. 2016, *ApJ*, 832, 42.
- Alonso, E. R., McGuire, B. A., Kolesníková, L., et al. 2019, *ApJ*, 883, 18.
- Altwegg, K., Balsiger, H., Bar-Nun, A., et al. 2016, *Science Advances*, 2, e1600285.
- Aponte, J. C., Whitaker, D., Powner, M. W., et al. 2019, *ACS Earth and Space Chemistry*, 3, 463.
- Belloche, A., Garrod, R. T., Müller, H. S. P., et al. 2009, *A&A*, 499, 215.
- Belloche, A., Garrod, R. T., Müller, H. S. P., et al. 2014, *Science*, 345, 1584.
- Belloche, A., Müller, H. S. P., Garrod, R. T., et al. 2016, *A&A*, 587, A91
- Belloche, A., Meshcheryakov, A. A., Garrod, R. T., et al. 2017, *A&A*, 601, A49 and references therein.
- Belloche, A., Garrod, R. T., Müller, H. S. P., et al. 2019, *A&A*, 628, A10.
- Bonfand, M., Belloche, A., Garrod, R. T., et al. 2019, *A&A*, 628, A27 Jin, M. & Garrod, R. T. 2020, *ApJS*, 249, 26
- Burkhardt, A. M., Long Kelvin Lee, K., Bryan Changala, P., et al. 2021, *ApJ*, 913, L18.
- Burton, A. S., Stern, J. C., Elsila, J. E. et al. 2012, *Chem. Soc. Rev.*, 41, 5459- 5472.
- Cernicharo, J., Agúndez, M., Cabezas, C., et al. 2021, *A&A*, 649, L15.
- Daly, A., Kolesníková, L. Mata, S., J. Alonso, J. L. 2014, *Journal of Molecular Spectroscopy*, 306, 11.
- Dunning, T.H. 1989, *J. Chem. Phys.*, 90, 1007.
- Durig, J. R., Guirgis, G. A., Little, T. S., et al. 1989, *J. Chem. Phys.*, 91, 738.
- Dwivedi, Y. & Rai, S.B. 2008, *Vibrational Spectroscopy*, 49, 2, 278-283.
- Gaussian 16, Revision A.03, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2016.
- Garrod, R. T. 2013, *ApJ*, 765, 60
- Garrod, R. T., Belloche, A., Müller, H. S. P. & Menten, K. M. 2017, *A&A*, 601, A48
- Garrod, R. T., Jin, M., Matis, K. A., Jones, D., Willis, E. R. & Herbst, E. 2021, *ApJS*, accepted

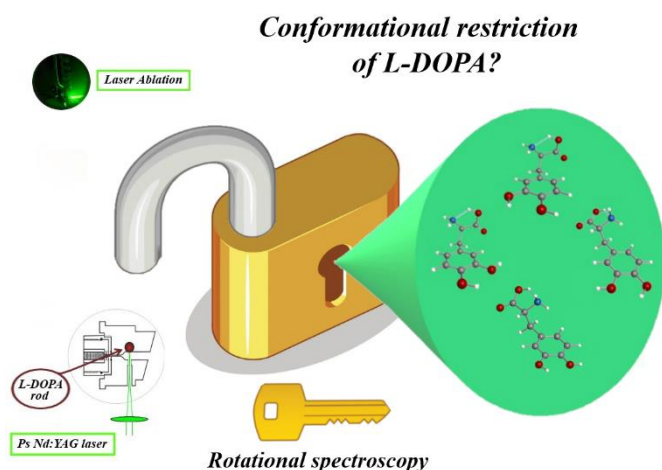
- Gottlieb, C. A. 1973, *Molecules in the Galactic Environment*, 181
- Gordy, W., Cook, R. L. *Microwave Molecular spectra*, Interscience Pub, 1970, New York, Chap. III, 58-70.
- Grimme, S.J. 2006, *Chem. Phys.*, 124, 034108.
- Grimme, S.J., Ehrlich, S., & Goerigk, L. 2011, *J. Comp. Chem.*, 32, 1456.
- Hartwig, H. & Dreizler, H. 1996, *Zeitschrift Naturforschung Teil A*, 51, 923. doi:10.1515/zna-1996-0807.
- Helgaker, T., Klopper, W., Koch, H., & Noga, J. 1997, *JChPh*, 106, 9639.
- Hollis, J. M., Lovas, F. J., & Jewell, P. R. 2000, *ApJ*, 540, L107.
- Hollis, J. M., Jewell, P. R., Lovas, F. J., et al. 2004, *ApJ*, 610, L21.
- Hotopp, K. M., Vara, V. V., & Dian, B. C. 2012, *Journal of Molecular Spectroscopy*, 280, 104.
- Jiménez-Serra, I., Martín-Pintado, J., Rivilla, V. M., et al. 2020, *Astrobiology*, 20, 1048.
- Jørgensen, J. K., Belloche, A., & Garrod, R. T. 2020, *ARA&A*, 58, 727.
- Klimkowski, V.J., Van Nuffel, P., Van Den Enden, L., Van Alsenoy, C., Geise, H.J., Scarsdale, J.N. and Schäfer, L. 1984, *J. Comput. Chem.*, 5: 122-128.
- Kisiel, Z., Pszczółkowski, L., Medvedev, I. R., et al. 2005, *Journal of Molecular Spectroscopy*, 233, 231.
- Kisiel, Z., Pszczółkowski, L., Drouin, B. J., et al. 2012, *Journal of Molecular Spectroscopy*, 280, 134.
- Koga, T. & Naraoka, H. 2017, *Scientific Reports*, 7, 636.
- Kolesniková, L., Alonso, E. R., Mata, S., et al. 2017, *ApJS*, 229, 26.
- Kolesniková, L., Alonso, E. R., Mata, S., et al. 2017, *ApJS*, 233, 24.
- Loomis, F. W. & Wood, R. W. 1928, *Physical Review*, 32, 223.
- Maret, S., Hily-Blant, P., Pety, J., et al. 2011, *A&A*, 526, A47
- Martín, S., Requena-Torres, M. A., Martín-Pintado, J., et al. 2008, *ApJ*, 678, 245.
- Martín, S., Martín-Pintado, J., Blanco-Sánchez, C., et al. 2019, *A&A*, 631, A159.
- Melosso, M., Belloche, A., Martin-Drumel, M.-A., et al. 2020, *A&A* in press, <https://doi.org/10.1051/0004-6361/202038466>
- McCarthy, M. C. & McGuire, B. A. 2021, *Journal of Physical Chemistry A*, 125, 3231.
- McCarthy, M. C., Lee, K. L. K., Loomis, R. A., et al. 2021, *Nature Astronomy*, 5, 176.
- McGuire, B. A. 2018, *ApJS*, 2.
- McGuire, B. A., Burkhardt, A. M., Kalenskii, S., et al. 2018, *Science*, 359, 202.
- McGuire, B. A., Loomis, R. A., Burkhardt, A. M., et al. 2021, *Science*, 371, 1265.
- Müller, H. S. P., Schlöder, F., Stutzki, J., et al. 2005, *Journal of Molecular Structure*, 742, 215
- Motiyenko, R. A., Belloche, A., Garrod, R. T., et al. 2020, *A&A*, 642, A29.
- Loomis, R. A., Shingledecker, C. N., Langston, G., et al. 2016, *MNRAS*, 463, 4175.

- Piart-Goypiro, A., Baron, M. H., Belloc, J., et al. 1991, *Spectrochimica Acta Part A: Molecular Spectroscopy*, 47, 363.
- Pickett, H. M. 1991, *Journal of Molecular Spectroscopy*, 148, 371.
- Pickett, H. M., Poynter, R. L., Cohen, E. A., et al. 1998, *J. Quant. Spectr. Rad. Transf.*, 60, 883
- Raghavachari, K., Trucks, G.W., Pople, J.A., & Head-Gordon, M. 1989, *CPL*, 157, 479.
- Pizzarello, S., Shock, E. 2010, *Cold Spinning Harb. Perspect. Biol.*, 2, 1-20.
- Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2019, *ApJ*, 885, 131.
- Remijan, A. J., Hollis, J. M., Lovas, F. J., et al. 2008, *ApJ*, 675, L85.
- Requena-Torres, M. A., Martín-Pintado, J., Martín, S., et al. 2008, *ApJ*, 672, 352.
- Rivilla, V. M., Martín-Pintado, J., Jiménez-Serra, I., et al. 2019, *MNRAS*, 483, L114.
- Rivilla, V. M., Martín-Pintado, J., Jiménez-Serra, I., et al. 2020, *ApJ*, 899, L28.
- Rivilla, V. M., Jiménez-Serra, I., Martín-Pintado, J., et al. 2021, *PNAS*, 118, 2101314118.
- Rivilla, V. M., Jiménez-Serra, I., García de la Concepción, J., et al. 2021, *MNRAS*, 506, L79.
- Rodríguez-Almeida, L. F., Jiménez-Serra, I., Rivilla, V. M., et al. 2021, *ApJ*, 912, L11.
- Rodríguez-Almeida, L. F., Rivilla, V. M., Jiménez-Serra, I., et al. 2021, *A&A*, 654, L1.
- Sbrana, G. & Schettino, V. 1970, *Journal of Molecular Spectroscopy*, 33, 100.
- Smirnov, I. A., Alekseev, E. A., Ilyushin, V. V., et al. 2014, *Journal of Molecular Spectroscopy*, 295, 44
- Snyder, L. E., Buhl, D., Zuckerman, B., et al. 1969, *Phys. Rev. Lett.*, 22, 679.
- Stanton, J.F., Gauss, J., Harding, M.E., & Szalay, P.G. 2013, *CFOUR*, A Quantum Chemical Program Package.
- Stiefvater, O. L. 1986, *Zeitschrift Naturforschung Teil A*, 41, 483 and references therein.
- Tsang, W. & Hampson, R. F. 1986, *J. Phys. Chem. Ref. Data*, 15, 1087
- Tsang, W. 1988, *J. Phys. Chem. Ref. Data*, 17, 887
- Watson, J. K. G. 1977, *Journal of Molecular Spectroscopy*, 66, 500.
- Woon, D.E. & Dunning, T.H. 1993, *J. Chem. Phys.*, 98, 1358.
- Zeng, S., Jiménez-Serra, I., Rivilla, V. M., et al. 2018, *MNRAS*, 478, 2962.
- Zingsheim, O., Müller, H. S. P., Lewen, F., et al. 2017, *Journal of Molecular Spectroscopy*, 342, 125

## CHAPTER VII. THE SHAPE OF BIOMOLECULES, FLUXIONAL SYSTEMS AND ORGANIC SYNTHONS

### CHAPTER VIIa. UNLEASHING THE SHAPE OF L-DOPA AT LAST

Adapted from Sanz Novo et al. 2021 Phys. Chem. Chem. Phys., 2021, Accepted Manuscript.

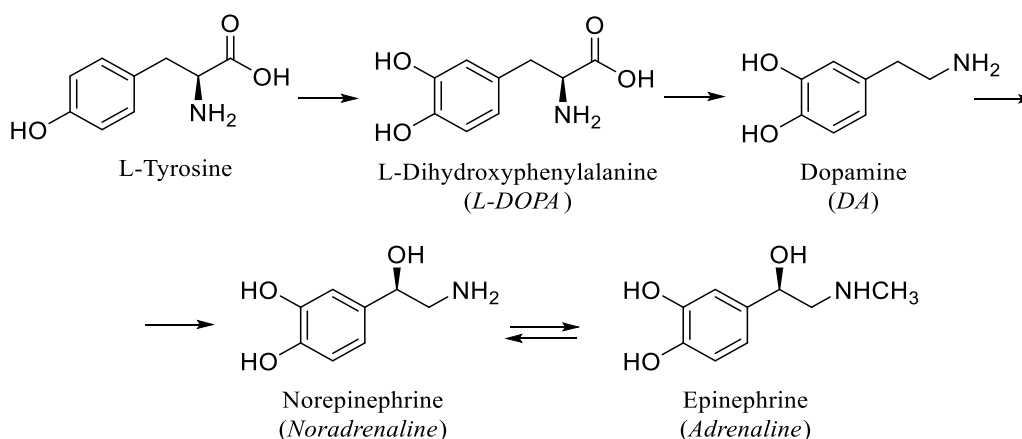


Herein we report the first rotational study of neutral L-DOPA, an extensively used supramolecular synthon and an amino acid precursor of the neurotransmitters dopamine, norepinephrine (noradrenaline) and epinephrine (adrenaline), using broadband and narrowband Fourier transform microwave spectroscopies coupled with a laser ablation

vaporization system. The spectroscopic parameters derived from the analysis of the rotational spectrum conclusively identify the existence of three distinct conformers of L-DOPA in the supersonic jet. The analysis of the  $^{14}\text{N}$  nuclear quadrupole coupling hyperfine structure further revealed the orientation of the N-bearing functional group, proving the existence of stabilizing N-H $\cdots\pi$  interactions for the observed structures.

## 1. INTRODUCTION

Neurotransmitters are specific molecules that transfer messages between neurons or from neurons to (target) muscle cells. The molecular recognition undergoing this process is commonly explained using a “key and lock” analogy; a particular neurotransmitter emitted from a presynaptic neuron can bind to and activate a specific receptor on a postsynaptic neuron (Wieland et al. 1996; Liapakis et al. 2000). Among all neurotransmitters, dopamine (DA) plays a vital role in motivational control (rewarding, aversive and alerting) in the brain (Bromberg-Martin et al. 2010). Calabresi et al. described the role of DA in several brain processes, such as the modulation of behavior and cognition, sleep, attention and learning (Calabresi et al. 2007).



**Figure 1.** Schematic classical biosynthesis and degradation pathway of dopamine [adapted from (Meiser et al. 2013)].

As shown in Fig. 1, the amino acid L-DOPA is one of the initial pieces in the metabolic route of neurotransmitters dopamine, norepinephrine (noradrenaline) and epinephrine (adrenaline) (Meiser et al. 2013), and it is also entangled in the biosynthesis of other secondary metabolites, such as melanines and betalaines (Kamal et al. 2014). This molecule belongs to the well-known catecholamine family, another member of the enediol family, which is of great relevance in biomedical imaging and in developing titanium dioxide (TiO<sub>2</sub>)-based solar cells (Rajh et al. 2004, Clarke et al. 2006; Li et al. 2009). It is also employed to treat Parkinson’s disease since it can cross the blood-brain barrier (BBB) and get decarboxylated into dopamine (Simuni et al. 2008; Shulman et al. 2001).

Like other catecholamines, L-DOPA presents high flexibility, leading to a large number of stable conformations (see Fig. 2). The knowledge of its most favored forms is thus of great importance because if a specific conformation is much more stable than the other conformations, this particular conformation will be preferred in the “key and lock” recognition process at the receptor site. In addition, of paramount relevance is that L-DOPA and its derivatives are able to self-assemble into a plethora of materials of outstanding physical properties, showing a wide range of applications within the field of Material Science (Giui et al. 2021). What is more, our understanding of bio-adhesion and bio-mineralization processes strongly depends on the mechanism through which amino acids attach to mineral surfaces (Hamhey et al. 2012). L-DOPA is implied in one bio-adhesion mechanism that takes place through proteins containing this amino acid (Lee et al. 2006; Lin et al. 2007; Guvendiren et al. 2009). In this context, this molecular system has been investigated in the condensed phase by several spectroscopic techniques, highlighting surface-enhanced Raman spectroscopy (SERS) and scanning tunneling microscopy (STM) (Weinhold et al. 2006).

It is well known that the balance of intramolecular interactions drives the conformational panoramas of neurotransmitters and their precursors. Moreover, their competence with intermolecular interactions plays a vital part in the molecular recognition process at the receptor site. Such a fragile equilibrium of multiple inter- and intramolecular forces results in a topic of major interest for biomedical chemistry. To better understand this interplay, it is often mandatory to study these properties in the absence of intermolecular interactions. Thus, the environment provided by gas-phase experiments, where molecules are isolated and free of intermolecular interactions, is needed (Snoek et al. 2003; Carbacal et al. 2005). This approach can be considered a first step to analyze neurotransmitters' conformation in a more complex environment such as that of a biological medium.

Within this framework, several gas-phase spectroscopic techniques, such as ultraviolet-ultraviolet (UV-UV) hole burning and infrared (IR) dip spectroscopy, have been used in the past to explore the conformational landscape of L-DOPA (Mitsuda et al. 2010, Ishiuchi et al. 2011). In these studies, the authors claimed a significant conformational reduction for L-DOPA based on the observation of one single conformer, compared to the previous results reported for tyrosine using the same approach [21]. Also, they further suggested that this severe conformational restriction is induced by the presence of a

catechol ring (an additional phenolic OH group than tyrosine). However, these results are in sharp contrast to the gas-phase study of dopamine, its direct catecholamine product (Cabezas et al. 2013), where seven distinct conformations are conclusively identified; no experimental evidence of any conformational restriction is observed. Delving into this idea, Ishiuchi et al. reported a subsequent low-resolution assignment of the previously detected L-DOPA conformer (Ishiuchi et al. 2011), which merely relies on a comparison between the experimental and computed IR spectra. Nevertheless, this approach can be puzzling since the vibrational band-shifts associated with distinct conformations are often difficult to model accurately. Considering all of this, a conclusive high-resolution spectroscopic study in the gas phase is therefore mandatory to unveil the structure of L-DOPA.

Among all high-resolution experimental techniques, microwave spectroscopy coupled with an ultra-fast laser ablation vaporization source in a supersonic expansion has proved to be a robust tool in the investigation of the conformational landscape of neurotransmitters in the gas phase (López et al 2007; Cabezas et al. 2012, 2013) as well as deciphering the conformational panorama of several relevant biomolecules (Alonso et al. 2015; Alonso et al. 2021). One remarkable feature of microwave studies is that every conformer possesses its own discrete set of transitions. Hence, the conformational assignments are based on the consistency of many observed rotational lines. The different conformers that are populated enough in the supersonic jet possess distinct spectroscopic parameters and can be therefore unequivocally identified through the analysis of the rotational spectrum.

In the last decade, several attempts have been performed in our laboratory to record the rotational spectrum of L-DOPA without success. L-DOPA is a solid with a very high melting point (276 - 278 °C) and chemically unstable, promptly reacting when exposed to the atmosphere. Moreover, it experiments decomposition problems when exposed to laser fluencies, preventing an easy transfer to the gas phase with laser ablation. Although the investigation of its closely related dopamine (Cabezas et al. 2013) has already been performed several years ago, the high-resolution study of L-DOPA has remained a challenging task for rotational spectroscopy so far. It was not until very recently that the developments of new sample preparation procedures and careful control of the laser-induced fragmentation processes (Kolesníková et al. 2021) have enabled its first rotational characterization. We have implemented the abovementioned improvements in



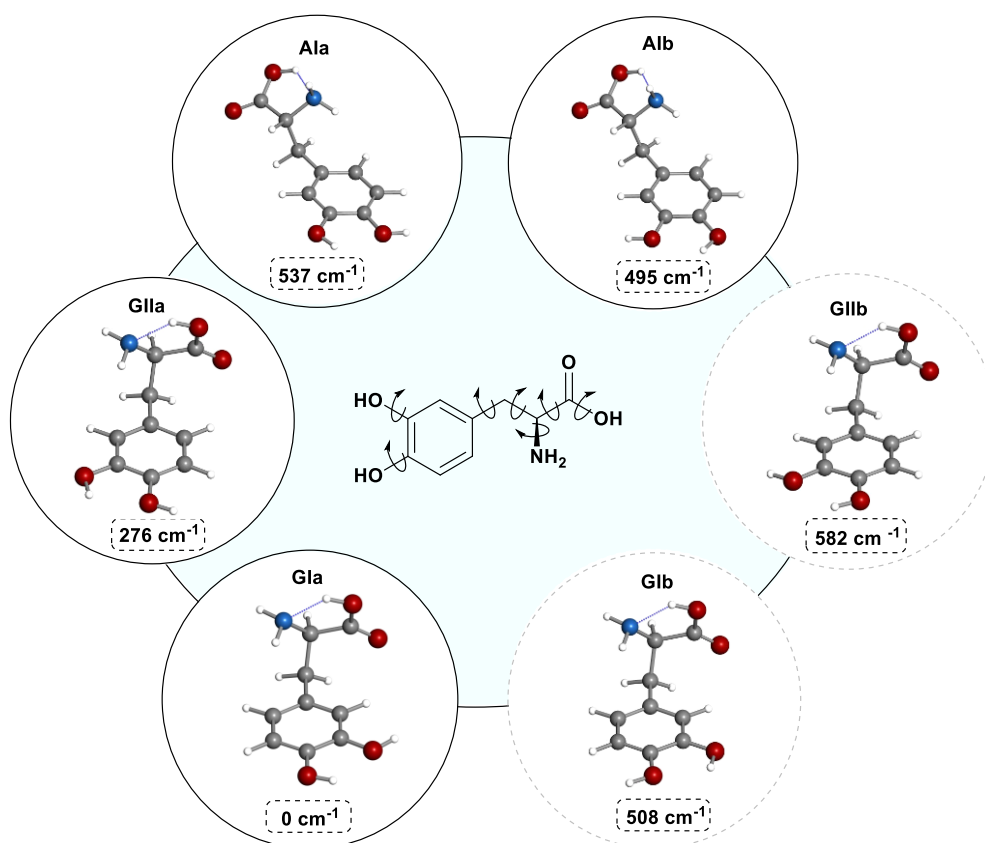
our broadband chirped pulse Fourier transform microwave spectrometer equipped with a laser ablation vaporization system (LA-CP-FTMW) to measure the rotational spectrum of L-DOPA.

We will show in the following sections a complete theoretical and experimental exploration of the conformational landscape of L-DOPA. Moreover, the role of the N–H $\cdots\pi$  interaction will be discussed based on the analysis  $^{14}\text{N}$  nuclear quadrupole hyperfine structure

## 2. RESULTS AND DISCUSSION

### 2.1. *Molecular modeling*

The conformational space of L-DOPA was explored using a systematic approach to guide the spectral searches. In a first step, fast molecular mechanics methods (MMFFs) with the “Large scales Low Mode” and a Monte Carlo-based search algorithms in MacroModel ([www.schrodinger.com](http://www.schrodinger.com)) were used to locate all possible minima energy structures on the potential hypersurface. More than 100 distinct structures were found below an energetic window of 35 kJ/mol. Each molecular conformation obtained in this way was further geometrically optimized at the B3LYP/ aug-cc-pVTZ level (Lee et al. 1988, Dunning, 1989), with Grimme dispersion (Grimme et al. 2010) taken into consideration. Frequency calculations were finally performed to verify that no imaginary frequencies occurred and to compute the Gibbs Free Energies. The Gaussian 16 package (Gaussian 16, Revision A.03) was employed to carry out all the theoretical computations.



**Figure 2.** Chemical structure of the six lowest-energy conformers of L-DOPA predicted by theoretical calculations. Spectroscopically characterized conformers are marked with a solid black circle and tentatively observed forms are indicated with a dashed gray circle. The relative total energy is given with respect to the global minimum (in  $\text{cm}^{-1}$ ).

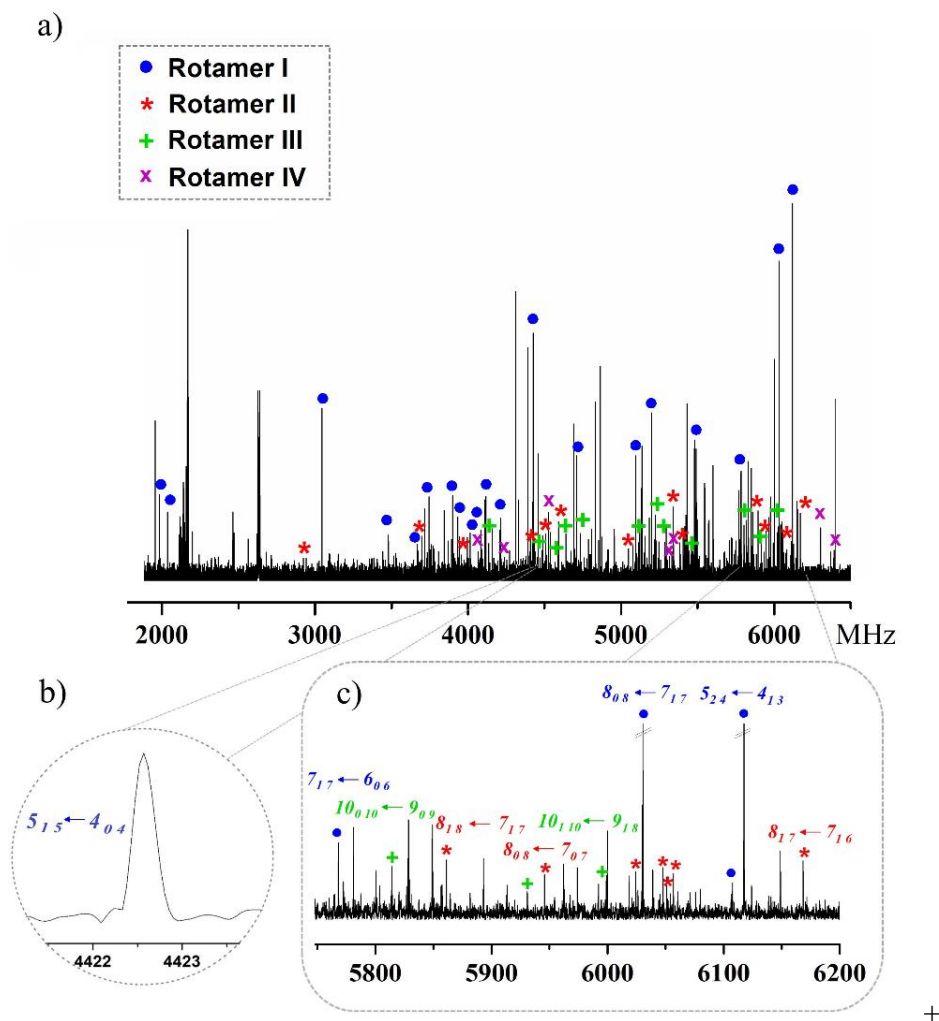
The structures of the six lowest-energy conformers are shown in Fig. 2 and their spectroscopic properties are collected in Table 1 (for additional information see Appendix D, Section D.1: Theoretical benchmark). To label these conformers, we follow a similar notation to that employed in previous studies of structurally similar amino acids tyrosine and phenylalanine (Pérez et al. 2011, 2015) together with that of dopamine (Cabezas et al. 2013). A combination of different symbols is used: a capital letter G or A indicate the gauche (folded) or anti (extended) arrangement of the  $-\text{CH}_2\text{COOH}$  chain. It is noticeable that all the theoretically predicted structures bellow an energetic window of  $1000 \text{ cm}^{-1}$  exhibit type II  $\alpha$ -amino acid configuration, attributed to the formation of stabilizing  $\text{O-H}\cdots\text{N}$  interactions. Finally, the Roman number denotes the energy order ( $\Delta E$ ) for the different conformations associated with the rotation of the cis  $-\text{COOH}$  group and the subscripts *a* and *b* specify the different orientations of the catechol  $-\text{OH}$  groups.

Finally, we have employed non-Covalent Interactions (NCI) methods to analyze and

further probe the existence of non-covalent bonds. These computations are recently implemented theoretical strategies to estimate and visualize weak non-covalent bonding from the topological analysis of the electron density ( $\rho$ ) and of its reduced gradient (Chaudret et al. 2014).

## 2.2. Broadband rotational spectrum and conformational analysis

In Figure 3(a), we show the broadband spectrum of L-DOPA in the 1.8-6.4 GHz frequency range. At first glance, the lines corresponding to known photo-fragmentation products were identified and eliminated. We obtained a noticeably poor spectrum due to the chemical instability of L-DOPA mentioned above. Following the theoretical predictions of Table 1, the six likely conformers (see Fig. 2) are expected to be asymmetric rotors close to the prolate limit. Moreover, all of them present moderate to significant  $\mu_b$  electric dipole-moment components. Thus, we can forecast *b*-type rotational patterns typified by sets of *R*-branch  $(J+1)_{1,J+1} \leftarrow J_{0,J}$  transitions (spaced by  $\sim 2C$ ) (Gillies 1982). In addition, some of the rotational transitions appeared broadened [see Figure 3 (b)] or even split into several close hyperfine components. This effect is characteristic of  $^{14}\text{N}$  bearing molecules and will be discussed below. Initially, the polarization power was set to polarize the transitions associated with moderate values of  $\mu_b$  optimally. Hence, recognizing an intense progression attributable to a first rotamer, labeled as I, was straightforward. The initial assignment was extended to other *b*-type *R*-branch and *Q*-branch transitions (up to  $J = 14$  and  $k_a = 4$ ). Notice, we also predicted *a*- and *c*-type lines but failed in observing them. After an iterative process of fittings and prediction (Pickett 1991), 58 frequency lines were measured and submitted to an analysis employing Watson's asymmetric top semirigid-rotor Hamiltonian in *A*-reduction and *F*-representation (Watson 1977). The resulting experimental rotational constants are collected in the first column of Table 1.



**Figure 3.** (a) Broadband spectrum of L-DOPA from 1.0 to 6.4 GHz; b) Zoomed view showing the non-resolved  $^{14}\text{N}$  hyperfine structure for the  $5_{1,5} \leftarrow 4_{0,4}$  transition of rotamer **I** (**G1a** L-DOPA); c) Section of the spectrum showing characteristic *a*- and *b*-type *R*-branch transitions for rotamers **I-III** for comparison.

After excluding these lines from the spectrum, a weaker progression of  $^aR$ -branch transitions separated by approximately  $B + C$  was discovered and assigned to another rotamer, labeled as **II** [see Figure 3(c)]. The assignments were confirmed by predicting and measuring additional *a*- and *b*-type rotational lines for this rotameric species. The second column of Table 1 shows the resulting spectroscopic parameters. Once the rotational lines of rotamers **I** and **II** were removed from the spectrum, weaker lines following new  $^aR$ -branch patterns were found. An iterative procedure of fittings and predictions made it possible to identify new *a*-type *R*-branch transitions belonging to a third rotamer labeled as **III**. Notice that no *b*- or *c*-type transitions were observed. The spectroscopic constants resulting from a rigid-rotor analysis are collected in the third column of Table 1. Finally, following the same approach, we managed to identify the

hindmost rotamer (**IV**), exhibiting both *a*- and *b*-type rotational transitions of remarkably low intensity. The fourth column of Table 1 lists the resulting spectroscopic parameters for this rotameric species. All measured lines for all four rotamers are included in Tables D2.1-D2.4 of Appendix D (Section D.2).

**Table 1.** Experimental and predicted spectroscopic parameters for the L-DOPA conformers.

Parameter	Experimental				Theory <sup>[a]</sup>					
	Rotamer I	Rotamer II	Rotamer III	Rotamer IV	GIa	GIIa	GIIb	GIIb	AIb	AIa
<i>A</i> /MHz <sup>[b]</sup>	1192.74716 (58) <sup>[b]</sup>	1324.1739 (68)	1726.19 (17)	1750.3440 (23)	1195	1323	1203	1343	1743	1750
<i>B</i> /MHz	437.50936 (49)	397.0161 (11)	307.72855 (43)	305.95858 (11)	433	395	428	389	306	305
<i>C</i> /MHz	376.05121 (59)	357.8806 (10)	280.09432 (42)	280.91514 (53)	373	355	369	351	279	280
$\Delta_J$ /kHz <sup>[c]</sup>	-0.0332 (21)	... <sup>[j]</sup>	...	...	...	...	...	...	...	...
$ \mu_a /D$ <sup>[d]</sup>	Not observed <sup>[i]</sup>	Observed	Observed	Observed	0.3	2.6	2.6	2.1	3.6	5.7
$ \mu_b /D$	Observed	Observed	Not observed	Observed	3.0	2.1	6.4	6.5	0.2	4.1
$ \mu_c /D$	Not observed	Not observed	Not observed	Not observed	0.3	0.5	1.3	0.0	2.1	1.9
<i>N</i> <sup>[e]</sup>	58	15	17	17	...	...	...	...	...	...
$\sigma_{\text{rms}}$ /kHz <sup>[f]</sup>	25.7	20.6	20.9	33.7	...	...	...	...	...	...
$\Delta E$ / cm <sup>-1</sup> <sup>[g]</sup>	...	...	...	...	0	276	495	582	508	537
$\Delta G$ / cm <sup>-1</sup>	...	...	...	...	0	237	442	517	371	381

[a] Calculated at the B3LYP/aug-cc-pVTZ level of the theory with Grimme D3 dispersion. [b] *A*, *B*, and *C* are the rotational constants; [c]  $\Delta_J$  and  $\Delta_{JK}$  are the quartic centrifugal distortion constants (*A*-reduction); [d]  $|\mu_a|$ ,  $|\mu_b|$  and  $|\mu_c|$  are the absolute values of the electric dipole moment components along the inertial axis *a*, *b*, *c*; [e] *N* represents the number of distinct frequency lines in the fit; [f]  $\sigma_{\text{rms}}$  is the root mean square deviation of the fit; [g]  $\Delta E$  and  $\Delta G$  represent the relative energy and Gibbs free energy (*T* = 298 K), respectively, with respect to the global minimum; [h] The numbers in parentheses are 1 $\sigma$  uncertainties in units of the last decimal digit; [i] Experimental observation of a given type of transition; [j] The values of  $\Delta_J$  are not included in the rigid rotor analysis.

We achieved the conformational identification of the observed rotamers by comparing the experimental spectroscopic properties with those predicted theoretically for the candidate structures shown in Fig. 2. The primary evidence comes from the comparison between the experimental and predicted values of the rotational constants collected in Table 1. The experimental values for rotamer **I** are at first in good agreement with those calculated for conformers **GIa** and **GIIb**, respectively. Nevertheless, the differences in the values are not significant enough to discriminate between rotamers that differ in the orientation of the catechol OHs (i.e. **GIa** and **GIIb**), which induces an almost negligible modification of the rotational constants values. However, it causes a drastic

change in the values of the electric dipole moment and, consequently, in the selection rules. For the **G1a** conformer, the electric dipole-moment coincides with the *b* principal axis and, consequently, only *b*-type spectrum is expected for this conformer. On the other hand, *a*-type rotational transitions are also expected for **G1b** conformer due to its high  $\mu_a$  dipole moment component. However, rotamer **I** do not show any *a*-type rotational transition, so it can be unambiguously assigned to conformer **G1a**.

Following the same arguments, rotamer **II** can be ascribed to conformers **G11a** or **G11b**. Both conformers exhibit *a*- and *b*-type spectra, but the-line's intensities correlate exclusively with the predicted dipole moment values of conformer **G11a**. Finally, the rotational constants of rotamers **III** and **IV** match nicely with those predicted for conformers **A1b** and **A1a**, respectively. Moreover, the absence of *b*-type lines allowed us to conclude straightforwardly the presence of conformer **A1b** in our supersonic expansion. Scaling factors ranging from 0.9897 to 1.0097 bring the predicted B3LYP-D3 rotational constants values to the experimental ones, which reinforce the global consistency of the conformational assignment.

According to the energy difference between the predicted conformers (see Table 1), **G1b** and **G11b** forms could also be expected in the jet. However, we experimentally observed conformer **A1b** and **A1a**, which are slightly higher in energy. This behavior can be easily rationalized given that these conformers exhibit an anti (extended) configuration of the amino acid chain, which implies a significant entropic contribution (i.e.,  $\Delta G(A1b) = 371 \text{ cm}^{-1}$ ) that further stabilizes these structures over the other high energy forms. Nevertheless, once we discarded the rotational signatures of conformers **G1a**, **G11a**, **A1b** and **A1a**, we performed thorough searches for **G1b** and **G11b** around the predicted transitions using the above scaling. Although we did some tentative assignments with a few observed lines, which most likely belong to these rotamers, we had not collected enough lines to achieve a proper fit. It must be noted that these weak rotational transitions may also appear split by the effect of the  $^{14}\text{N}$  nuclear quadrupole coupling, challenging their analysis

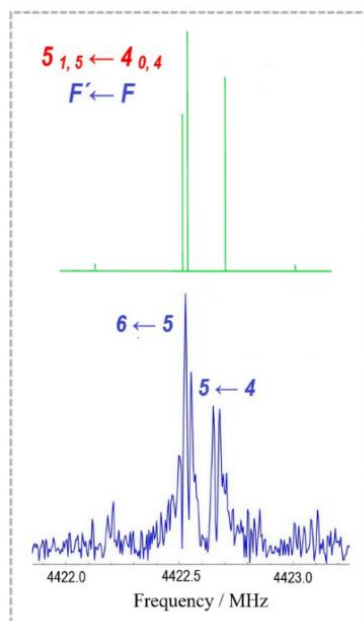
The four structures characterized in this work and two additional tentative observations show a significantly rich conformational panorama for L-DOPA, in accordance with the observations for other catecholamine neurotransmitters, such as dopamine (Cabezas et al. 2013), but in sharp contrast with the previous work of Mitsuda et al. 2010 Based on the observation of a single conformation by electronic spectroscopy,

these authors asserted the existence of a conformational reduction for L-DOPA due to the presence of a catechol ring. Nevertheless, given the experiment's difficulties, the control of the laser desorption process and the abovementioned chemical instability of the molecule could be at the root of these results. The exceptional agreement between the theoretical and experimental parameters shown in our study rule out any sort of conformational restriction.

### 2.3. $^{14}\text{N}$ Hyperfine quadrupole hyperfine structure: the $\text{N-H}\cdots\pi$ interaction

All of the observed conformers exhibit type II amino acid configurations characterized by the formation of stabilizing  $\text{O-H}\cdots\text{N}$  interactions, which are also shown for several  $\alpha$ -amino acids (Blanco et al. 2007; Sanz et al. 2010; Cabezas et al. 2012) Moreover, a common feature of aromatic amino acids (Pérez et al. 2001, 2015, Sanz et al. 2014) along with most catechol amines (Cabezas et al. 2013), is the so-called  $\text{N-H}\cdots\pi$  interactions in which an  $sp^2$ -hybridized covalent bond can act as a donor group and the  $\pi$ -electrons can function as an acceptor group. This interaction depends on the orientation of the N-bearing group and is expected to play a critical role in the conformational preferences of this biomolecule; it further stabilizes its structure by allowing the formation of reinforced intramolecular networks.

We went a step forward in the characterization of L-DOPA and studied the role of the  $\text{N-H}\cdots\pi$  interaction. In this context, the use of the  $^{14}\text{N}$  nuclear quadrupole coupling interactions is an elegant way to experimentally establish the orientation of the amino group and, therefore, to help identify the intramolecular interactions in which this group is involved. The values of the nuclear quadrupole coupling constants are strongly correlated with the chemical environment of the  $^{14}\text{N}$  nucleus. Thus, in a quest to evidence the intramolecular interactions and determine the disposition of the  $-\text{NH}_2$  group relative to the rest of the molecule, we then proceeded to interpret the hyperfine structure arising from the rotational transitions of L-DOPA.



**Figure 4.** Resolved hyperfine pattern for the  $5_{1,5} \leftarrow 4_{0,4}$  transition of **GIa** L-DOPA (LA-MB-FTMW experiment) along with the B3LYP/aug-cc-pVTZ simulations (in green). Each transition appears as Doppler doublets due to the sub-Doppler resolution of the spectrometer. The intensity is given in arbitrary units.

This task emerged as a more challenging endeavor since the resolution of the LA-CP-FTMW spectrometer was insufficient to resolve the hyperfine structure [see Figure 3 (b)]. Consequently, we decided to benefit from the high resolution of our narrowband LA-MB-FTMW spectrometer (León et al. 2021). We started the analysis by re-measuring the previously non-resolved  $5_{1,5} \leftarrow 4_{0,4}$  transition [see Figure 3(b)]. Using the cavity-based technique, we managed to fully resolve its  $^{14}\text{N}$  hyperfine pattern, as shown in Figure 4. This complex structure is derived from the interaction of the electric quadrupole moment of the  $^{14}\text{N}$  ( $I = 1$ ) nucleus with the electric field gradient originated at the location of the nitrogen by the remaining electronic and nuclear charges of the molecule (Gordy & Cook 1970). This interaction leads to a coupling between the  $^{14}\text{N}$  nuclear spin and the overall angular momentum, originating a splitting in the rotational energy levels that decreases the overall intensity of each rotational transition. We then extended the analysis to five intense  ${}^bR$ -branch transitions of rotamer **I**. We collected these rotational lines and fitted them to a Hamiltonian in the following form:  $H = H_{SR}^{(A)} + H_Q$ , where  $H_{SR}^{(A)}$  describes the rotational parameters using a semirigid rotor approximation and  $H_Q$  considers the quadrupole coupling interactions (Foley 1947; Robinson 1953). The energy levels involved in each transition are labelled according to the quantum numbers  $J$ ,  $K_{-1}$ ,  $K_{+1}$ , and  $F$ , following the coupled basis set  $F = J + I$ . The derived quadrupole coupling



constants are listed in Table 2. Nevertheless, given the experimental difficulties obtaining the rotational spectrum, only transitions belonging to the most stable form, showing the most intense lines, were affordable during the narrowband experiment. A list of the measured hyperfine components of **GIa** is presented in Table D2.5 of Appendix D (Section D2).

**Table 2.** Experimental and predicted  $^{14}\text{N}$  nuclear quadrupole constants for **GIa** L-DOPA.

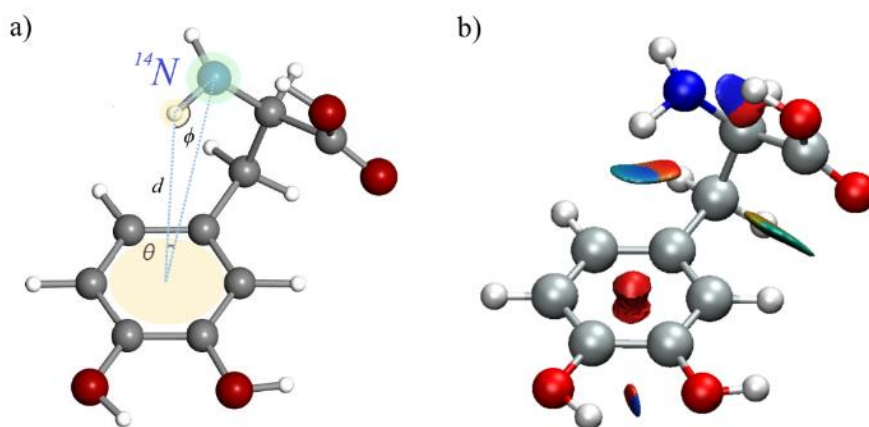
<b>GIa</b> L-DOPA		
Parameter	Experimental	Theoretical [a]
$\chi_{aa}$ /MHz <sup>[b]</sup>	0.675 (34) <sup>c</sup>	0.661
$\chi_{bb}$ /MHz	0.184 (32)	0.379
$\chi_{cc}$ /MHz	-0.858 (32)	-1.040

[a] Calculated at the B3LYP/aug-cc-pVTZ level of the theory with Grimme D3 dispersion factor. [b]  $\chi_{aa}$ ,  $\chi_{bb}$  and  $\chi_{cc}$  are the diagonal elements of the  $^{14}\text{N}$  nuclear quadrupole coupling tensor; [c] The numbers in parentheses are  $1\sigma$  uncertainties in units of the last decimal digit.

As shown in Table 2, the same conclusion mentioned above for the rotational constants can be reached by comparing the theoretical and experimental values of the  $^{14}\text{N}$  quadrupole coupling constants. The experimental values of rotamer **I** are in good agreement with those predicted for the **GIa** conformer. The small discrepancy found between the experimental and predicted values of quadrupole coupling constants (Table 2) is due to a slight variation in the actual orientation of the amino group compared to that predicted via *DFT* methods. All in all, this exceptional tool allowed us to establish the orientation of the  $-\text{NH}_2$  group with respect to the molecular framework (see Fig. 5(a))

Concurrent to the experimental work, we executed non-Covalent Interactions (NCI) computations based on the B3LYP/aug-cc-pVTZ optimized structures to corroborate and further probe the nature of the intramolecular bonding. Thus, we can visualize weak non-covalent interactions from the topological analysis of the electron density ( $\rho$ ) and its reduced gradient. The conformational preferences showed by the four conformers, studied in detail for the **GIa** form [see Fig. 5(b)], can be therefore rationalized in terms of different intramolecular forces: the  $\text{OH}\cdots\text{OH}$  hydrogen bond of the hydroxyl groups

in the catechol ring, the strong O-H $\cdots$ NH<sub>2</sub> hydrogen bond of the side chain, and the NH $\cdots$  $\pi$  interaction between the amino group and the aromatic ring, which is one of the principal structural motifs stabilizing the structure.



**Figure 5.** (a) Structure of **G1a** L-DOPA computed at the B3LYP/aug-cc-pVTZ level showing the <sup>14</sup>N nucleus environment along with the geometry of the N–H $\cdots$  $\pi$  intramolecular interaction ( $d=3.204$  Å). (b) Visual molecular dynamics (VMD) graphics of the **G1a** L-DOPA NCI analysis (Humphrey et al. 1996). Different NCI type I, II and III isosurfaces were found. NCI type I isosurfaces corresponding to strong stabilizing interactions, attributable to intramolecular hydrogen bonds and the relevant N–H $\cdots$  $\pi$  interaction are labeled in blue; NCI type II isosurfaces account for strong destabilizing interactions, corresponding to ring closures and steric crowding, which are displayed in red; NCI type III gathers the delocalized weak Van der Waals interactions and are labeled in green (Guillet et al. 2012). NCI color code: BGR.

### 3. CONCLUSIONS

The present study constitutes the first complete experimental investigation of the gas-phase structures of the elusive amino acid L-DOPA. Thus, the conformational behavior of this neurotransmitter precursor and supramolecular synthon has been deciphered at last in the isolation conditions of the gas phase based on a comprehensive high-resolution rotational study. A remarkably rich conformational space has been found, and four distinct structures of L-DOPA have been unequivocally characterized, along with two additional tentative assignments. Our investigation highlights that no conformational restriction is shown for this catecholamine and conclusively refuses the previously reported catechol ring-induced conformational restriction (Mitsuda et al 2010).

Also, in consonance with previous structural studies of closely related  $\alpha$ -amino acids, all of the observed structures show type II amino acid configurations (stabilizing O-H $\cdots$ N interactions), which allow the formation of multiple intramolecular bonds that significantly stabilize each structure. Moreover, a thorough analysis of the  $^{14}\text{N}$  nuclear quadrupole hyperfine structure enabled us to unveil the orientation of the N-bearing functional group and confirm the presence of stabilizing N-H $\cdots\pi$  interactions. Finally, the intramolecular forces that drive the conformational panorama of L-DOPA have been thoroughly described using a combination of NCI and theoretical computations.

The current investigation provides fundamental knowledge on the shape of the isolated molecule and should help to elucidate and design novel supramolecular materials that involve L-DOPA as a multifunctional building block. Also, this study should shed light on the characterization of systems with increasing complexity levels, i.e., water clusters. Future studies will be directed towards adding individual water molecules to mimic the transition to the biological environment (i.e., an aqueous medium). It shall ultimately aid to clarify the structural dependence in the decarboxylation process that leads to dopamine in our organism

#### 4. EXPERIMENTAL SECTION

In the experimental procedure, we mixed finely powdered L-DOPA with a small amount of a commercial copolymeric binder and pressed it into cylindrical rods, which were ablated using the third harmonic (355 nm, 4 mJ per pulse) of an Nd-YAG picosecond laser. The vaporized products were seeded in neon at a backing pressure of 10 bar, adiabatically expanded into the spectrometer's vacuum chamber, and then probed by broadband chirped pulse Fourier transform microwave spectroscopy (LA-CP-FTMW) (Bermúdez et al. 2016). A high-power excitation pulse of 300 W was used to polarize the molecules at frequencies from 1 to 6.4 GHz. Up to 80000 individual free induction decays (4 FIDs on each valve cycle at a 2 Hz repetition rate) were averaged in the time domain, and Fourier transformed to obtain the broadband frequency domain spectrum. A second experiment was then carried out from 7.4 to 12.6 GHz to extend the frequency range and allow to collect more rotational lines. In this case 100k FIDs were recorded.

In a second step, we employed a laser ablation molecular beam Fourier transform

microwave (LA-MB-FTMW) spectrometer, which operates between the frequency range from 2-8 GHz, described elsewhere (León et al. 2019), to resolve the  $^{14}\text{N}$  hyperfine structure of L-DOPA. The laser-ablated products were seeded in Ne (stagnation pressure 10 bar) and expanded adiabatically to form a supersonic jet into a Fabry-Pérot resonator. A short microwave radiation pulse (0.3  $\mu\text{s}$ ) was then applied to polarize the vaporized molecules macroscopically. The microwave transient FID was again registered in the time domain, and Fourier transformed into the frequency domain.

## 5. BIBLIOGRAPHY

- Alonso, E. R., León, I., Alonso, J. L., *Intra- and Intermolecular Interactions between Non-covalently Bonded Species*. Elsevier, **2021**, 93-141.
- Alonso, J. L.; Sanz, M. E.; López, J. C.; Cortijo, V. *J. Am. Chem. Soc.* **2009**, 131, 4320–4326.
- Alonso, J. L. López, J. C. *In Gas-Phase IR Spectroscopy and Structure of Biological Molecules*, Springer International Publishing, Cham, **2015**, p. 335.
- Bermúdez, C., Peña, I., Mata, S., Alonso, J. L. *Chem. Eur. J.* **2016**, 22, 16829 and references therein.
- Blanco, S., Sanz, M.E., López, J.C., Alonso, J.L., *Proc. Natl. Acad. Sci. U.S.A.*, **2007**, 104, 20183.
- Bromberg-Martin, E. S., Matsumoto, M. and Hikosaka, O., *Neuron*. **2010**; 68(5): 815–834.
- Cabezas, C.; Varela, M.; Peña, I.; Lopez, J. C.; Alonso, J. L. *Phys. Chem. Chem. Phys.* **2012**, 14, 13618–13623, and references therein
- Cabezas, C.; Peña, I. López, J. C.; Alonso, J.L. *J. Phys. Chem. Lett.* **2013**, 4, 486-490.
- Calabresi, P., Picconi, B., Tozzi, A. and Filippo, M. D. *Trends in Neurosciences*, **2007**, 30, 5, pp. 211–219, and references therein.
- Caminati, W. *Phys. Chem. Chem. Phys.*, **2004**, 6, 2806-2809.
- Çarcabal, P.; Snoek, L. C.; Van Mourik, T. *Mol. Phys.* **2005**, 103, 1633–1639.
- Chaudret, R., de Courcy, B., Contreras-García, J., Gloaguen, E., Zehnacker-Rentien, A. Mons, M., and Piquemal, J.-P. *Phys. Chem. Chem. Phys.* **2014**, 16, 9876.
- Clarke, S. J.; Hollmann, C. A.; Zhang, Z. J.; Suffern, D.; Bradforth, S. E.; Dimitrijevic, N. M.; Minarik, W. G.; Nadeau, J. L. *Nat. Mater.* **2006**, 5, 409.
- Dunning, T.H., *J. Chem. Phys.* **1989**, 90, 1007.
- Fernandez-Pastor, I., Luque-Muñoz, A., Rivas, F., et al. *Phytochemical Analysis*, **2019**, 30:89-94.
- Foley, H. M. **1947**, *Phys. Rev.*, 71, 747.
- Gaussian 16, Revision A.03, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi,

- R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian, Inc., Wallingford CT, **2016**.
- Gillet, R. Chaudret, J. Contreras-Garcia, W. T. Yang, B. Silvi and J. P. Piquemal, *J. Chem. Theory Comput.*, **2012**, 8, 3993–3997.
- Gillies, C.W., *Applied Spectroscopy Reviews*, **1982**, 18(1), 1.58.
- Giui, D., Ravarino, P., Tomasini, C., L-Dopa in small peptides: an amazing functionality to form supramolecular materials, *Org. Biomol. Chem.*, **2021**, 19, 4622.
- Gordy, W., Cook, R. L. *Microwave Molecular spectra*, Interscience Pub, **1970**, New York.
- Grimme, S., Antony, J., Ehrlich, S. and Krieg, S., *J. Chem. Phys.* **2010**, 132, 154104.
- Guvendiren, M.; Brass, D. A.; Messersmith, P. B.; Shull, K. R. *J. Adhes.* **2009**, 85, 631.
- Humphrey, W., Dalke, A., Schulten, K. *J. Molec. Graphics*, **1996**, 14, 33-38.
- Ishiuchi S., Mitsuda, H., Asakawa, T. Mizayaki, M., Fujji, M. *PCCP*, **2011**, 13, 7812-7820.
- Kamal Uddin Zaidi, Ayesha S. Ali, Sharique A. Ali, and Ishrat Naaz, *Biochemistry Research International*, **2014**, 16.
- Kolesníková, L., León, I., Alonso, E.R., Mata, S., Alonso, J.L. *Angew. Chem. Ed. Int.* **2021**, accepted DOI: [10.1002/anie.202110325](https://doi.org/10.1002/anie.202110325)
- Lee, C. T. Yang, W.T. Parr, R.G. *Phys. Rev.* **1988**, 37, 785–789.
- Lee, H.; Scherer, N. F.; Messersmith, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, 103, 12999.
- León, I.; Alonso, E. R.; Cabezas, C.; Mata, S.; Alonso, J. L. *Commun. Chem.* **2019**, 2, 1-6 and references therein.
- Li, S. C.; Wang, J. G.; Jacobson, P.; Gong, X. Q.; Selloni, A.; Diebold, U. *J. Am. Chem. Soc.* **2009**, 131, 980.
- Liapakis, G.; Ballesteros, J. A.; Papachristou, S.; Chan, W. C.; Chen, X.; Javitch, J. A. *J. Biol. Chem.* **2000**, 275, 37779–37788.
- Lin, Q.; Gourdon, D.; Sun, C. J.; Holten-Andersen, N.; Anderson, T. H.; Waite, J. H.; Israelachvili, J. N.. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, 104, 3782.
- López, J. C.; Cortijo, V.; Blanco, S.; Alonso, J. L.. *Phys. Chem. Chem. Phys.* **2007**, 9, 4521– 4527.
- Meiser, J., Weindl, D. & Hiller, K. Complexity of dopamine metabolism. *Cell Commun Signal* **2013**, 11, 34.
- Mitsuda, H., Miyazaki, M., Nielsen, I. B., Carcabal, P., Dedonder, C., Jouvet, C., Ishiuchi, S., and Fujii, M., *J. Phys. Chem. Lett.* **2010**, 1, 1130–1133.
- Namhey Lee, Daniel R. Hummer Dimitri A. Sverjensky Tijana Rajh, Robert M. Hazen, Andrew Steele, and George D. Cody, *Langmuir*, **2012**, 28, 50, 17322–17330 and references therein.
- Pérez, C., Mata, S., Cabezas, C., López, J.C., Alonso, J.L. *J. Phys. Chem. A.* **2015**, 3731-3735.
- Pérez, C., Mata, S., Blanco, S., López, J.C., Alonso, J.L. *J. Phys. Chem. A.* **2011**, 115, 34, 9653-9657.
- Pickett, H. M., *Journal of Molecular Spectroscopy*, **1991**, 148, 371.
- Rajh, T.; Saponjic, Z.; Liu, J. Q.; Dimitrijevic, N. M.; Scherer, N. F.; Vega-Arroyo, M.; Zapol, P.; Curtiss, L. A.; Thurnauer, M. C. *Nano Lett.* **2004**, 4, 1017.
- Robinson, G. W., & Cornwell, C. D. **1953**, *J. Chem. Phys.*, 21, 1436.
- Sanz, M. E., López, J.C., Alonso, J.L., *PCCP*, **2010**, 12, 3573-3578.

Sanz, M. E., Cabezas, C., Mata, S., Alonso, J.L., *JCP*, **2014**, 140, 204308.

Shulman J M, De Jager P L and Feany M B, *Annu Rev Pathol*, **2011**, 193.

Simuni T and Hurtig H, Levodopa, *A Pharmacologic Miracle Four Decades Later, in Parkinson's disease: Diagnosis and Clinical Management*, edited by S A Factor and W J Weiner, Demos Medical Publishing, New York, **2008**.

Snoek, L. C.; Van Mourik, T.; Simons, J. P. *Mol. Phys.* **2003**, 101, 1239–1248.

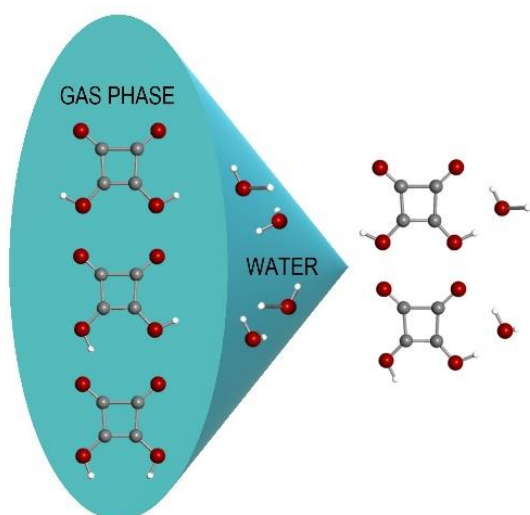
Watson, J. K. G. *In Vibrational Spectra and Structure*; Durig, J. R., Ed., Elsevier: Amsterdam, The Netherlands, **1977**; Vol. 6, pp 1–89.

Weinhold, M., Soubatch, S., Temirov, R. Rohlfig, M. et al. *J. Phys. Chem. B*, **2006**, 110, 23756-23769.

Wieland, K.; Zuurmond, H. M.; Krasel, C.; Ijzerman, A. P.; Lohse, M. J.. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, 93, 9276–9281.

## CHAPTER VIIb. THE SHAPE OF THE ARCHETYPICAL OXOCARBON SQUARIC ACID AND ITS WATER CLUSTERS

Adapted from Sanz-Novo et al. 2019, *Chem. Eur. J.* 25, 10748-10755



Herein, a full structural description is presented for the archetypical supramolecular synthone squaric acid (3,4-dihydroxy-3-cyclobutene-1,2-dione), placed in the gas phase by laser ablation and characterized by chirped pulse Fourier transform microwave technique. Free from natural environmental disturbances, two different *anti-anti* and *syn-anti* planar forms and the corresponding water clusters have been revealed in a supersonic expansion. The substitution structure of the most stable *anti-anti*

conformer has also been extracted from the analysis of the rotational spectra of the  $^{13}\text{C}$  and  $^{18}\text{O}$  isotopic species in their natural abundance. The interplay between inter- and intramolecular interactions involving hydroxy and carbonyl groups has been analyzed by QTAIM (quantum theory of atoms in molecules) methods for squaric acid and its water clusters to understand their chemical behavior and further rationalize their role in the stabilization of these molecular systems.

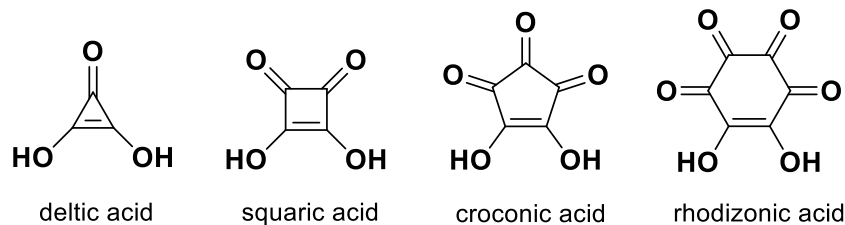
## 1. INTRODUCTION

Cyclic oxocarbons and their derivatives have attracted the attention of many theoretical and experimental researchers in the last half-century. Among them, cyclic oxocarbon acids ( $C_nO_nH_2$ , shown in Figure 1) and their conjugated bases, the oxocarbon dianions, have been extensively used to design metal-organic frameworks (MOFs) and supramolecular ensembles (Das et al. 2005; Wang et al. 2007). One of the most iconic members of the cyclic oxocarbon acids family, squaric acid (3,4-dihydroxy-3-cyclobutene-1,2-dione, SA from now), was first synthesized in the 1950s (Cohen et al. 1959; Park et al. 1962). Its structure is quite remarkable owing to its four-membered cyclic ring and its capability to establish four intermolecular hydrogen bonds. This high affinity for hydrogen bonding, in combination with its inherent high structural rigidity, allows squaric acid and its derivatives to reach a variety of applications in organic chemistry (Badiola et al. 2014), biomedical chemistry (Liu et al. 2013; Hill et al. 2015), and material science (Skopinska-Wisniewska et al. 2016). Hence, squaric acid is used as a rigid ligand to ensure the directionality of the transformation of meta-organic macrocycles (Li et al. 2018). In the field of asymmetric organocatalysis, Chen et al. (2014) designed novel bifunctional squaramide-catalyzed cascade reactions. Moreover, Pedrosa et al. (2017) reported the synthesis of enantiopure chromenes and pirochromanes utilizing supported and unsupported chiral squaramides as organocatalysts. Later, Liu et al. (2013) proposed 3,4-diaryl squaric acid analogs as a new class of potent anticancer agents (Liu et al. 2013) noticing that advances in the synthesis of squaric acid derivatives have gradually encouraged biomedical chemists to incorporate this small molecule into new medicinal chemistry programs. Finally, Skopinska-Wisniewska et al. (2016) exploited the acidity and the cyclic symmetrical configuration of squaric acid acting as a cross-linker agent in collagen/elastin hydrogels.

Despite the importance of its applications, no detailed structural information of the isolated molecule has been yet reported. Squaric acid is a solid that decomposes at 245°C (Lee et al. 2012); it cannot be transferred intact into the vapor phase by using conventional heating methods. It has been investigated in the condensed phases, by IR (West & Niu 1963), X-ray diffraction, and complementary neutron diffraction spectroscopies (Demmingsen et al. 1995), showing a two-dimensional sheet structure where each  $C_4O_4H_2$  group is bonded to four different units. A strong hydrogen-bond network is, therefore, created in the crystal. Moreover, Nowak et al. (2013)



trapped squaric acid in solid argon by matrix isolation techniques inducing conformational changes by using a near-IR laser.



**Figure 1.** Cyclic oxocarbon acids of formula  $(C_nO_nH_2)$ .

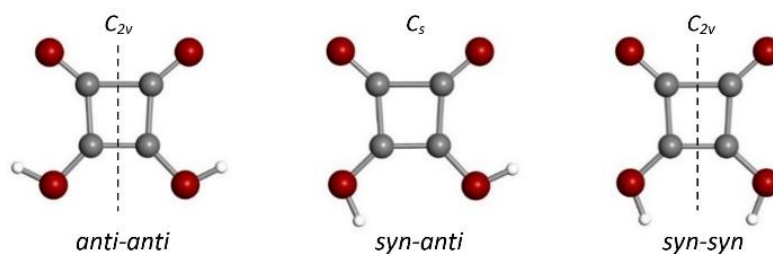
Squaric acid stands as a challenging problem for high-resolution rotational studies, remaining unexplored in the isolation conditions of the gas phase. Fourier transform (FT) microwave spectroscopy is extremely sensitive to molecular geometry and has been proved to be a crucial tool in the conformational analysis (Schermann, 2008). Coupled with laser ablation techniques, it enables the transfer of organic and biological molecules from the solid state to the gas phase and allows them to be probed in the isolated environment of a supersonic expansion (Alonso et al. 2014). By using this powerful strategy, most of the amino acids and several relevant biological molecules (Alonso et al. 2015), as well as species of astrophysical interest, such as glycineamide (Alonso et al. 2018), have been recently revealed. Also, new gas-phase experiments exploring not only the conformational landscape of squaric acid but its microsolvated complexes should represent an essential source of reference data to understand its inherent inter- and intramolecular interactions. The chemical interpretation of those might pave the way to the elucidation of the hydration phenomena in liquids at the molecular level.

## 2. RESULTS AND DISCUSSION

### 2.1. Conformational panorama

As seen in Figure 2, squaric acid possesses a diketocyclobutene ring with two additional hydroxy moieties. The position of the hydrogen atom of the hydroxy groups is the unique contribution to the conformational panorama of this molecule. Thus, SA can achieve three different configurations: *anti-anti* ( $C_{2v}$  symmetry), *syn-anti* ( $C_s$  symmetry), and *syn-syn* SA ( $C_{2v}$  symmetry).

The conformational landscape of squaric acid was explored with the help of theoretical calculations to guide the spectral search. As anticipated, the candidates mentioned above were found in our conformational search and were then optimized at different levels of theory (see Tables D3.1-3.2 in Section D.3 of Appendix D). Each conformer was confirmed to be a local minimum in the potential energy surface by checking that its Hessian matrix did not have any imaginary eigenvalues. The interconversion barrier heights from the *syn-anti* to the *anti-anti* conformer [25.0 kJ mol<sup>-1</sup>, calculated at the B3LYP/6-311++G(d,p) level] were estimated to be similar to that reported elsewhere (Miller et al. 2005) for the *cis-trans* conversion barrier of glycine (closer to 24 kJ mol<sup>-1</sup>), which clearly points to the possibility of having *syn-anti* SA in the supersonic jet. Two individual relaxed potential-energy scans (see Figure D1 in Section D.3 of Appendix D.) were computed, choosing the CCOH (*syn-OH*) and CCOH (*anti-OH*) torsional angles, as the driving coordinate. The conformational switching, by rotation of one of the -OH groups, seems to be hindered owing to the plausible formation of long-range intramolecular hydrogen bonds (stabilizing both *anti-anti* and *syn-anti* conformers).

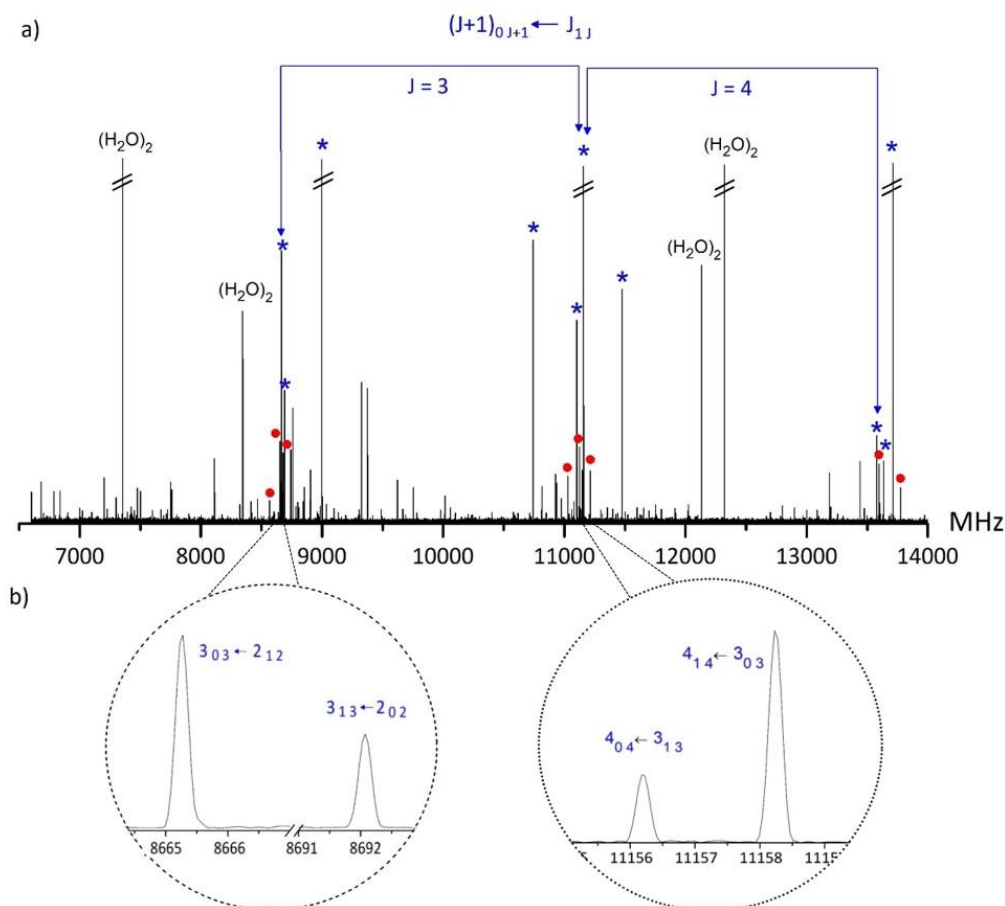


**Figure 2.** Structures of *anti-anti* SA (left), *syn-anti* SA (middle) and *syn-syn* SA (right) conformers. Configurations are labelled following the nomenclature proposed by Semmingsen et al. (1995): considering the molecule as consisting of two crosslinked O=C-C=C-O-H chains, the two chains can differ in the position of the hydroxyl proton with respect to the rest of the arrangement.

Besides the relative energy of conformers, molecular conformational properties relevant to the investigation of the rotational spectrum, mainly rotational constants and electric dipole moment components, were predicted to guide the spectral searches. Also, quantum theory of atoms in molecules (QTAIM) and non-covalent interactions (NCI) methods (Bader 2002; Matta & Boyd 2007; Chaudret et al. 2014), recently developed

theoretical strategies to visualize weak non-covalent interactions from the topological analysis of the electron density ( $\rho$ ) and of its reduced gradient, have been performed to analyze the non-covalent bonding. A more detailed description is given in Section D.5 of Appendix D,

## 2.2. Rotational spectrum



**Figure 3.** (a) Broadband LA-CP-FTMW spectrum of squaric acid from 6.5 to 14 GHz. We depict strong *b*-type R-branch  $(J+1)_{0,J+1} \leftarrow J_{1,J}$  and  $(J+1)_{1,J+1} \leftarrow J_{0,J}$  transitions of rotamer **I** (*anti-anti* SA, in blue) as well as some transitions of rotamer **II** (*syn-anti* SA, in red). (b) Zoom-in on the broadband spectrum showing the alternation of the 3:1 statistical weights (see text) for the  $3_{03} \leftarrow 2_{12}$  (*eo-oe*),  $3_{13} \leftarrow 2_{02}$  (*oo-ee*) and the  $4_{04} \leftarrow 3_{13}$  (*ee-oo*),  $4_{14} \leftarrow 3_{03}$  (*oe-oo*) transitions of rotamer **I**. The intensity is given in arbitrary units.

The broadband spectrum of SA from 6.5 to 14 GHz is shown in Figure 3(a). It appeared dominated by intense water dimer  $(\text{H}_2\text{O})_2$  signals and other lines belonging to known

photofragmentation products, which were identified and discarded from the analysis. According to the predictions of Table 1, the three plausible conformers of SA are asymmetric rotors close to the oblate limit. We can, therefore, anticipate rotational spectra patterns characterized by pairs of  ${}^bR$ -branch  $(J+1)_{0,J+1} \leftarrow J_{1,J}$  and  $(J+1)_{1,J+1} \leftarrow J_{0,J}$  transitions spaced at approximately  $2C$ , arising from conformers predicted with significant  $\mu_b$  electric dipole-moment components. Hence, it was straightforward to recognize an intense progression attributable to a first rotamer, labeled as **I** in Figure 3. The initial assignment was extended to other  $b$ -type transitions (with  $J$  ranging from 2 to 5). Notice,  $a$ - and  $c$ -type lines were also predicted but not observed. After an iterative procedure of fitting and prediction, a total of 13 rotational lines were measured. They were least-squares fitted to a rigid rotor Hamiltonian (Pickett 1991).

**Table 1.** Theoretical and experimental spectroscopic parameters for the conformers of Squaric Acid.

	<i>Experiment</i>		<i>Theory</i> <sup>[a]</sup>		
	<i>rotamer I</i>	<i>rotamer II</i>	<i>anti-anti</i>	<i>syn-anti</i>	<i>syn-syn</i>
$A$ <sup>[b]</sup>	2585.8670 (19) <sup>[h]</sup>	2502.2160 (18)	2583	2501	2495
$B$	2382.1261 (19)	2443.5205 (26)	2384	2445	2434
$C$	1239.5813 (15)	1235.98520(99)	1240	1236	1232
$\Delta I$ <sup>[c]</sup>	0.10789 (54)	0.09089 (42)	-0.0074	-0.0073	-0.0073
$ \mu_a $ <sup>[d]</sup>	No	Yes	0.0	2.0	0.0
$ \mu_b $	Yes	Yes	3.5	5.6	7.7
$ \mu_c $	No	No	0.0	0.0	0.0
$N$ <sup>[e]</sup>	13	13	-	-	-
$\sigma$ <sup>[f]</sup>	7.8	12.3	-	-	-
$\Delta E_{TOT}$	-	-	0	769	2836
$\Delta G$ <sup>[g]</sup>			0	762	2752

[a] Theoretical calculations at B3LYP/6-311++G (d,p) level of theory with Grimme Dispersion. [b]  $A$ ,  $B$ , and  $C$  represent the rotational constants (in MHz). [c]  $\Delta I$  is the inertial defect (in  $\text{u}\text{\AA}^2$ ), conversion factor:  $505379.1 \text{ MHz} - \text{u}\text{\AA}^2$ . [d]  $\mu_a$ ,  $\mu_b$ , and  $\mu_c$  are the electric dipole moment components (in D). [e]  $N$  is the number of measured transitions. [f]  $\sigma$  is the root mean square (rms) deviation of the fit (in kHz). [g]  $\Delta E_{TOTAL}$  and  $\Delta G$  are the electronic and Gibbs energies (in  $\text{cm}^{-1}$ ) at 298 K relative to the global minimum, taking into account the zero-point vibrational energy (ZPE) for the electronic energy ( $E_{TOTAL} = E + E_{ZPE}$ ) calculated at the same level. [h] Standard error in parentheses in units of the last digit.

The resulting experimental spectroscopic constants are listed in the first column of Table 1. After excluding these lines from the spectral analysis, another weaker progression of *b*R-branch transitions was discovered and assigned to another rotamer, labeled as **II**. The assignments were confirmed by predicting and measuring new lines (additional rotational lines). Contrary to rotamer **I**, rotamer **II** exhibits a weak *a*-type rotational spectrum. After correspondent fitting, a total of 13 rotational lines in the same frequency region were assigned to rotamer **II**. The resulting rotational constants (*A*, *B*, *C*) from the rigid rotor analysis are also collected in the second column of Table 1. All measured lines for both rotamers are listed in Tables D5.1-D5.2 in Section D.5 of Appendix D.

The first piece of evidence for the identification of the observed rotamers comes from the values of the rotational constants. Both rotamers **I** and **II** share the same skeletal framework. However, owing to the different orientation of the hydroxy groups, the differences in the values of the rotational constants are significant enough to allow conclusive discrimination. The excellent match between the experimental and theoretical values shown in Table 1 enables the identification of rotamers **I** and **II** as *anti-anti* and *syn-anti* SA, respectively. Scaling factors ranging from 0.99998 to 1.00092 bring the values of the predicted rotational constants of Table 1 to those experimentally obtained. Thus, we can infer that its actual geometry should be close to that calculated at the DFT level. Surprisingly, we noticed that B3LYP/6-311++G(d,p) gives the most accurate rotational constants and thus a better structural prediction, even compared with higher level computations.

Further proof for the identification of the rotamers is provided by the observed dipole-moment selection rules. For both *anti-anti* and *syn-syn* conformers with  $C_{2v}$  symmetry, the electric dipole-moment coincides with the  $C_2$ -symmetry axis (*b* principal axis) and, consequently, only the *b*-type spectrum is observed for rotamer **I**, assigned as conformer *anti-anti*. Otherwise, the *syn-anti* conformer, having  $C_s$  symmetry, presents an additional permanent dipole-moment component along the *a* principal axis. Accordingly, rotamer **II**, identified as the *syn-anti* conformer, shows both *a*-type and *b*-type rotational spectrum. Therefore, the dipole-moment selection rules enabled us to confirm the conformational assignment.

In this case, another interesting spectroscopic tool is taken into account: the nuclear statistics. It is well known that the presence of identical nuclei in the molecule can have

significant consequences for the determination of statistical weights of energy levels and the relative intensities of rotational lines (Gordy & Cook 1970). Considering that the  $C_{2v}$  *anti-anti* conformer of SA has one pair of equivalent protons (obeying Fermi--Dirac statistics), the antisymmetric spin functions must be used with rotational functions that are symmetric with respect to a rotation around the  $b$ -symmetry axis ( $K_a, K_c$  equals "even, even" or "odd, odd") and antisymmetric spin functions with symmetric rotational functions ( $K_a, K_c$  equals "even, odd" or "odd, even"). The proper nuclear statistical weight is 1:3, respectively. In fact, this ratio is supported by the alternation of the intensities observed along with the pairs of  $^bR$ -branch  $(J+1)_{0,J+1} \leftarrow J_{1,J}$  and  $(J+1)_{1,J+1} \leftarrow J_{0,J}$  transitions, which are shown in Figure 3(b). No similar effect is observed for rotamer **II** transitions (see Figure D6 in Appendix D, Section D.6.). This enabled us to further corroborate the *syn-anti* conformer as rotamer **II**. In conclusion, by using the above-mentioned spectroscopic tools, an unambiguous conformational characterization of SA was achieved.

One of the singular features of squaric acid is the planar and rigid structure of the diketocyclobutene ring, which has not been experimentally proved yet for the bare molecule, although it has been studied for various squaric acid derivates (Davis et al 1999; Rotger et al 2004). It is particularly fruitful to make use of a combination of the principal inertial moments called "inertial defect",  $\Delta I = I_c - I_a - I_b$ , which is given in Table 1 for both rotamers, to establish the planarity of squaric acid ( $ab$  plane of symmetry). This quantity depends on the square of the out-of-plane  $c$  coordinates of the  $i$ -th nucleus (Gordy & Cook 1970), giving the mass extension out of the  $ab$  plane. Small positive values of the inertial defect ( $\Delta I(\textit{anti-anti})=0.10789 \text{ u}\text{\AA}^2$ ,  $\Delta I(\textit{syn-anti})=0.09089 \text{ u}\text{\AA}^2$ ) are observed for both monomers. This is a direct indicator of the planarity of small molecules, where the positive values are attributed to vibration-rotation interactions, dissected to the harmonic, anharmonic, and Coriolis contributions (Oka et al. 1995). This analysis finally leads us to the conclusion that both equilibrium conformations are indeed planar. It is of interest to see how this planarity allows squaric acid molecules to crystallize into 2D sheets where each SA group is hydrogen bonded to four different units configuring a macromolecular network. Neutron diffraction studies carried out by Semmingsen et al. (1995) showed *syn-anti* SA as the predominant monomer of the crystal structure. However, it is not the most abundant species in the gas phase.

### 2.3. Isotopic species and substitution structure of the *anti-anti* conformer

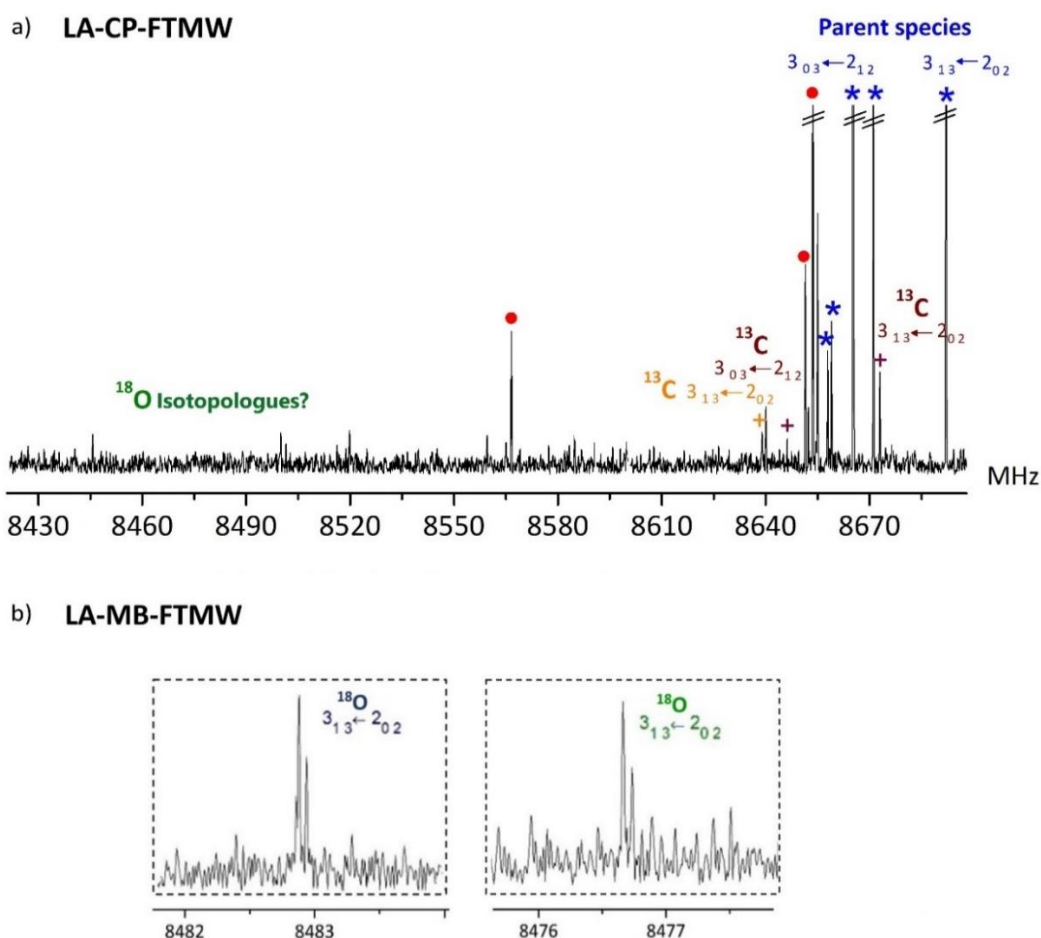
A closer inspection into the low-intensity background of the spectrum did not reveal the existence of spectral signatures attributable to the *syn-syn* conformer of SA. It lies 2800  $\text{cm}^{-1}$  above the *anti-anti* global minimum (see Table 1) and is expected not to be populated enough in our supersonic expansion. This fact is also supported by the low interconversion barrier from *syn-anti* to *syn-syn* SA (9.3  $\text{kJ mol}^{-1}$ ). However, on the low-frequency side of each  $\mu_b$ -type *R*-branch transition of the *anti-anti* conformer, very weak sets of the same *R*-branch progressions could be identified and were indeed ascribed to the  $^{13}\text{C}$  isotopologs in their natural abundance (1.109 %) with doubled intensity owing to the existence of two equivalent carbons [see Figure 4(a)]. The sensitivity attained with our laser ablation broadband chirped pulse Fourier transform microwave spectroscopy (LA-CP-FTMW) technique was not sufficient to observe the  $^{18}\text{O}$  monosubstituted species despite its enhanced abundance (0.410 % due to the  $C_{2v}$  symmetry). We took advantage of the high sensitivity of our narrowband LA-MB (molecular beam)-FTMW spectrometer (Alonso et al 2009; Bermúdez et al 2014) for the observation of  $^{18}\text{O}$  isotopologs [see Figure 4(b)]. A total of 19 rotational transitions were measured (Tables D5.3-D5.6 in Appendix D, Section D.5) for the four isotopic species and used in a rigid rotor analysis to give the experimental spectroscopic constants presented in Table 2. Notice that isotopic substitutions,  $^{12}\text{C}$  ( $^{13}\text{C}$ ) and  $^{16}\text{O}$  ( $^{18}\text{O}$ ), do not alter appreciably the quantity of the inertial defects, which further corroborate the planarity of *anti-anti* SA.

The substitution structure ( $r_s$ ) of the heavy atom skeleton of the *anti-anti* form of SA was derived from the set of observed rotational constants in Table 2. By using the Kraitchman equations (Kraitchman 1953), the cartesian coordinates of the substituted atoms are determined in the principal axis system. The  $r_s$  substitution coordinates together with the derived bond lengths and angles are given in Table 3. A satisfactory agreement was achieved between the  $r_s$  structure and the predicted equilibrium ( $r_e$ ) structure calculated at the B3LYP/6-311++G(d,p) level.

**Table 2.** Experimental spectroscopic parameters for the observed isotopologues of *anti-anti* SA.

Parameters	<i>anti-anti</i>	<i>anti-anti</i> $^{13}\text{C}=\text{O}$	<i>anti-anti</i> $^{13}\text{C}-\text{OH}$	<i>anti-anti</i> $\text{C}=\text{}^{18}\text{O}$	<i>anti-anti</i> $\text{C}-\text{}^{18}\text{OH}$
$A$ <sup>[a]</sup>	2585.8670 (19) <sup>[e]</sup>	2577.5799 (54)	2580.2216 (72)	2540.2913 (16)	2537.48 (10)
$B$	2382.1261 (19)	2375.2038 (59)	2376.7693 (39)	2304.6537 (14)	2311.098 (93)
$C$	1239.5812 (15)	1235.8132 (23)	1236.8595 (16)	1208.04001 (21)	1209.118 (13)
$\Delta I$ <sup>[b]</sup>	0.10789 (54)	0.1043 (10)	0.09928 (84)	0.11464 (20)	0.1224 (13)
$N$ <sup>[c]</sup>	13	5	6	4	4
$\sigma$ <sup>[d]</sup>	7.8	16.3	15.4	0.3	18.3

[a]  $A$ ,  $B$ , and  $C$  represent the rotational constants (in MHz); [b]  $\Delta I$  is the inertial defect (in  $\text{u}\text{\AA}^2$ ), conversion factor: 505379.1 MHz -  $\text{u}\text{\AA}^2$ ; [c]  $N$  is the number of measured transitions; [d]  $\sigma$  is the rms deviation of the fit (in kHz); [e] Standard error in parentheses in units of the last digit.



**Figure 4.** (a) Zoom-in section (8430-8685 MHz) on the broadband LA-CP-FTMW spectrum showing  $^{13}\text{C}=\text{O}$  (orange) and  $^{13}\text{C}-\text{OH}$  (garnet) *anti-anti* SA isotopologues in comparison to the parent species (*anti-anti* SA in blue and *syn-anti* SA in red). No lines were observed from 8475 to 8485 MHz in the broadband spectrum. (b) We depict the rotational transition  $3_{13} \leftarrow 2_{02}$  of *anti-anti*  $\text{C}-\text{}^{18}\text{OH}$  (green) and *anti-anti*  $\text{C}=\text{}^{18}\text{O}$  (dark blue) SA isotopologues observed during the LA-MB-FTMW experiment. The resonance frequency is determined by the arithmetic mean of two Doppler components. The LA-MB-FTMW spectrum was obtained by averaging 300 experimental cycles (4 free induction decay signals per cycle).

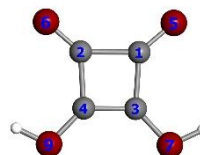


**Table 3.** Substitution structure of *anti-anti* SA.

Atom	a	b	c  <sup>[c]</sup>
C <sub>1</sub>	0.7724 (20) <sup>[a]</sup>	0.8082 (19)	0.0
C <sub>3</sub>	0.6814 (22)	0.6615 (23)	0.0
O <sub>5</sub>	1.69174 (89)	1.59456 (94)	0.0
O <sub>7</sub>	1.6046 (13)	1.6250 (17)	0.0

Parameter	r <sub>s</sub>	r <sub>e</sub> <sup>[b]</sup>
r(C <sub>1</sub> -O <sub>5</sub> )	1.2097 (21)	1.2016
r(C <sub>3</sub> -O <sub>7</sub> )	1.3344 (27)	1.3219
r(C <sub>3</sub> -C <sub>4</sub> )	1.3629 (32)	1.3744
r(C <sub>1</sub> -C <sub>2</sub> )	1.5448 (28)	1.5566
r(C <sub>1</sub> -C <sub>3</sub> )	1.4726 (30)	1.4887
a(C <sub>2</sub> -C <sub>1</sub> -C <sub>3</sub> )	86.46 (15)	86.49
a(C <sub>3</sub> -C <sub>1</sub> -O <sub>5</sub> )	134.08 (18)	133.89
a(C <sub>2</sub> -C <sub>1</sub> -O <sub>5</sub> )	139.46 (17)	139.61
a(C <sub>1</sub> -C <sub>3</sub> -O <sub>7</sub> )	132.48 (19)	132.78
a(C <sub>1</sub> -C <sub>3</sub> -C <sub>4</sub> )	93.54 (18)	93.51



[a] Standard error in parenthesis in units of the last digit. [b] Equilibrium structure calculated at the B3LYP/6-311++G(d,p) level. [c] Assumed planar skeleton (see text). Bond distances are given in Angstroms (Å) and bond angles in degrees.



## 2.4. Water clusters

Once all the spectral signatures corresponding to both *anti-anti* and *syn-anti* conformers, as well as their corresponding isotopomers, were discarded from the spectrum, a reasonable number of rotational lines remained unassigned. Based on the presence of the strong water dimer transitions [see Figure 3(a)], these lines presumably belong to the solvated SA species formed with residual water. Thus, we carried out conformational searches to predict water clusters, and after correspondent geometry optimization, only two of them lay in an energy window of 1000 cm<sup>-1</sup>.

According to DFT predictions, SA water clusters can adopt two different configurations denoted as *w-anti-anti* and *w-syn-anti* forms, depicted in Table 4. The estimated values of the rotational constants and electric dipole-moment components are also collected in Table 4. On this basis, we were able to identify independent sets of *b*-type *R*-branch transitions (spaced at approximately 2C) belonging to the complexes above. This assignment was extended to other *a*- and *b*-type rotational lines. Careful searches were performed looking for *c*-type transitions for both rotamers, but they were

not observed. All measured lines (listed in Tables D2.5-D2.6 in Appendix D, Section D.6) were analyzed by using a rigid rotor approximation. The determined spectroscopic parameters are also listed in Table 4. The values of the inertial defect clearly suggest nearly planar structures, with the skeleton framework lying on the *ab* inertial plane. Furthermore, the small negative values of -0.62567 and -0.60995 uÅ<sup>2</sup> are consistent with the out-of-plane contributions of the water hydrogens (Villamañán, et al 1989; Alonso & Villamañán 1989). The excellent agreement between experimental and calculated values of the rotational constants is reflected in the scaling factors, which range from 0.99773 to 1.00045. Hence, for both complexes, the DFT geometry can be assumed as the actual structure.

**Table 4.** Theoretical and experimental spectroscopic parameters for the conformers of the SA-water cluster.

	<i>Experiment</i>	<i>DFT</i> <sup>[a]</sup>	<i>Experiment</i>	<i>DFT</i>
				
<i>A</i> <sup>[b]</sup>	2537.7599 (15) <sup>[h]</sup>	2541	2456.9117 (32)	2462
<i>B</i>	1166.5963 (73)	1168	1192.5704 (18)	1192
<i>C</i>	799.9969 (38)	801	803.6433 (66)	804
$\Delta I$ <sup>[c]</sup>	-0.6256 (42)	-0.74	-0.6099 (88)	-0.72
$ \mu_a / \mu_b / \mu_c $ <sup>[d]</sup>	Yes/Yes/No	1.2/2.2/1.1	Yes/Yes/No	3.2/4.2/1.1
<i>N</i> <sup>[e]</sup>	17	-	17	-
$\sigma$ <sup>[f]</sup>	11.2	-	17.5	-
$\Delta E_{TOTAL}/\Delta G$ <sup>[g]</sup>	-	0/0	-	568/567

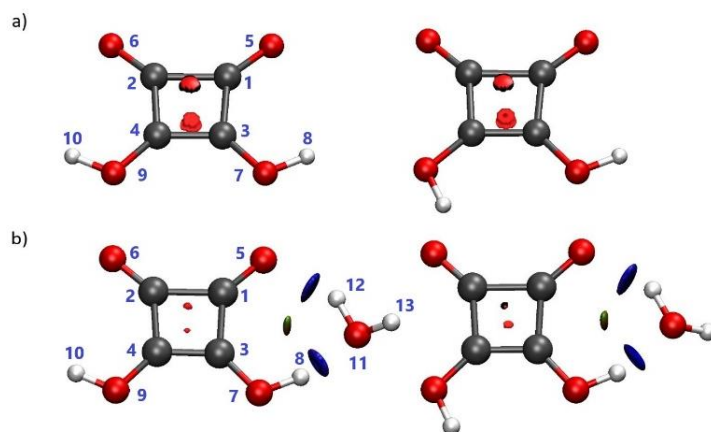
[a] Theoretical calculations at B3LYP/6-311++G (d,p) level of theory; [b] *A*, *B*, and *C* represent the rotational constants (in MHz); [c]  $\Delta I$  is the inertial defect (in uÅ<sup>2</sup>), conversion factor: 505379.1 MHz - uÅ<sup>2</sup>; [d]  $\mu_a$ ,  $\mu_b$ , and  $\mu_c$  are the electric dipole moment components (in D); [e] *N* is the number of measured transitions; [f]  $\sigma$  is the rms deviation of the fit (in kHz); [g]  $\Delta E_{TOTAL}$  and  $\Delta G$  are the electronic and Gibbs energies (in cm<sup>-1</sup>) at 298 K relative to the global minimum calculated at the B3LYP/6-311++G(d,p) level of theory with Grimme Dispersion, taking into account the zero-point vibrational energy (ZPE) for the electronic energy ( $E_{TOTAL} = E + E_{ZPE}$ ) calculated at the same level; [h] Standard error in parentheses in units of the last digit.

### 2.5. The interplay between inter- and intramolecular interactions

It is known that molecules having intramolecular hydrogen bonds are usually expected to form addition and insertion complexes with H<sub>2</sub>O molecules. On the one hand, the addition complex implies the formation of a hydrogen bond between the water molecule and one of the available sites of the target molecule, without disruption of the existing intramolecular hydrogen bond. In the generation process of the insertion complex, breaking of the intramolecular hydrogen bond and further reconstruction of the intermolecular hydrogen-bond network caused by the H<sub>2</sub>O insertion need to arise simultaneously. Usually, the addition complexes are higher in energy than the insertion complexes. Squaric acid possesses close donor (hydroxy) and acceptor (carbonyl) groups, so it could be an ideal entity to form multiple inter/intramolecular hydrogen bonds.

The observation of these novel aggregates stabilized by plausible inter- and intramolecular hydrogen-bond networks (different non-covalent interactions) raises a few questions. a) What is the interplay between inter- and intramolecular interactions? b) Do SA monomers form strong intramolecular hydrogen bonds? c) Is it reasonable to think about a significant competitiveness between intramolecular hydrogen bonds from the monomers and the insertion of water molecules?

In a quest to confirm the existence of intramolecular interactions for the monomers, QTAIM and complementary NCI methods were executed on the B3LYP/6-311++G(d,p) optimized structures. We can, therefore, visualize weak non-covalent interactions from the topological analysis of the electron density ( $\rho$ ) and its reduced gradient. As is shown in Figure 5(a) for the monomers, *anti-anti* and *syn-anti* conformers do not exhibit any critical bond points (BCPs) for plausible *anti*-O<sub>9</sub>-H<sub>10</sub>⋯O<sub>6</sub> nor *syn*-O<sub>9</sub>-H<sub>10</sub>⋯O<sub>7</sub> intramolecular hydrogen bonds through QTAIM analysis. For detailed information, see Section D.5 of the Appendix D, Topological Analysis of Chemical Bonding [SA and (SA-water) clusters]. However, the calculated B3LYP/6-311++G(d,p) distance of the *anti*-O<sub>9</sub>-H<sub>10</sub>⋯O<sub>6</sub> interaction (2.926 Å) is shorter than that of *syn*-O<sub>9</sub>-H<sub>10</sub>⋯O<sub>7</sub> (2.960 Å), in good agreement with the calculated relative stabilities.



**Figure 5.** a) VMD representations of *anti-anti* (left) and *syn-anti* (right) NCI analysis (Humphrey et al. 1996). Only NCI type **II** isosurfaces are found (labeled in red), which corresponds to strong destabilizing interactions due to a sum of contributions from nearby C atoms (ring closure and steric crowding). (Gillet et al. 2012) b) VMD representations of *w-anti-anti* (left) and *w-syn-anti* (right) NCI analysis. NCI Type **I** isosurfaces (strong stabilizing interactions, archetypal case of the H-bond) are labelled in blue; NCI type **II** isosurfaces (strong destabilizing interactions, according to the diketocyclobuten- ring closures) are represented in red; NCI type **III** (delocalized weak Van der Waals interactions) are labelled in green. NCI color code: BGR.

In the second step of our study, we were able to rationalize the enhanced stabilization of SA-water clusters. In this case, QTAIM and NCI calculations note the presence of BCPs for  $O_7-H_8 \cdots O_{11}$  and  $C_1-O_5 \cdots H_{12}$  interactions in *w-anti-anti* and *w-syn-anti*, although no  $O_9-H_{10} \cdots O_7$  BCPs are found [see Figure 5(b)]. Water molecules interact through the carbonyl group, simultaneously acting as a proton donor to the carbonyl oxygen atom and accepting a proton from the hydroxy group. This configuration benefits from two intermolecular hydrogen bonds, with water bridging the carbonyl and hydroxy groups of SA in a cyclic planar structure, and it is confirmed by the identification of ring critical points (RCPs) in the QTAIM analysis. The structures of both conformers of the SA--water cluster retain the preferred *anti-anti* and *syn-anti* conformation of unsolvated *anti-anti* and *syn-anti* SA, respectively. Plausible formation of hydrogen-bond networks, in fact, reinforce their stability as the water molecule in the complex does not enter into competition with any strong intramolecular hydrogen bonds in squaric acid. This is also supported by the solvation free energy values [ $\Delta G_{\text{solv}}$  (295.15 K, 1 atm)] calculated at the B3LYP/6-311++G(d,p) level,  $\Delta G_{\text{solv}}(\textit{anti-anti}) = -24.0 \text{ kJ mol}^{-1}$ ,  $\Delta G_{\text{solv}}(\textit{syn-anti}) = -17.2 \text{ kJ mol}^{-1}$ , which correspond to exothermic processes for both cases. In conclusion, we can infer that hydrogen bonds are one of the most important forces driving the interactions in

more complex SA-based molecular or even supramolecular systems as was already proposed for the solid structures of squarate ions (Prohens et al. 2014).

### 3. CONCLUSIONS

The molecular shape of squaric acid has been investigated in the isolated conditions of the gas phase. The most stable structures have been unequivocally identified in the rotational spectrum, showing that the most abundant conformer (*anti-anti* SA) differs from the *syn-anti* conformation present in the hydrogen-bonding chains extended along with the crystal structure.<sup>[8]</sup> The orientation of the hydroxy groups is relevant to distinguish between the different conformers of squaric acid and might account for the differences in the chemical and physical properties of this biologically relevant molecule. Furthermore, the planarity of both squaric acid conformers has been further corroborated by evaluating the values of the inertial defects. Also, the substitution structure of the heavy atom skeleton of *anti-anti* SA has been analyzed. This precise structural data may be useful in future studies as a starting point in the design of new SA-based synthones in supramolecular chemistry.

Water clusters formed owing to the presence of residual water have been further unveiled in the supersonic expansion. Moreover, plausible inter/intramolecular hydrogen bonds involving hydroxy and carbonyl groups from squaric acid and water molecules have been fully described by using a combination of theoretical calculations and a topological analysis by QTAIM and NCI methods. This study is of utmost importance not only to describe the squaric acid conformational landscape and its structure but also to serve as the simplest molecular model of squaric acid-water interactions, which represents the initial step of the hydration process.

Structural studies of labile organic molecules have started a new era with the development of new rotational spectroscopic techniques combined with laser ablation. It is now possible to place nonvolatile molecules in the gas phase and explore the most stable conformers and their hydrated clusters generated in a supersonic expansion. Future investigations are presently directed toward the study of squaric acid derivatives, such as squaramides, and another relevant bio/inorganic agents and organocatalysts. This would, therefore, contribute to improving our understanding of the factors that modulate their interactions in a considerable number of synthetic and biological processes.

## 4. EXPERIMENTAL SECTION

### *LA-CP-FTMW spectroscopy*

In an attempt to provide rotational signatures of the isolated SA free from environmental disturbances to unravel its intrinsic conformational preferences, we used broadband chirped pulse Fourier transform microwave (CP-FTMW) spectroscopy equipped with a laser ablation vaporization system (Bermúdez et al. 2016). In the experimental procedure, finely powdered SA was mixed with a small amount of a commercial copolymeric binder and pressed into cylindrical rods, which were ablated by using the third harmonic (355 nm, 20 mJ per pulse) of an Nd-YAG picosecond laser. The vaporized products were seeded in neon at a backing pressure of 10 bar and adiabatically expanded into the vacuum chamber of the spectrometer, and then probed by laser ablation broadband chirped pulse Fourier transform microwave spectroscopy (LA-CP-FTMW). A high-power excitation pulse of 300 W was used to polarize the molecules at frequencies from 6.5 to 14 GHz. Up to 75.000 individual free induction decays (4 FIDs on each valve cycle at a 2 Hz repetition rate) were averaged in the time domain, and then Fourier transformed to obtain the broadband frequency domain spectrum.

### *LA-MB-FTMW spectroscopy*

In a second step, we employed a laser ablation molecular beam Fourier transform microwave (LA-MB-FTMW) spectrometer, which operates in the frequency range 8-18 GHz, described elsewhere (Alonso et al. 2009; Bermúdez et al. 2014), to search for SA isotopologs. The laser-ablated products were seeded in Ne (stagnation pressure 10 bar) and expanded adiabatically to form a supersonic jet into a Fabry-Pérot resonator. A short microwave radiation pulse (0.3  $\mu$ s) was then applied to polarize the vaporized molecules macroscopically. The microwave transient FID was again registered in the time domain and Fourier transformed to the frequency domain. The pulsed molecular beam was introduced parallel to the axis of the resonator, so each observed transition appears as a Doppler doublet.

## 5. BIBLIOGRAPHY

- Alonso, E. R., Kolesníková, L., Bialkowska-Jaworska, E. et al., *Astrophys. J.* 2018, 861, 70.
- Alonso, J. L., Lozoya, M. A., Peña, I., López, J. C., Cabezas, C., Mata, S., Blanco, S., *Chem. Sci.* 2014, 5, 515–522.
- Alonso, J. L., Pérez, C., Sanz, M. E., López, J. C., Blanco, S. *Phys. Chem. Chem. Phys.* 2009, 11, 617–627.
- Alonso, J. L., López J. C., in *Gas-Phase IR Spectroscopy and Structure of Biological Molecules* (Eds.: A. M. Rijs, J. Oomens), Springer International Publishing, Cham), 2015, p. 335.
- Alonso, J. L. & Villamañán, R. M. *J. Chem. Soc. Perkin Trans. 2* 1989, 85, 137–149.
- Andrés, J. M., Losada, J., Maestro, A., Rodríguez-Ferrer, P. Pedrosa, R. *J. Org. Chem.* 2017, 82, 8444 – 8454.
- Bader, R. F., *Atoms in molecules*, in *Encyclopedia of Computational Chemistry*, Wiley, Hoboken, 2002
- Badiola, E., Fiser, B., Gjmez-Bengoia, E., et al. *J. Am. Chem. Soc.* 2014, 136, 17869 – 17881.
- Bermúdez, C., Mata, S., Cabezas, C., Alonso, J. L., *Angew. Chem. Int. Ed.* 2014, 53, 11015–11018; *Angew. Chem.* 2014, 126, 11195 – 11198.
- Bermúdez, C., Peña, I., Mata, S., Alonso, J. L., *Chem. Eur. J.* 2016, 22, 16829 – 16837 and references therein.
- Chaudret, R., de Courcy, B., Contreras-García, J., Gloaguen, E. et al. *Phys. Chem. Chem. Phys.* 2014, 16, 9876 – 9891.
- Chen, S., Pan, J., Wang, Y., Zhou, Z., *Eur. J. Org. Chem.* 2014, 7940 – 7947.
- Cohen, S., Lacher, J. R., Park, J. D. *J. Am. Chem. Soc.* 1959, 81, 3480.
- Das, N., Ghosh, A. Arif, A. M., Stang, P. J., *Inorg. Chem.* 2005, 44, 7130 – 7137.
- Davis, A. P., Draper, S. M., Dunne, G., Ashton, P., *Chem. Commun.* 1999, 2265 – 2266.
- Gillet, N., Chaudret, R., Contreras-Garcia, J., Yang, W. T. et al. *Chem. Theory Comput.* 2012, 8, 3993–3997.
- Gordy, W. & Cook, R. L. *Microwave Molecular Spectra*, Interscience Pub., New York, 1970, Chap. III, pp. 58–70 and Chap. XIII, pp. 495–547.
- Hill, N. D., Bunata, K., Hebert, A., *Clin. Dermatol.* 2015, 33, 300 – 304.
- Humphrey, W., Dalke, A., Schulten, K., *J. Molec. Graphics* 1996, 14, 33 –38.
- Kraitchman, J. *Am. J. Phys.* 1953, 21, 17.
- Matta, C. F., Boyd, R. J., *An introduction to the quantum theory of atoms in molecules*, in *The Quantum Theory of Atoms in Molecules: From Solid State to DNA and Drug Design*, (Eds.: C. F. Matta, R. J. Boyd), Wiley-VCH, Weinheim, 2007, Chap. I.
- Miller, T. F., Clary, D. C., Meijer, A. J. H. M., *J. Chem. Phys.* 2005, 122, 244323.
- Lapinski, L., Reva, I., Rostkowska, J., et al., *J. Phys. Chem. A* 2013, 117, 5251–5259.
- Lee, K.-S., Kweon, J. J., Oh, I.-H., Lee, C. E., *J. Phys. Chem. Solids* 2012, 73, 890–895.
- Li, S.-C., Zhang, T., Deng, X.-P., Guo, X.-Q., et al., *Inorg. Chem. Commun.* 2018, 92, 59 –73.
- Liu, Z., Wang, U., Han, W., Liu, L., et al., *Eur. J. Med. Chem.* 2013, 65, 187 – 194.
- Oka, T. *J. Mol. Struct.* 1995, 352/353, 225 – 233 and references therein.

Park, J. D., Cohen, S., Lacher, J. R. *J. Am. Chem. Soc.* 1962, 84, 2919 – 2922.

Pickett, H. M. *J. Mol. Spectrosc.* 1991, 148, 371 – 377.

Prohens, R., Portell, A., Font-Barda, M., Bauz, A., Frontera, A., *Cryst. Growth Des.* 2014, 14, 2578– 2587 and references therein.

Rotger, M. C., Pina, M. N., Frontera, A., et al. *J. Org. Chem.* 2004, 69, 2302– 2308.

Schermann, J. P. *Spectroscopy and Modelling of Biomolecular Building Blocks*, Elsevier, Amsterdam, 2008.

Semmingsen, D., Tun, Z., Nelmes, R. J., et al., *Z. Kristallogr.* 1995, 210, 934– 947.

Skopinska-Wisniewska, J., Kuderko, J., Bajek, A. et al., *Mater. Sci. Eng. C* 2016, 60, 100–109.

Villamanan, R. M., López, J. C., Alonso, J. L. *J. Am. Chem. Soc.* 1989, 111, 6487 – 6491.

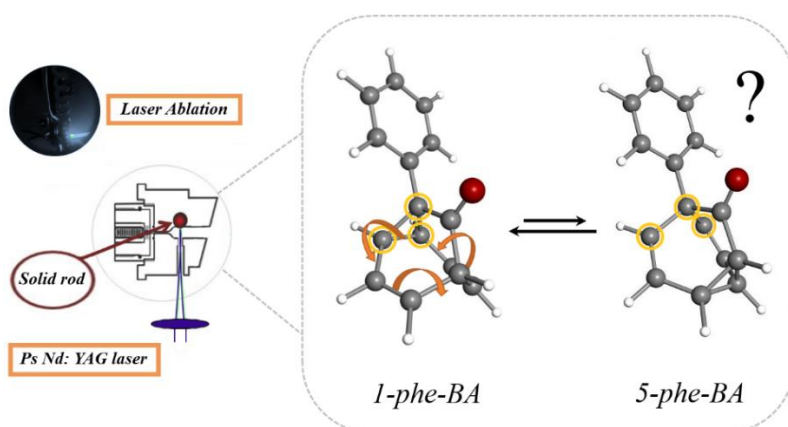
Wang, C.-C., Kuo, C.-T., Yang, J.-C., et al. *Cryst. Growth Des.* 2007, 7, 1476 – 1482.

West, R. & Niu, H.Y., *J. Am. Chem. Soc.* 1963, 85, 2580 – 2588.



## CHAPTER VIIc. SHAPE-SHIFTING MOLECULES: UNVELING THE VALENCE TAUTOMERISM PHENOMENA IN BARBARALONES

Adapted from Sanz Novo et al. 2021 under review.

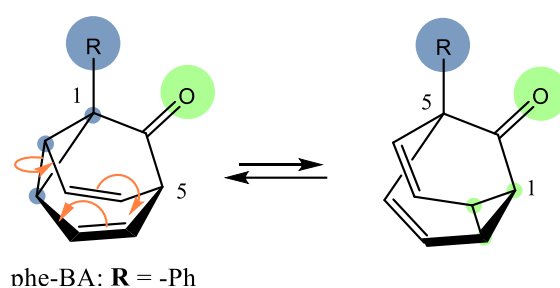


Herein we report a state-of-the-art spectroscopic study of an archetypical barbaralone, conclusively revealing the valence tautomerism phenomena for this bistable molecular system. The two distinct 1- and 5- substituted valence

tautomers have been isolated in a supersonic expansion for the first time and successfully characterized by high-resolution rotational spectroscopy. This work provides irrefutable experimental evidence of the [3, 3]-rearrangement in barbaralones and highlights the use of rotational spectroscopy to analyze shape-shifting mixtures. Moreover, this observation opens the window toward the characterization of new fluxional systems in the isolation conditions of the gas phase and should serve as a reference point in the general understanding of the valence tautomerism.

## 1. INTRODUCTION

Dynamic covalent chemistry (DCvC) is a coveted synthetic strategy that has attracted many theoretical and experimental studies for around a century (Rowan et al. 2002; Jin et al. 2013). It plays a key role in developing intricate supramolecular assemblies from discrete molecular building blocks (Jin et al. 2014; Nikitin et al 2019). Among all DCvC processes, of paramount importance are the sigmatropic rearrangements. This phenomena, also known as valence tautomerism, consists in a particular intramolecular mechanism involving just one reactive species where an intramolecular bond migrates giving rise to a usually different molecular structure (see Fig. 1).



**Figure 1.** a) Valence tautomerism of 1-substituted barbaralones, showing the redistribution of bonding electrons ascribed to the sigmatropic equilibrium. As shown, a  $\sigma$ -bond migrates leading to a reorganization of the two  $\pi$ -bonds, but the sequence of atoms within the carbon cage remains unchanged.

In this context, barbaralone, bullvalone, and bullvalene, widely well-known fluxional or “shape-shifting” molecules, have been deeply ingrained in understanding the valence tautomerism phenomena (Doering et al. 1963; Williams et al 2001; Hrovat et al 2005). These structurally unique dynamic molecules can interconvert among different constitutional isomers through low energy strain-assisted Cope rearrangements ([3, 3]-sigmatropic rearrangements). For instance, among 1.2 million degenerate isomers are anticipated for bullvalene (He et al. 2011). However, a lower number of valence tautomers exist for substituted bullvalenes (Volkman et al. 1984; Rebsamen et al. 1993; Poupko et al. 1996; Mgller et al. 1996; Luz et al. 1998; Lippert et al. 2006, 2009, 2010; Larson et al. 2012; Teichert et al. 2013), and only two discrete species are expected for barbaralone (Lambert et al. 1963; Nakanishi et al. 1975; Johnston et al. 1998). So far,

valence tautomerism phenomena has been investigated in the condensed phases by NMR and X-ray diffraction experiments (Katz et al. 1970; Turnblom et al. 1973; Poupko et al. 1984; Weissman et al. 1986. Meller et al. 1994; He et al. 2013).

It is well known that modern NMR spectroscopy has become the chemist's "choice by default" when it comes to the description of a chemical structure as well as dynamics of molecular systems. The successful determination of dynamic parameters depends on whether the target process is faster or slower than the corresponding NMR time scale. For fast processes, a simplified NMR spectrum is observed, which typically de-coalesce into a more complex, and therefore informative, array of peaks when we decrease the temperature of the experiment. Nevertheless, the dynamic equilibria may be perturbed due to interactions with the environment. The prior scenario highlights the relevance of a benchmark investigation on a fluxional system to decipher its intrinsic structural properties, therefore precluding laborious spectroscopic interpretations. Thus, to understand the structure of these molecular architectures, it is desirable first to study these systems in the absence of intermolecular interactions. Gas-phase experiments will enable to study of each structure involved in the dynamic equilibria separately, instead of trying to deconvolute those equilibria in the condensed phase. They should represent an essential source of reference data to unveil the phenomena of valence tautomerism and shall display exciting facets of even well-known systems which have not yet been realized.

In the present work, we report the first experimental investigation of the archetypal phenyl-substituted barbaralone (phe-BA from now) isolated in the gas phase to shed light on these shape-shifting molecular systems. Rotational spectroscopy has proven a powerful tool for unambiguously identifying chemical species in the gas phase and is acknowledged among the most robust techniques to distinguish between subtle changes in molecular geometry (Schermann 2008, Alonso et al. 2021). Since every molecule has a unique three-dimensional structure, it possesses a characteristic set of rotational constants. The major problem of working with gas-phase barbaralones is due to the labile nature of their solid samples and inherent vaporization difficulties. Fourier transform microwave spectroscopy, coupled with laser ablation techniques, breaks through the limitation of placing nonvolatile molecules in the gas phase, unraveling the most stable conformers generated in a supersonic expansion (Cabezas et al. 2013a). To date, detailed structural and conformational information has been reported on many building blocks using this experimental approach (Cabezas et al. 2013b, 2015, 2017; León et al. 2017, 2019).

The beauty of this approach is that the spectral assignments rely on the consistency of hundreds of observed rotational lines. In our particular case of study, different constitutional isomers, if are enough populated in the supersonic expansion, will present distinct spectroscopic constants and can be therefore unequivocally identified. The usefulness of rotational spectroscopy is such that not only it allows to distinguish between different conformers, but also to discern between similar molecules with minimal structural changes, such as an exchange in the position of one hydrogen atom (i.e. tautomers). This fact is strongly supported by the experimental observation of the tautomeric properties of several nitrogen basis as well as for creatinine (Alonso et al. 2009, Bermúdez et al. 2014, León et al. 2021). Even very similar compounds such as the 1- and 5- valence tautomers of phe-BA (see Fig.1) should be readily distinguished on this basis.

## 2. RESULTS AND DISCUSSION

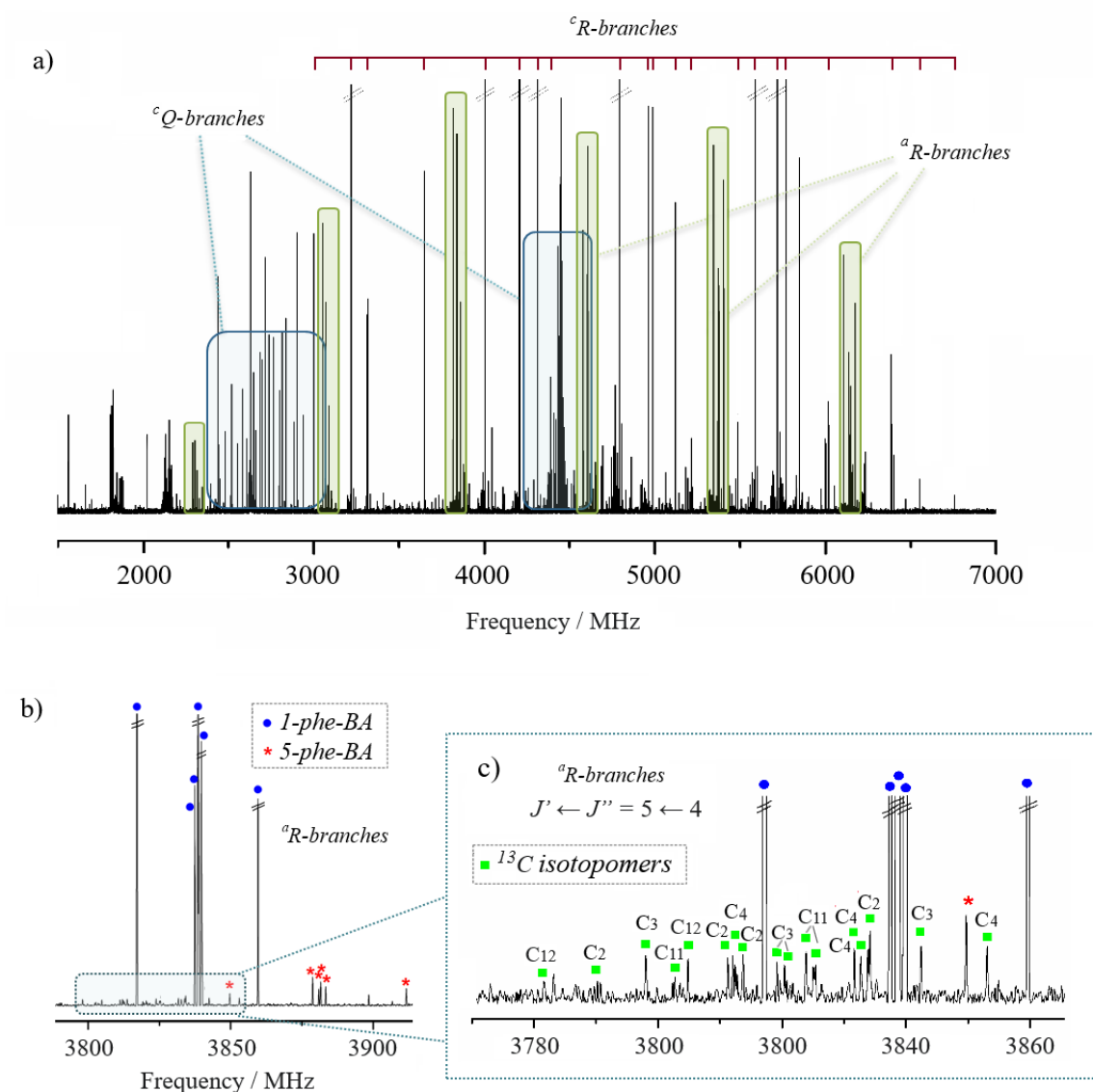
Phe-BA was synthesized following the general strategy previously reported by Ferrer & Echavarren (2016) to generate substituted barbaralones by a gold(I)-catalyzed oxidative cyclization of 7-ethynyl-1,3,5-cycloheptatrienes. We employed our laser ablation chirped-pulse Fourier transform microwave (LA-CP-FTMW) spectrometer (Bermúdez et al. 2016) to record the broadband spectrum of phe-BA from 1.0 to 7.0 GHz, which is presented in Fig. 2. Among the spectrum's dominating features, besides several intense  ${}^cR$ -branch lines, it was straightforward to identify strong  ${}^cQ$ -branch progressions ascribable to a first rotameric species (rotamer **I**, see Figure 2). Also, characteristic progressions of  ${}^aR$ -branch transitions ranging from  $J'' = 1$  to  $J'' = 7$  were quickly recognized and attributed to the same rotamer. Up to 147 rotational transitions were collected, and subsequently, least-squares fitted (Pickett et al. 1991) to semirigid-rotor Hamiltonian (Watson's A-reduction, I<sup>r</sup>-Representation). The obtained spectroscopic constants are listed in the first column of Table 1.

The assignment of an observed rotameric species to a discrete molecular structure is generally achieved by comparing the experimentally determined spectroscopic constants with those predicted theoretically for a set of candidate structures. With this aim, and concurrent to the experimental work, theoretical calculations were carried out on the two plausible 1- and 5- substituted phe-BA tautomers. We explore the potential energy surface

(PES) for each tautomer separately using fast molecular mechanics methods. For both systems, only one single conformation was found below an energetic window of 1000  $\text{cm}^{-1}$ . Then, we executed geometry optimizations for 1- and 5-phe-BA using high-level DFT and *ab initio* theoretical computations. The predicted spectroscopic constants for the global minimum of each PES are collected in Table 1.

The values of the rotational constants of rotamer **I** are, at first, consistent with those expected for a sizeable molecular species such as both the 1- and 5-phe-BA tautomers. Nevertheless, the two candidate structures exhibit minimal structural differences in the carbon cages (see Figure 1) that should induce a small but specific change in rotational constants' values, significant enough to allow conclusive discrimination. As we can see in Table 1, we find an excellent matching between the rotational constants' experimental values of rotamer **I** and those theoretically predicted for 1-phe-BA. Scaling factors ranging from 1.001 to 1.003 bring the predicted DFT rotational constants values to the experimental ones, highlighting a remarkable resemblance which further supports the unequivocal identification.

At this point, an intriguing question arises: Is the 5-phe-BA tautomer present in the supersonic expansion? In a quest to confirm the existence of the higher-in-energy tautomer, we removed the lines belonging to rotamer **I** and performed a thorough inspection in the spectrum's low-intensity background. Strikingly, using the above scaling factors, much weaker rotational features attributable to a second rotamer (rotamer **II**) following *a*- and *c*- type dipole moment selection rules were found. Thus, we were able to first assign *a*-branch  $(J+1)_{0J+1} \leftarrow (J+1)_{0J}$  transitions, which are spaced approximately  $B + C$  [see Figure 2(b)]. Subsequently, we extended the analysis to other *a*- and *c*-type lines and performed a least-square fit using a rigid-rotor approximation. We present the corresponding spectroscopic parameters in the second column of Table 1.



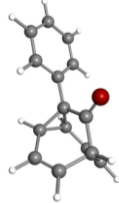
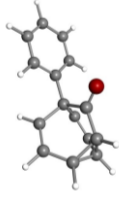
**Figure 2.** a) Jet-cooled broadband LA-CP-FTMW rotational spectrum of BA from 1.5 to 7.0 GHz, where intense  ${}^cQ$ -branches are remarked in blue,  ${}^aR$ -branch progressions are highlighted in green and intense  ${}^cR$ -branch lines are indicated in garnet. b) Section of the broadband spectrum showing a comparison of strong  ${}^aR$ -branch transitions ( $J' \leftarrow J'' = 5 \leftarrow 4$ ) for both tautomers, 1-phe-BA (in blue) and 5-phe-BA (in red); c) Zoomed view showing the same  ${}^aR$ -branch transitions of the  ${}^{13}\text{C}$  isotopomers of 1-phe-BA (in green). The intensity is in arbitrary units.

In this case, the experimental rotational constants for rotamer **II** are in excellent agreement with those calculated for 5-phe-BA. Thus, following the abovementioned arguments, rotamer **II** can be directly ascribed to 5-phe-BA. Moreover, for rotamer **II**,  $a$ - and  $c$ -type lines showed relatively similar intensities, which is in good agreement with the resemblance of the  $\mu_a$  and  $\mu_c$  calculated dipole-moment components of 5-phe-BA

(given in Table 1 for comparison), corroborating the conformational identification. Hence, in the present work we have provided the first experimental evidence of the higher in energy tautomer based on a line-by-line assignment of over 40 discrete spectroscopic features. The complete list of the measured experimental frequencies for both rotamers is shown in Table D7.1-D7.2 of [Appendix D \(Section D.7\)](#).

Regarding the valence tautomerism equilibrium, the activation energies for the [3.3]-sigmatropic rearrangements of barbaralones are relatively low (between 30 and 50 kJ/mol) compared to Cope rearrangements of acyclic systems (150 kJ/mol) (Bismillah et al. 2020). This low energy barriers imply that the interconversion rates between the possible constitutional isomers are way too rapid for room temperature  $^1\text{H}$  NMR. For instance, using this spectroscopic technique, we can only observe the chemical shift of the resonance for the distinct proton environments as a weighted average of those present in 1- and 5-phe-BA (the  $^1\text{H}$  NMR spectrum is presented [Section D.9 of Appendix D](#)). Another possible approach consists in forecasting the equilibrium distribution of isomers using low-temperature NMR (He & Bode 2013). However, as it has been shown, the isolation conditions granted by the supersonic expansion enable to study each individual structure separately, instead of trying to deconvolute the dynamic equilibria in the condensed phase (Bismillah et al. 2020), which may be perturbed due to interactions with the environment. In this context, we attempted to estimate the relative abundances of both tautomers using the intensity of the rotational transitions. In our experiment, the laser ablation process implies an ultrafast vaporization of the sample that does not allow for any sort of interconversion between tautomers. Thus, assuming that the population of both tautomers in the crystal is maintained during the course of the experiment (Alonso et al. 2014), we can estimate ratios of approximately  $97:3 \pm 1$  between the 1- and 5-phenyl-barbaralone based on a thorough analysis of the line's intensity. It is known that the presence and chemical nature of any possible substituent should induce isomerization restrictions. Within this framework, this work offers definite experimental proof suggesting that the synthetic approach carried out by Ferrer & Echevarren 2016 is highly selective on the formation of 1-substituted barbaralones.

**Table 1** Experimental spectroscopic parameters of the phe-BA tautomers compared with those theoretically predicted for the lowest in energy structures.

Parameters	Experimental		Theory <sup>[f]</sup>	
	Rotamer I	Rotamer II	1-phe-BA	5-phe-BA
				
$A^{[a]}$ / MHz	1274.35677 (46) <sup>[e]</sup>	1331.6505 (10)	1271.7	1327.5
$B$ / MHz	388.10352 (12)	394.31620 (45)	387.5	392.3
$C$ / MHz	379.62560 (12)	381.92829 (41)	378.5	380.8
$\Delta_{JK}$ / kHz	0.0742 (22)	-	-	-
$\delta_K$ / kHz	-5.519 (28)	-	-	-
$ \mu_a $ / D	observed	observed	1.2	2.5
$ \mu_b $ / D	not observed	not observed	0.0	0.0
$ \mu_c $ / D	observed	observed	3.1	2.6
$\sigma^{[b]}$ / kHz	10.7	14.4	-	-
$N^{[c]}$	147	43	-	-
$\Delta E^{[d]}$ / kcal mol <sup>-1</sup>	-	-	0.0	2.6

<sup>[a]</sup>  $A$ ,  $B$ , and  $C$  represent the rotational constants (in MHz);  $\mu_a$ ,  $\mu_b$ , and  $\mu_c$  are the electric dipole moment components (in D); <sup>[b]</sup> RMS deviation of the fit (in kHz). <sup>[c]</sup> Number of measured transitions. <sup>[d]</sup> Relative energies (in kcal/mol) concerning the global minimum taking into account the zero-point energy (ZPE). <sup>[e]</sup> Standard error in parentheses in units of the last digit. <sup>[f]</sup> Theoretical computations were conducted at the B3LYP-GD3/aug-cc-pVTZ level of theory.

Finally, we removed the lines belonging to 1- and 5-phe-BA and on the low-frequency side of each <sup>a</sup>R-branch transition of rotamer **I** very weak sets of the same  $R$ -branch progressions were identified [shown in Figure 2(b) for the <sup>a</sup>R-branch  $J' \leftarrow J'' = 5 \leftarrow 4$  progression], which could only be ascribed to five distinct <sup>13</sup>C isotopomers of the parent species in their natural abundance (1.109 %), with doubled intensity due to the existence of equivalent carbons (molecule with  $C_s$  symmetry). Hence, the isotopic substitution will allow us not only to facilitate the search for other species in the spectra but also to determine relevant structural parameters (i.e., bond distances). The final proof for the identification of rotamer **I** comes from the partial substitution structure ( $r_s$ ) of the heavy atom skeleton which was derived from the set of observed rotational constants in Table 1. Using the Kraitchman equations (Kraitchman 1953), the cartesian coordinates of



the substituted atoms are determined in the principal axis system (see Table 2). The derived C-C bond distances of the bonds implied in the electron migration are  $d(\text{C}_2\text{-C}_3) = 1.4805 (21) \text{ \AA}$  and  $d(\text{C}_3\text{-C}_4) = 1.3431 (81) \text{ \AA}$ , which coincides with typical C-C and C=C bond distances, respectively, irrefutably corroborating the identification of rotamer **I** as 1-phe-BA. The complete list of the measured experimental frequencies for both the  $^{13}\text{C}$  isotopomers of 1-phe-BA are shown in Tables S8.2-S8.6 of Appendix D (Section 8). Finally, after removing the rotational lines of 1- and 5-phe-BA and the  $^{13}\text{C}$  isotopomers of the parent species, a very humble rotational spectrum was obtained. Although we tried to search for other experimental rotamers, such as plausible photo-fragmentation products, (Kolesníková et al. 2021) no significant rotational signals remained unassigned.

**Table 2.** Substitution structure of 1-phe-BA

Atom	lal	lbl	lcl	Predicted equilibrium structure ( $r_e$ ) <sup>[b]</sup>	
$\text{C}_2, \text{C}_8$	0.9350 (17) <sup>[a]</sup>	0.7768 (21)	1.2714 (13)		
$\text{C}_3, \text{C}_7$	2.16219 (79)	1.5260 (11)	0.9186 (19)		
$\text{C}_4, \text{C}_6$	2.88180 (57)	1.2017 (14)	0.1681 (98)		
$\text{C}_{11}, \text{C}_{15}$	1.9800 (12)	1.2050 (19)	0.105 (22)		
$\text{C}_{12}, \text{C}_{14}$	3.37737 (70)	1.20190 (20)	0.107 (23)		
<b>Parameter</b>	<b><math>r_s</math></b>	<b><math>r_e</math></b>			
$r(\text{C}_2\text{-C}_3)$	1.4805 (21)	1.4713			
$r(\text{C}_3\text{-C}_4)$	1.3431 (81)	1.3344			

[a] Standard error in parenthesis in units of the last digit. [b] Equilibrium structure calculated at the B3LYP-GD3/aug-cc-pVTZ level. Bond distances are given in Angstroms ( $\text{Å}$ ). We show the equivalent carbon atoms due to the  $\text{C}_s$  symmetry of the molecule.

### 3. CONCLUSIONS

In summary, the molecular shape of a phenyl-substituted barbaralone has been investigated in the isolated conditions of the gas phase for the first time. Hence, the two distinct valence tautomers (1- and 5- substituted forms) have been unequivocally identified in the rotational spectrum, showing that the most abundant species (1-phe-BA) matches with the conformation expected along with the crystal structure. In addition, the C-C bond distances involved in the  $\pi$ -bond reorganization have been precisely

determined for the parent species, corroborating the presence of 1-phe-BA in the jet. Taken together, these observations provide compelling evidence of the shape-shifting rearrangement in barbaralones. Our investigation further corroborates that substituted barbaralones are chemically robust yet structurally dynamic, which is an important step toward their further use as starting materials and may bear implications in their chemical-physical properties in Material Science. So far, NMR spectroscopy has been employed for the analysis of shape-shifting mixtures, but it is a tedious task to conclusively identify the individual structures rather than derive averaged chemical properties. We're working furiously to explore the full scope and capabilities of rotational spectroscopy coupled to a laser ablation vaporization source as a reliable characterization technique to unravel dynamic molecular systems. Following on from this, further gas-phase studies exploring the valence tautomerism equilibrium of relevant fluxional molecules will be undertaken to pave the route in the characterization of new shape-shifting species.

#### 4. EXPERIMENTAL SECTION

Phenyl-barbaralone was prepared following a slight modification of the previously reported procedure in Ferrer & Echevarren (2016). Catalyst loading was reduced to 2 mol % for scaling up (the  $^1\text{H}$  NMR spectrum of the product is shown in Section D.9 of Appendix D).

In the experimental procedure, 350 mg of the finely powdered solid were mixed with a small amount of a commercial copolymeric binder and pressed it into cylindrical rods. Afterward, the third harmonic (355 nm, 15 mJ per pulse) of an Nd-YAG picosecond laser was employed to laser ablate the rods. The vaporized products were seeded in neon (backing pressure of 12 bar) and adiabatically expanded into the vacuum chamber of the spectrometer. Finally, the molecules populating the jet were probed by broadband chirped pulse Fourier transform microwave spectroscopy (LA-CP-FTMW) (Bermúdez et al. 2014). We used a high-power excitation pulse of 200 W (amplified using a solid-state amplifactory) to polarize the molecules at frequencies from 1.5 to 6.4 GHz. Up to 96000 individual free induction decays (4 FIDs on each valve cycle at a 2 Hz repetition rate) were averaged in the time domain, and Fourier transformed to obtain the final broadband frequency domain spectrum.

## 5. BIBLIOGRAPHY

- Alonso, J.L., Peña, I., López, J.C., Vaquero, V., *Angew. Chemie. Int. ed.* 2009, 48, 33, 6141-6143
- Alonso, J. L. Lozoya, M. A. Peña, I. et al. *Chem. Sci.* 2014, 5, 515–522.
- Alonso, E. R., León, I., Alonso, J. L., *Intra- and Intermolecular Interactions between Non-covalently Bonded Species*. Elsevier, 2021, 93-14
- Alonso, J. L. López, J. C. *In Gas-Phase IR Spectroscopy and Structure of Biological Molecules*, Springer International Publishing, Cham, 2015, p. 335
- Bermúdez, C., Mata, S., Cabezas, C., Alonso, J.L., *Angew. Chem. Int. Ed.* 2014, 53, 11015-11018;
- Bermúdez, C., Peña, I., Mata, S., Alonso, J. L. *Chem. Eur. J.* 2016, 22, 16829 and references therein.
- Bismillah, A. N., Chapin, B M., Hussein, B.A McGonigal, P. R., *Chem. Sci.*, 2020, 11, 324.
- Cabezas, C.; Peña, I. López, J. C.; Alonso, J.L., *J. Phys. Chem. Lett.* 2013, 4, 486-490.
- Cabezas, C. Varela, M. Cortijo, V. et al., *Phys. Chem. Chem. Phys.*, 2013, 15, 2580.
- Cabezas, C. Robben, M. A. T. Rijs, A. M et al., *Phys. Chem. Chem. Phys.*, 2015, 17, 20274–20280.
- Cabezas, C. Varela M. and Alonso, J. L. *Angew. Chem., Int. Ed.*, 2017, 129, 6520–6525.
- Doering, W. v. E. Roth, W. R. *Angew. Chem. Int. Ed. Engl.* 1963, 2, 115 – 122;
- Engdahl, C. Ahlberg, P. J. *Am. Chem. Soc.* 1979, 101, 3940 – 3946
- Ferrer, S., Echavarren, A. M., *Angew. Chem. Int. Ed.* 2016, 55, 11178 –11182.
- Hrovat, D. A. Brown, E. C. Williams, R. V. et al., *J. Org. Chem.* 2005, 70, 2627 – 2632
- Jin, Y., Yu, C., Denman, R. J., Zhang, W., *Chem. Soc. Rev.*, 2013, 42, 6634.
- Jin. Y., Taynton, P., Zhang, W., *Acc. Chem. Res.* 2014, 47, 5, 1575-1586;
- Johnston, E. R. Barber, J. S. Jacomet, M. Barborak, J. C. J. *Am. Chem. Soc.* 1998, 120, 1489 – 1493
- Katz, T. J. Carnahan, J. C. Clarke, G. M. Acton, N. *J. Am. Chem. Soc.* 1970, 92, 734–73
- Kolesníková, L., León, I., Alonso, E.R., Mata, S., Alonso, J.L. *Angew. Chem. Int.*, 2021, in publication.
- Lambert, J. B. *Tetrahedron Lett.* 1963, 4, 1901 – 1906
- Larson, K. K. He, M. Teichert, J. F. et al., *Chem. Sci.* 2012, 3, 1825 – 1828
- León, I. Alonso, E. R. Mata, S. et al., *Phys. Chem. Chem. Phys.*, 2017, 19, 24985–24990.
- León, I., Tasinato, N., Spada, L. et al. *ChemPlusChem*, 2021, 86, 1-14.
- León, I. Alonso, E. R. Mata, S. et al., *Angew. Chem., Int. Ed.*, 2019, 58, 16002–16007.
- Lippert, A. R. Kaeobamrung, J. Bode, J. W. *J. Am. Chem. Soc.* 2006, 128, 14738 – 14739
- Lippert, A. R. Keleshian, V. L. Bode, J. W. *Org. Biomol. Chem.* 2009, 7, 1529 – 1532
- Lippert, A. R. Naganawa, A. Keleshian, V. L. Bode, J. W. *J. Am. Chem. Soc.* 2010, 132, 15790 – 15799
- Luz, Z. Olivier, L. Poupko, R. et al., H. Zimmermann, *J. Am. Chem. Soc.* 1998, 120, 5526 – 5538.

- He, M., Bode, J. W., *Proc. Natl. Acad. Sci. USA* 2011, 108, 14752 – 14756.
- He, M. Bode, J. W. *Org. Biomol. Chem.* 2013, 11, 1306 – 1317.
- Kraitichman, J., *Am. J. Phys.* 1953, 21, 17.
- Meller, A. Haeberlen, U. Zimmerman, H. et al., *Mol. Phys.* 1994, 81, 1239–1258.
- Mgller, K. Zimmermann, H. Krieger, C. et al., *J. Am. Chem. Soc.* 1996, 118, 8006 – 8014
- Nakanishi, H. Yamamoto, O. *Chem. Lett.* 1975, 513 – 516
- Nikitin, K., O’Gara, R., *Chem. Eur.J.*2019, 25,4551 –4589.
- Pickett, H. M. *J. Mol. Spectrosc.* 1991, 148, 371–377.
- Poupko, R. Zimmermann, H. E. Luz, *Z. J. Am. Chem. Soc.* 1984, 106, 5391–5394.
- Poupko, K. Mgller, C. Krieger, H. Zimmermann, Z. Luz, *J. Am. Chem. Soc.* 1996, 118, 8015 – 8023
- Poupko, R. Zimmermann, H. Mgller, K. Luz, *Z. J. Am. Chem. Soc.* 1996, 118, 7995 – 8005;
- Rebsamen, K. Schröder, G. *Chem. Ber.* 1993, 126, 1419 – 1423;
- Rowan, S. J., Cantril S. J., Cousins, G. R. L., Sanders, J. K. M., Stoddart, J. F., *Angew. Chem. Int. Ed.* 2002, 41, 898 – 952;
- Schermann, J. P. *Spectroscopy and Modelling of Biomolecular Building Blocks*, Elsevier, Amsterdam, 2008;
- Teichert, J. F. Mazunin, D. Bode, J. W. *J. Am. Chem. Soc.* 2013, 135, 11314 – 11321.
- Turnblom, E. W. Katz, T. J. *J. Am. Chem. Soc.* 1973, 95, 4292–4311;
- Volkman, B. Schröder, G. *Chem. Ber.* 1984, 117, 2226 – 2232;
- Weissman, S. A. Baxter, S. G. Arif, A. M. et al., *J. Am. Chem. Soc.* 1986, 108, 529–531.
- Williams, R. V. *Chem. Rev.* 2001, 101, 1185 – 1204;

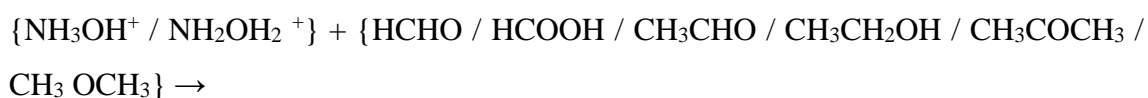
## CHAPTER VIII. CONCLUSIONS AND FUTURE TRENDS

---



The present memory gathers several research works involving molecules of mainly astrochemical but also organic and biochemical interest (Chapter II: Introduction). In this context, a myriad of experimental and theoretical studies has been carried out during this Ph.D. Most of them are in the framework of rotational spectroscopy and were performed employing different experimental and computational approaches and the micro- and millimeter-wave spectrometers presented in Chapter III. The following general conclusions and prospects are stated regarding both astro- and biochemical or organic subjects.

We have carried out a theoretical investigation of several ion-molecule gas-phase reactions between protonated hydroxylamine and acetic acid and between protonated hydroxylamine and cyanoacetaldehyde, respectively. These studies aim to analyze the feasibility under interstellar conditions of the gas-phase formation processes of **several isomers of protonated glycine** (Chapter IV) and examine the generation of **protonated cyanoacetamide** as a credible source of cyanoacetamide in the ISM (Chapter V). For each system, we have performed a full exploration of the corresponding singlet potential energy surfaces. Hence, we have found an exothermic process with no net activation barrier, and therefore feasible under the typical conditions of the ISM initiated in both cases by the higher in energy tautomer of protonated hydroxylamine, leading to protonated acetohydroxamic acid and protonated cyanoacetamide, respectively. These results suggest that the corresponding neutral systems, acetohydroxamic acid and cyanoacetamide should be interesting targets for interstellar detection. In the near future, we intend to extend these mechanistic studies to the formation of other plausible interstellar species employing, for instance, protonated hydroxylamine as starting compound:



On this basis, we reported the first spectroscopic studies in the centimeter-wave region of several COMs: **acetohydroxamic acid** ( $\text{CH}_3\text{C}(\text{O})\text{NHOH}$ , Chapter IV), **cyanoacetamide** ( $\text{CH}_2(\text{CN})\text{C}(\text{O})\text{NH}_2$ , Chapter V), and **cyanoacetic acid** ( $\text{CH}_2(\text{CN})\text{C}(\text{O})\text{OH}$ , Chapter V) as a preliminary step to identify these molecular systems in the ISM. We are starting to fill the gap between laboratory and observational data that lasted for several decades. We use state-of-the-art time-domain spectroscopic techniques equipped, if needed, with a laser ablation vaporization system (LA-CP-FTMW and LA-MB-FTMW). We have provided accurate experimental values of the rotational and  $^{14}\text{N}$  nuclear quadrupole coupling constants, which act as molecular probes for astronomical identification in low-frequency line surveys. Also, seeking to reproduce the rotational spectrum at higher frequencies and therefore enable a conclusive interstellar detection, we have extended our measurements up to 40 GHz using a newly configured double-resonance technique, which allowed us to determine the quartic centrifugal distortion

constants. In this context and motivated by the lack of high-frequency data of several proposed systems, we plan to build a new spectrometer that extends the laser ablation techniques to the millimeter-wave regime (75-120 GHz). This new instrument shall represent an essential source of new laboratory data of solid biomolecules at higher frequency ranges. For instance, we plan to extend the measurements of **acetohydroxamic acid** to the millimeter-wave region.

Afterward, we have carried out **radio astronomical searches** aiming to the detection of new COMs in collaboration with several leading institutions in the field of Astrophysics. In this context, we used the frequency-modulated millimeter- and submillimeter-wave spectrometer at the University of Valladolid to investigate the rotational spectrum of **glycolamide** ( $\text{CH}_2(\text{OH})\text{C}(\text{O})\text{NHOH}$ , Chapter IV), another glycine isomer, in a large spectral range (up to 460 GHz) as well as to analyze the rotational spectra of **normal-** and **iso-butyraldehyde** ( $\text{CH}_3(\text{CH}_2)_2\text{C}(\text{O})\text{H}$  and  $(\text{CH}_3)_2\text{CH}_2\text{C}(\text{O})\text{H}$ , Chapter VI) from 75 to 325 GHz. We have provided precise ground state spectroscopic constants up to sextic centrifugal distortion constants for each molecular system from the fit of thousands of transitions. These are the most precise descriptions of these molecules accessible to date, which have been used to prepare accurate line catalogs to be used in their subsequent search in the ISM. The most important conclusions of the astronomical results are presented as follows:

- 1) We have searched for glycolamide toward Sgr B2(N2) with ALMA and the envelope of Sgr B2(N) with the Effelsberg 100-m radio telescope. Although we reported a nondetection of glycolamide in both sources, the former lower limit ( $2.4 \times 10^{16} \text{ cm}^{-2}$ ) is only slightly higher than the column density ratio of acetaldehyde to glycolaldehyde. Our astrochemical models pointed out that the relative fractional abundances are pretty close to the observed upper limit for glycolamide in Sgr B2(N2), suggesting that a slightly more sensitive search in this source could indeed yield a future detection.
- 2) Moreover, we have searched for *n*- and *i*-butyraldehyde toward Sgr B2(N1) with ALMA and the G+0.693-0.027 molecular cloud using Yebes 40m and IRAM 30m observations. Additionally, we completed the study with a wider analysis including the search for shorter aldehydes toward both astronomical sources. We report a nondetection of *n*- and *i*-butyraldehyde toward Sgr B2(N1S) with ALMA

and toward G+0.693 using Yebes 40m and IRAM 30m observations. Our astronomical results indicate that the family of interstellar aldehydes in the Galactic center region is characterized by a drop of one order of magnitude in abundance at each incrementation in the level of molecular complexity. Comparison with astrochemical models indicates good agreement between observed and simulated abundances (where available). Grain-surface chemistry appears sufficient to reproduce aldehyde ratios in G+0.693-0.027; gas-phase production may play a more active role in Sgr B2(N1S). Model estimates for the larger aldehydes indicate that the observed upper limits may be close to the underlying values. In this regard, the inclusion of these molecules in a more comprehensive gas-grain chemical network would be highly valuable in providing more accurate predictions for future observational searches. Thus, we plan to perform gas-phase theoretical studies on the formation of these large aldehydes to help with the construction of the abovementioned chemical networks.

- 3) Finally, we have used our microwave data to search for **cyanacetamide** and **cyanoacetic acid** toward the hot core Sgr B2(N1) (A. Belloche priv. comm.) using ALMA as well as toward the cold cloud G+0.693-0.027 (V. M. Rivilla priv. comm.) using the Yebes 40-m and the IRAM 100-m telescopes without yielding any clear detection.

A highly exotic prospect derived from this Ph.D. is the study of interstellar COMs clustered, for instance, with one water molecule (i.e., the observed **cyanoacetic-water cluster** presented in **Chapter V**), aiming to the eventual identification of the water cluster in the ISM. Then a very intriguing question arises. Are such clusters expected to form in the gas phase, for instance at the typical high densities found in hot cores? This is an arduous question to answer. We expect that it would be complicated to form entirely in the gas-phase, given that the rate of three-body association reactions is relatively small for small molecules. But what about desorption of an already formed cluster from the grain mantles? In principle, it is not impossible that these water clusters could form in amorphous icy-mantles, and be desorbed perhaps in shocks. Thus, they should be plausible interstellar candidates in prominent sources such as the G+0.693-0.027 molecular cloud, where desorptions are non-thermal and supposed to be dominated by shocks, but not in hot cores. Within this framework, we plan to investigate several COMs-



water clusters of very abundant interstellar molecules, such as **methylcyanide-water**, **ethanol-water** and **formamide-water clusters**) and extend their frequency coverage to the millimeter-wave region to enable an eventual interstellar detection.

Besides the molecular systems presented in this thesis, mainly focused on interesting COMs that are candidates to be detected in the ISM, another important family of chemical species are molecules containing metals. Since the early detection of metal halides in the ISM, particularly in IRC+10216, the number of metal-bearing species identified in the ISM has been steadily growing and extended even to transition metal-bearing species, such as FeO, FeCN, TiO, and TiO<sub>2</sub>. Of great relevance was the detection of hydromagnesium isocyanide (HMgNC, Cabezas, et al., 2013). Following a previous work carried out in our group regarding candidate interstellar molecules that contain titanium (Redondo et al. 2019), we will therefore explore the full PESs and predict the corresponding structural and rotational/vibrational spectroscopic parameters of several new molecular systems following the HMNC/HMCN general formula (i.e., M= Al).

In parallel to the astrophysical backbone of this Dissertation and using the same experimental approach (LA-CP-FTMW and LA-MB-FTMW spectroscopies), we investigated the shape of the amino acid and supramolecular synthon **L-DOPA** (Chapter VII.a), to corroborate or conclusively dismiss its previously reported catechol ring-induced conformational restriction. We found a remarkably rich conformational panorama; more than four conformers have been irrefutably characterized, highlighting that no conformational restriction is presented for this molecular system. Moreover, we have thoroughly described the balance of intramolecular forces that drive its conformational space using a combination of theoretical computations and a topological analysis by NCI methods.

In the same context, we presented a comprehensive rotational investigation of the three-dimensional structure and conformational behavior of **squaric acid** (Chapter VII.b), another relevant supramolecular synthon, whose molecular structure has been deciphered at last in the isolation conditions of the gas phase. In addition, we have unveiled two squaric acid-water clusters in the supersonic expansion, which is the first step toward the study of this system in a more complex environment (i.e., an aqueous or even biological medium). This precise structural information should be helpful in future studies as a starting point in the design of new L-DOPA or squaric acid-based synthons in supramolecular chemistry.

Finally, we faced the fascinating topic of fluxional molecules, which has not been handled in our spectroscopic field so far. Hence, we have revealed the molecular shape of a **phenyl-substituted barbaralone** (Chapter VII.c) using high-resolution rotational spectroscopy. The two distinct valence tautomers (1- and 5- substituted forms) have been isolated in the supersonic jet and conclusively identified in the rotational spectrum, suggesting the most abundant tautomer (1-phe-barbaralone) matches the conformation expected along with the crystal structure. This study is of utmost importance since it definitively unveils the shape-shifting rearrangement in this fluxional systems and also describes the three-dimensional structure of this archetypal barbaralone.

Gas-phase structural studies of solid organic molecules have established a new era due to the development of new rotational spectroscopic techniques combined with laser ablation. We can transfer nonvolatile molecules to the gas phase and explore the most stable conformers and even their hydrated clusters. Thus, we are currently working to explore the full scope and impact of rotational spectroscopy coupled to a laser ablation vaporization source as a reliable characterization technique to unravel dynamic molecular systems, which seems a priori superior to NMR spectroscopy for the analysis of shape-shifting mixtures. As a continuation of our investigation of phenyl barbaralone, several substituted barbaralones will be undertaken by LA-CP-FTMW. These new projects will contribute to improving our understanding of the valence tautomerism phenomena and should shed some light on the characterization of more interesting fluxional or “shape-shifting” molecules, as well as other “hot topics” of physical-organic chemistry.

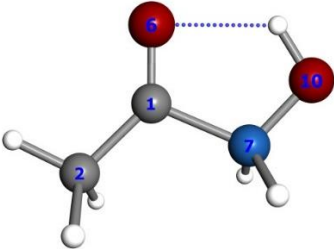
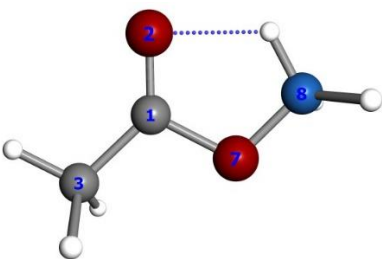
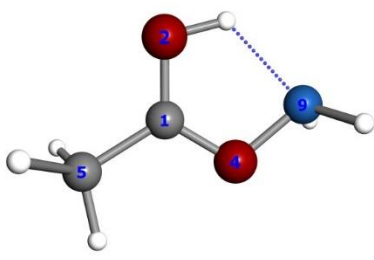
## REFERENCES

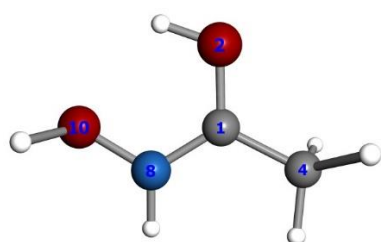
- Cabezas, C.; Cernicharo, J.; Alonso, J.L.; Agúndez, M.; Mata, S.; Guélin, M.; Peña, I. *Apj*, 2013, 775, 133.
- Redondo, P., Barrientos, C., Largo, A. *Apj.*, 2019, 871:180 (8pp).

*A.1 Complementary tables for the theoretical study of the protonated and neutral glycine isomers*

Tables A1 and A2 show the optimized geometry as well as the harmonic and anharmonic vibrational frequencies of the P1-P8 isomers, respectively, while Tables A3-A4 present the optimized geometry as well as the harmonic and anharmonic vibrational frequencies of the P1N-P8N isomers.

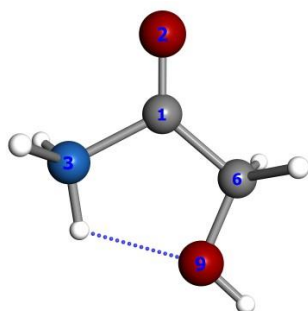
**Table A1.** Geometry optimized at both B3LYP/aug-cc-pVTZ and MP2/aug-cc-pVTZ levels of theory for P1-P8 isomers. Bond lengths are given in Angstroms and bond angles in degrees.

Isomer	Parameter	B3LYP	MP2
 <p><b>P1</b></p>	r (C <sub>2</sub> - O <sub>5</sub> )	1.2072	1.1824
	r (C <sub>1</sub> - C <sub>2</sub> )	1.4626	1.4767
	r (C <sub>1</sub> - H <sub>6</sub> )	1.0885	1.0861
	r (N - C <sub>2</sub> )	1.6249	1.6065
	r (N - O <sub>4</sub> )	1.3889	1.3867
	r (N - H <sub>9</sub> )	1.0254	1.0267
	a (N - C <sub>2</sub> - C <sub>1</sub> )	114.05	112.48
	a (O <sub>4</sub> - N - C <sub>1</sub> )	112.18	111.10
	a (O <sub>5</sub> - C <sub>2</sub> - N)	111.98	113.01
	a (H <sub>6</sub> - C <sub>1</sub> - C <sub>2</sub> )	112.12	109.94
	a (H <sub>7</sub> - C <sub>1</sub> - C <sub>2</sub> )	109.51	108.84
	a (H <sub>8</sub> - C <sub>1</sub> - C <sub>2</sub> )	108.12	109.73
	a (H <sub>9</sub> - N - C <sub>2</sub> )	109.34	110.34
	a (H <sub>10</sub> - N - C <sub>2</sub> )	110.15	110.41
a (H <sub>11</sub> - O <sub>4</sub> - N)	104.59	103.63	
 <p><b>P2</b></p>	r (C <sub>2</sub> - O <sub>3</sub> )	1.4475	1.4455
	r (C <sub>1</sub> - C <sub>2</sub> )	1.4792	1.4775
	r (C <sub>1</sub> - H <sub>7</sub> )	1.0918	1.0891
	r (C <sub>1</sub> - H <sub>8</sub> )	1.0858	1.0845
	r (C <sub>2</sub> - O <sub>5</sub> )	1.1896	1.1969
	r (N - O <sub>3</sub> )	1.4093	1.4033
	r (N - H <sub>10</sub> )	1.0240	1.0234
	r (N - H <sub>11</sub> )	1.0488	1.0485
	a (O <sub>3</sub> - C <sub>2</sub> - C <sub>1</sub> )	109.85	109.28
	a (N - O <sub>3</sub> - C <sub>2</sub> )	109.07	108.28
	a (O <sub>5</sub> - C <sub>2</sub> - O <sub>3</sub> )	117.91	118.15
	a (H <sub>6</sub> - C <sub>2</sub> - C <sub>1</sub> )	110.26	109.31
	a (H <sub>7</sub> - C <sub>2</sub> - C <sub>1</sub> )	108.85	108.83
	a (H <sub>8</sub> - C <sub>2</sub> - C <sub>1</sub> )	109.37	109.34
a (H <sub>9</sub> - N - O <sub>3</sub> )	108.63	108.40	
a (H <sub>10</sub> - N - O <sub>3</sub> )	108.68	108.40	
a (H <sub>11</sub> - N - O <sub>3</sub> )	105.32	105.37	
 <p><b>P3</b></p>	r (C <sub>2</sub> - O <sub>3</sub> )	1.3025	1.3054
	r (C <sub>1</sub> - C <sub>2</sub> )	1.4692	1.4673
	r (C <sub>1</sub> - H <sub>7</sub> )	1.0853	1.0841
	r (C <sub>1</sub> - H <sub>9</sub> )	1.0922	1.0904
	r (C <sub>2</sub> - O <sub>5</sub> )	1.2698	1.2682
	r (N - O <sub>3</sub> )	1.4596	1.4517
	r (N - H <sub>10</sub> )	1.0195	1.0197
	r (O <sub>5</sub> - H <sub>6</sub> )	0.9961	1.0020
	a (O <sub>3</sub> - C <sub>2</sub> - C <sub>1</sub> )	120.77	117.52
	a (N - O <sub>3</sub> - C <sub>1</sub> )	111.71	110.38
	a (O <sub>5</sub> - C <sub>2</sub> - O <sub>2</sub> )	121.43	120.69
	a (H <sub>6</sub> - O <sub>5</sub> - C <sub>2</sub> )	107.91	105.93
	a (H <sub>7</sub> - C <sub>2</sub> - C <sub>1</sub> )	111.01	110.62
	a (H <sub>8</sub> - C <sub>2</sub> - C <sub>1</sub> )	108.81	108.27
a (H <sub>9</sub> - C <sub>2</sub> - C <sub>1</sub> )	108.45	108.27	
a (H <sub>10</sub> - N - O <sub>3</sub> )	103.00	102.87	
a (H <sub>11</sub> - N - O <sub>3</sub> )	103.00	102.86	
	r (C <sub>1</sub> - O <sub>2</sub> )	1.2917	1.2921
	r (C <sub>1</sub> - C <sub>4</sub> )	1.4771	1.4745
	r (O <sub>2</sub> - H <sub>3</sub> )	0.9765	0.9783
	r (C <sub>4</sub> - H <sub>5</sub> )	1.0918	1.0899



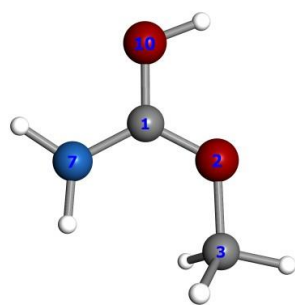
**P4**

r (C <sub>4</sub> - H <sub>7</sub> )	1.0868	1.0893
r (N - C <sub>1</sub> )	1.3102	1.3079
r (N - O <sub>10</sub> )	1.3790	1.3749
r (N - H <sub>9</sub> )	1.0139	1.0142
r (O <sub>10</sub> - H <sub>11</sub> )	0.9723	0.9735
a (H <sub>3</sub> - O <sub>2</sub> - C <sub>1</sub> )	112.19	111.00
a (C <sub>4</sub> - C <sub>1</sub> - O <sub>2</sub> )	118.21	118.32
a (H <sub>5</sub> - C <sub>4</sub> - C <sub>1</sub> )	108.68	108.36
a (H <sub>6</sub> - C <sub>4</sub> - C <sub>1</sub> )	108.76	108.40
a (H <sub>7</sub> - C <sub>4</sub> - C <sub>1</sub> )	112.18	111.79
a (N - C <sub>1</sub> - O <sub>2</sub> )	120.41	120.40
a (H <sub>9</sub> - N - C <sub>1</sub> )	122.76	122.81
a (O <sub>10</sub> - N - C <sub>1</sub> )	118.25	118.02
a (H <sub>11</sub> - O <sub>10</sub> - N)	107.02	106.32



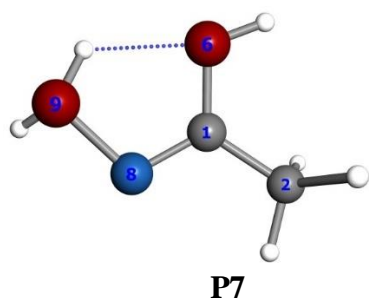
**P5**

r (C <sub>1</sub> - O <sub>2</sub> )	1.1701	1.1818
r (C <sub>1</sub> - C <sub>6</sub> )	1.5071	1.5066
r (C <sub>6</sub> - H <sub>7</sub> )	1.0957	1.0930
r (C <sub>6</sub> - O <sub>9</sub> )	1.4096	1.4112
r (N - C <sub>1</sub> )	1.5898	1.5638
r (N - H <sub>5</sub> )	1.0226	1.0230
r (N - H <sub>11</sub> )	1.0273	1.0278
r (O <sub>9</sub> - H <sub>10</sub> )	0.9637	0.9652
a (N - C <sub>1</sub> - O <sub>2</sub> )	117.42	117.68
a (H <sub>4</sub> - N - C <sub>1</sub> )	109.18	108.96
a (H <sub>5</sub> - N - C <sub>1</sub> )	109.17	108.88
a (C <sub>6</sub> - C <sub>1</sub> - O <sub>2</sub> )	130.62	130.08
a (H <sub>7</sub> - C <sub>6</sub> - C <sub>1</sub> )	106.73	106.92
a (H <sub>8</sub> - C <sub>6</sub> - C <sub>1</sub> )	106.74	106.97
a (O <sub>9</sub> - C <sub>6</sub> - C <sub>1</sub> )	108.34	107.35
a (H <sub>10</sub> - O <sub>9</sub> - C <sub>6</sub> )	110.19	110.38
a (H <sub>11</sub> - N - C <sub>1</sub> )	108.98	109.52

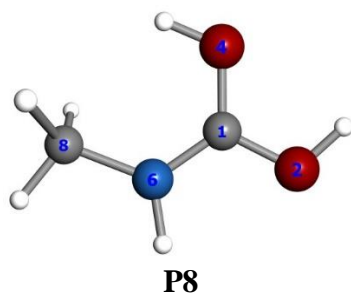


**P6**

r (C <sub>1</sub> - O <sub>10</sub> )	1.2990	1.2987
r (C <sub>3</sub> - H <sub>4</sub> )	1.0879	1.0872
r (C <sub>3</sub> - H <sub>5</sub> )	1.0830	1.0818
r (C <sub>3</sub> - O <sub>2</sub> )	1.4654	1.4615
r (C <sub>1</sub> - O <sub>2</sub> )	1.2870	1.2872
r (N - H <sub>8</sub> )	1.0068	1.0065
r (N - H <sub>9</sub> )	1.0103	1.0103
r (O <sub>10</sub> - H <sub>11</sub> )	0.9717	0.9730
a (C <sub>3</sub> - O <sub>2</sub> - C <sub>1</sub> )	121.80	119.63
a (H <sub>4</sub> - C <sub>3</sub> - O <sub>2</sub> )	109.32	109.11
a (H <sub>5</sub> - C <sub>3</sub> - O <sub>2</sub> )	104.51	104.43
a (H <sub>6</sub> - C <sub>3</sub> - O <sub>2</sub> )	109.33	109.11
a (H <sub>7</sub> - C <sub>6</sub> - C <sub>1</sub> )	125.36	125.22
a (N - C <sub>1</sub> - O <sub>2</sub> )	122.25	122.01
a (H <sub>9</sub> - N - C <sub>1</sub> )	119.24	118.86
a (O <sub>10</sub> - C <sub>1</sub> - O <sub>2</sub> )	117.52	117.75
a (H <sub>11</sub> - O <sub>10</sub> - C <sub>1</sub> )	111.96	110.74



r (C <sub>1</sub> - O <sub>6</sub> )	1.3276	1.3349
r (C <sub>1</sub> - C <sub>2</sub> )	1.4876	1.3809
r (C <sub>1</sub> - H <sub>3</sub> )	1.0919	1.0902
r (C <sub>1</sub> - H <sub>4</sub> )	1.0867	1.0860
r (O <sub>6</sub> - H <sub>7</sub> )	0.9678	0.9690
r (N - C <sub>2</sub> )	1.2854	1.2896
r (O <sub>9</sub> - N)	1.5464	1.4971
r (O <sub>9</sub> - H <sub>10</sub> )	0.9817	0.9891
r (O <sub>9</sub> - H <sub>11</sub> )	0.9748	0.9785
a (H <sub>3</sub> - C <sub>2</sub> - C <sub>1</sub> )	109.13	109.16
a (H <sub>4</sub> - C <sub>2</sub> - C <sub>1</sub> )	110.59	109.91
a (H <sub>5</sub> - C <sub>2</sub> - C <sub>1</sub> )	109.57	109.50
a (O <sub>6</sub> - C <sub>1</sub> - C <sub>2</sub> )	122.36	122.54
a (H <sub>7</sub> - O <sub>6</sub> - C <sub>1</sub> )	113.10	112.21
a (N - C <sub>1</sub> - O <sub>6</sub> )	120.73	119.33
a (O <sub>9</sub> - N - C <sub>1</sub> )	107.71	107.45
a (H <sub>10</sub> - O <sub>9</sub> - N)	112.45	110.42
a (H <sub>11</sub> - O <sub>9</sub> - N)	104.46	104.87



r (C <sub>1</sub> - O <sub>2</sub> )	1.2976	1.2979
r (C <sub>1</sub> - O <sub>4</sub> )	1.3027	1.3026
r (O <sub>2</sub> - H <sub>3</sub> )	0.9717	0.9730
r (O <sub>4</sub> - H <sub>5</sub> )	0.9696	0.9710
r (N - C <sub>1</sub> )	1.3045	1.3032
r (N - H <sub>6</sub> )	0.9696	1.0122
r (N - C <sub>8</sub> )	1.4705	1.4664
r (C <sub>8</sub> - H <sub>9</sub> )	1.0880	1.0870
a (H <sub>3</sub> - O <sub>2</sub> - C <sub>1</sub> )	112.37	111.22
a (O <sub>4</sub> - C <sub>1</sub> - O <sub>2</sub> )	116.96	117.11
a (H <sub>5</sub> - O <sub>4</sub> - C <sub>1</sub> )	114.77	113.44
a (N - C <sub>1</sub> - O <sub>2</sub> )	118.37	118.48
a (H <sub>7</sub> - N - C <sub>1</sub> )	115.82	115.85
a (C <sub>8</sub> - N - C <sub>1</sub> )	125.08	124.09
a (H <sub>9</sub> - C <sub>8</sub> - N)	110.64	110.33
a (H <sub>10</sub> - C <sub>8</sub> - N)	107.77	107.74
a (H <sub>11</sub> - C <sub>8</sub> - N)	110.64	110.34

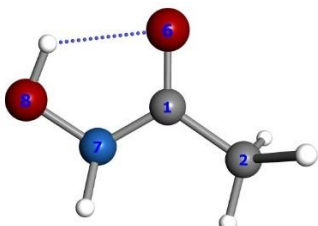
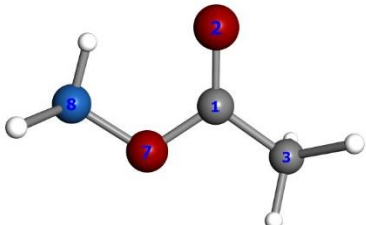
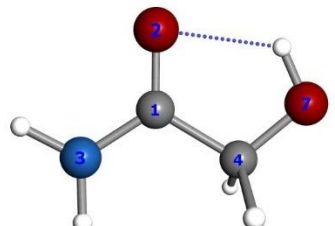
**Table A2.** Harmonic ( $\omega$ ), scaled harmonic ( $\omega_{sc}$ ), anharmonic ( $\nu$ ) vibrational frequencies and Infrared intensities (I) of P1-P8 isomers calculated at MP2/aug-cc-pVTZ level of theory. Frequencies are given in  $\text{cm}^{-1}$  and intensities in  $\text{km}\cdot\text{mol}^{-1}$ .

Isomer	Mode	$\nu / \text{cm}^{-1}$	I/ $\text{km}\cdot\text{mol}^{-1}$	$\omega / \text{cm}^{-1}$	$\omega_{sc} / \text{cm}^{-1}$	I/ $\text{km}\cdot\text{mol}^{-1}$	Vibration
<b>P1</b>	5	264	64	438	417	78	
	7	515	81	552	526	98	
	15	1275	186	1317	1256	207	
	19	1457	57	1510	1439	81	
	20	1571	26	1634	1558	40	NH <sub>2</sub> Bend
	21	1892	69	1921	1830	105	C=O Stretch
	22	2961	24	3072	2927	22	Symmetric C-H Stretch
	23	3013	5	3156	3008	5	Asymmetric C-H Stretch
	25	3225	60	3377	3219	82	Symmetric N-H Stretch
	26	2361	71	3454	3292	112	Asymmetric N-H Stretch
27	3352	118	3529	3363	137	O-H Stretch	
<b>P2</b>	8	720	175	739	705	173	C-O Stretch
	10	1013	15	1043	994	35	N-O Stretch
	12	1123	41	1167	1112	96	
	16	1400	167	1463	1394	260	NH <sub>3</sub> Bend
	19	1581	39	1648	1570	39	NH <sub>2</sub> Bend
	21	1852	59	1871	1783	133	C=O Stretch
	22	2984	44	3093	2947	18	Symmetric C-H Stretch
	23	2825	154	3098	2952	170	N-H Stretch
	24	3026	5	3179	3030	6	Asymmetric C-H Stretch
	25	3075	5	3225	3074	7	Symmetric C-H Stretch (CH <sub>2</sub> )
26	3274	153	3432	3271	152	Symmetric N-H Stretch (NH <sub>2</sub> )	
27	3331	137	3505	3340	152	Asymmetric N-H Stretch	
<b>P3</b>	8	884	11	905	862	19	N-O Stretch
	14	1239	92	1293	1232	162	
	17	1399	160	1449	1381	160	
	20	1591	89	1623	15467	132	C=O Stretch
	21	1611	128	1666	1587	177	NH <sub>2</sub> Bend
	22	2973	47	3085	2940	38	Symmetric C-H Stretch
	23	3049	7	3167	3018	16	Asymmetric C-H Stretch
	25	2832	59	3247	3094	323	O-H Stretch
	26	3308	61	3463	3300	63	Symmetric N-H Stretch
	27	3386	74	3559	3392	81	Asymmetric N-H Stretch
<b>P4</b>	4	270	87	393	375	112	
	9	702	55	740	705	73	
	14	1185	108	1233	1175	120	
	20	1540	43	1588	1514	78	C=O Stretch
	21	1664	94	1715	1634	216	C=N Stretch
	22	2977	21	3087	2942	22	Symmetric C-H Stretch (CH <sub>3</sub> )
	23	3034	8	3176	3026	12	Asymmetric C-H Stretch
	25	3408	208	3571	3403	201	N-H Stretch
	26	3438	146	3634	3463	175	O <sub>2</sub> -H <sub>3</sub> Stretch

	27	3519	171	3702	3528	189	O <sub>10</sub> -H <sub>11</sub> Stretch
<b>P5</b>	3	234	114	255	243	140	
	8	580	55	624	595	51	N-C Stretch
	12	972	36	1103	1051	61	C-O Stretch
	13	1193	72	1152	1098	158	
	16	1475	211	1401	1335	201	NH <sub>3</sub> Bend
	21	1886	109	1906	1816	138	C=O Stretch
	22	2960	12	3070	2925	14	Symmetric C-H Stretch
	23	2983	1	3126	2979	2	Asymmetric C-H Stretch
	24	3216	58	3354	3196	88	Symmetric N-H Stretch (NH <sub>3</sub> )
	25	3301	104	3474	3311	219	Symmetric N-H Stretch (NH <sub>2</sub> )
	26	3337	149	3510	3345	146	Asymmetric N-H Stretch
	27	3641	146	3819	3639	163	O-H Stretch
<b>P6</b>	6	561	303	540	514	327	
	10	878	73	905	862	74	
	12	1076	31	1106	1054	67	C-O Stretch
	16	1430	56	1464	1395	62	CH <sub>3</sub> Bend
	19	1555	392	1593	1518	68	NH <sub>2</sub> Bend
	20	1591	310	1630	1554	440	C=O Stretch (mixed)
	21	1714	391	1760	1677	441	N=C Stretch
	22	2992	2	3097	2951	2	Symmetric C-H Stretch
	25	3442	121	3578	3410	211	Symmetric N-H Stretch (NH <sub>2</sub> )
	26	3540	168	3715	3540	174	Asymmetric N-H Stretch
27	3540	240	3723	3548	260	O-H Stretch	
<b>P7</b>	2	155	30	233,1	222,1	87	
	6	449	50	438,5	417,9	110	
	11	866	42	941,0	896,8	102	
	16	1357	27	1392,1	1326,7	108	
	17	1398	116	1441,4	1373,7	102	C-C Stretch
	20	1438	31	1563,5	1490,0	100	OH <sub>2</sub> Bend
	21	1701	50	1736,6	1655,0	85	C=N Stretch
	22	2973	11	3080,9	2936,1	9	Symmetric C-H Stretch (CH <sub>3</sub> )
	23	3032	2	3169,7	3020,7	3	Asymmetric C-H Stretch
	25	3222	141	3455,2	3292,8	147	O-H Stretch
	26	3483	267	2669,5	2544,0	415	Asymmetric O-H Stretch
27	3585	147	3767,8	3590,7	225	O-H Stretch	
<b>P8</b>	4	441	287	451	430	309	C-O Stretch
	14	1142	80	1198	1142	153	
	15	1183	87	1236	1178	158	
	17	1436	27	1472	1402	26	CH <sub>3</sub> Bend
	20	1590	145	1631	1555	255	
	21	1732	384	1781	1697	518	C=N Stretch
	22	3000	2	3093	2948	2	Symmetric C-H Stretch
	24	3066	1	3203	3053	2	Asymmetric C-H Stretch
	25	3423	126	3594	3425	179	N-H Stretch
	26	3541	292	3723	3548	319	O <sub>2</sub> -H <sub>3</sub> Stretch
27	3558	138	37423	3567	150	O <sub>4</sub> -H <sub>5</sub> Stretch	



**Table A3.** Geometry optimized at both B3LYP/aug-cc-pVTZ and MP2/aug-cc-pVTZ levels of theory for P1N-P8N isomers. Bond lengths are given in Angstroms and bond angles in degrees.

Isomer	Parameter	B3LYP	MP2
 <p><b>P1N</b></p>	r (C <sub>2</sub> - O <sub>5</sub> )	1.2244	1.2299
	r (C <sub>1</sub> - C <sub>2</sub> )	1.5057	1.4997
	r (C <sub>2</sub> - H <sub>6</sub> )	1.0897	1.0879
	r (N - C <sub>2</sub> )	1.3603	1.3627
	r (N - O <sub>4</sub> )	1.3989	1.3970
	r (N - H <sub>10</sub> )	1.0070	1.0088
	r (O <sub>4</sub> - H <sub>9</sub> )	0.9795	0.9819
	a (N - C <sub>1</sub> - C <sub>2</sub> )	115.94	115.68
	a (O <sub>4</sub> - N - C <sub>2</sub> )	116.07	114.98
	a (O <sub>5</sub> - C <sub>2</sub> - N)	119.82	119.67
 <p><b>P2N</b></p>	a (H <sub>6</sub> - C <sub>1</sub> - C <sub>2</sub> )	112.98	112.42
	a (H <sub>7</sub> - C <sub>1</sub> - C <sub>2</sub> )	108.85	108.63
	a (H <sub>8</sub> - C <sub>1</sub> - C <sub>2</sub> )	108.86	108.69
	a (H <sub>9</sub> - O <sub>4</sub> - N)	100.82	100.29
	a (H <sub>10</sub> - N - C <sub>2</sub> )	121.82	120.07
	r (C <sub>1</sub> - O <sub>3</sub> )	1.3511	1.3540
	r (C <sub>1</sub> - C <sub>2</sub> )	1.5026	1.4976
	r (C <sub>2</sub> - H <sub>6</sub> )	1.0864	1.0850
	r (C <sub>2</sub> - H <sub>7</sub> )	1.0906	1.0889
	r (C <sub>2</sub> - O <sub>5</sub> )	1.2065	1.2126
r (N - O <sub>3</sub> )	1.4490	1.4442	
r (N - H <sub>9</sub> )	1.0172	1.0173	
r (N - H <sub>10</sub> )	1.0227	1.0238	
a (O <sub>3</sub> - C <sub>1</sub> - C <sub>2</sub> )	111.01	110.53	
a (N - O <sub>3</sub> - C <sub>2</sub> )	114.19	113.23	
a (O <sub>5</sub> - C <sub>2</sub> - O <sub>3</sub> )	123.36	123.41	
a (H <sub>6</sub> - C <sub>1</sub> - C <sub>2</sub> )	111.91	111.43	
a (H <sub>7</sub> - C <sub>1</sub> - C <sub>2</sub> )	108.58	108.28	
a (H <sub>8</sub> - C <sub>1</sub> - C <sub>2</sub> )	108.88	108.80	
a (H <sub>9</sub> - N - O <sub>3</sub> )	102.39	102.25	
a (H <sub>10</sub> - N - O <sub>3</sub> )	103.73	103.59	
 <p><b>P5N</b></p>	r (C <sub>1</sub> - O <sub>2</sub> )	1.2193	1.2250
	r (C <sub>1</sub> - C <sub>4</sub> )	1.5273	1.5207
	r (C <sub>1</sub> - N)	1.3537	1.3537
	r (C <sub>4</sub> - H <sub>5</sub> )	1.0971	1.0941
	r (C <sub>4</sub> - O <sub>7</sub> )	1.3991	1.4016
	r (O <sub>7</sub> - H <sub>8</sub> )	0.9719	0.9729
	r (N - H <sub>9</sub> )	1.0054	1.0051
	r (N - H <sub>10</sub> )	1.0038	1.0032
	a (N - C <sub>1</sub> - O <sub>2</sub> )	123.37	123.51
	a (C <sub>4</sub> - C <sub>1</sub> - O <sub>2</sub> )	120.14	120.34
a (H <sub>5</sub> - C <sub>4</sub> - C <sub>1</sub> )	108.77	108.85	
a (H <sub>6</sub> - C <sub>4</sub> - C <sub>1</sub> )	108.77	108.85	
a (O <sub>7</sub> - C <sub>4</sub> - C <sub>1</sub> )	110.29	110.01	
a (H <sub>8</sub> - O <sub>7</sub> - C <sub>4</sub> )	105.15	104.51	
a (H <sub>9</sub> - N - C <sub>1</sub> )	119.00	118.99	
a (H <sub>10</sub> - N - C <sub>1</sub> )	122.03	121.63	

<p><b>P6N</b></p>	r (C <sub>1</sub> - O <sub>2</sub> )	1.3564	1.3543
	r (C <sub>1</sub> - C <sub>3</sub> )	1.4333	1.4343
	r (C <sub>3</sub> - H <sub>4</sub> )	1.0886	1.0870
	r (C <sub>3</sub> - H <sub>5</sub> )	1.0860	1.0843
	r (C <sub>1</sub> - O <sub>10</sub> )	1.2085	1.2136
	r (N - C <sub>1</sub> )	1.3632	1.3644
	r (N - H <sub>8</sub> )	1.0040	1.0037
	r (N - H <sub>9</sub> )	1.0033	1.0035
	a (C <sub>3</sub> - O <sub>2</sub> - C <sub>1</sub> )	114.98	113.42
	a (H <sub>4</sub> - C <sub>3</sub> - O <sub>2</sub> )	110.75	110.42
	a (H <sub>5</sub> - C <sub>3</sub> - O <sub>2</sub> )	105.56	105.32
	a (H <sub>6</sub> - C <sub>3</sub> - O <sub>2</sub> )	110.70	110.39
	a (N - C <sub>1</sub> - O <sub>2</sub> )	110.10	109.92
a (H <sub>8</sub> - N - C <sub>1</sub> )	119.42	118.41	
a (H <sub>9</sub> - N - C <sub>1</sub> )	116.74	116.08	
a (O <sub>10</sub> - C <sub>1</sub> - O <sub>2</sub> )	124.57	124.57	
<p><b>P7N</b></p>	r (C <sub>1</sub> - O <sub>6</sub> )	1.3467	1.3465
	r (C <sub>1</sub> - C <sub>2</sub> )	1.4874	1.4831
	r (C <sub>2</sub> - H <sub>3</sub> )	1.0910	1.0854
	r (C <sub>2</sub> - H <sub>4</sub> )	1.0861	1.0890
	r (C <sub>1</sub> - N)	1.2778	1.2855
	r (O <sub>6</sub> - H <sub>7</sub> )	0.9699	0.9720
	r (N - O <sub>9</sub> )	1.4308	1.4280
	r (O <sub>9</sub> - H <sub>10</sub> )	0.9602	0.9627
	a (H <sub>3</sub> - C <sub>2</sub> - C <sub>1</sub> )	110.08	109.82
	a (H <sub>4</sub> - C <sub>2</sub> - C <sub>1</sub> )	109.89	109.57
	a (H <sub>5</sub> - C <sub>2</sub> - C <sub>1</sub> )	110.10	109.82
	a (O <sub>6</sub> - C <sub>1</sub> - C <sub>2</sub> )	115.10	115.30
	a (H <sub>7</sub> - O <sub>6</sub> - C <sub>1</sub> )	106.96	106.36
a (N - C <sub>1</sub> - O <sub>6</sub> )	124.38	124.30	
a (O <sub>9</sub> - N - C <sub>1</sub> )	108.94	108.14	
a (H <sub>10</sub> - O <sub>9</sub> - C <sub>1</sub> )	102.53	101.88	
r (C <sub>1</sub> - O <sub>2</sub> )	1.2114	1.2164	
<p><b>P8N</b></p>	r (C <sub>1</sub> - O <sub>3</sub> )	1.3660	1.3649
	r (O <sub>3</sub> - H <sub>4</sub> )	0.9652	0.9668
	r (N - C <sub>1</sub> )	1.3525	1.3513
	r (N - H <sub>6</sub> )	1.0046	1.0046
	r (N - C <sub>7</sub> )	1.4479	1.4459
	r (C <sub>7</sub> - H <sub>9</sub> )	1.0864	1.0860
	r (C <sub>7</sub> - H <sub>10</sub> )	1.0914	1.0889
	a (O <sub>3</sub> - C <sub>1</sub> - C <sub>2</sub> )	122.99	123.30
	a (N - O <sub>3</sub> - C <sub>2</sub> )	105.56	104.77
	r (C <sub>1</sub> - O <sub>3</sub> )	1.3660	1.3649
	a (O <sub>5</sub> - C <sub>2</sub> - O <sub>3</sub> )	126.60	126.53
	a (H <sub>6</sub> - C <sub>1</sub> - C <sub>2</sub> )	117.31	117.26
	a (H <sub>7</sub> - C <sub>1</sub> - C <sub>2</sub> )	122.63	121.03
a (H <sub>8</sub> - C <sub>1</sub> - C <sub>2</sub> )	107.84	107.48	
a (H <sub>9</sub> - N - O <sub>3</sub> )	111.22	111.05	
a (H <sub>10</sub> - N - O <sub>3</sub> )	111.27	111.10	

**Table A4.** Harmonic ( $\omega$ ), scaled harmonic ( $\omega_{sc}$ ), anharmonic ( $\nu$ ) vibrational frequencies and Infrared intensities ( $I$ ) of P1N-P8N isomers calculated at MP2/aug-cc-pVTZ level of theory. Frequencies are given in  $\text{cm}^{-1}$  and intensities in  $\text{km}\cdot\text{mol}^{-1}$ .

Isomer	Mode	$\nu / \text{cm}^{-1}$	$I / \text{km}\cdot\text{mol}^{-1}$	$\omega / \text{cm}^{-1}$	$\omega_{sc} / \text{cm}^{-1}$	$I / \text{km}\cdot\text{mol}^{-1}$	Vibration
<b>P1N</b>	4	484	17	414	395	78	
	5	251	149	429	409	158	
	6	425	46	490	467	51	
	12	1072	42	1099	1047	56	N-O Stretch
	15	1387	146	1428	1361	185	
	19	1702	151	1733	1651	203	C=O Stretch
	20	2982	3	3088	2943	3	Symmetric C-H Stretch ( $\text{CH}_3$ )
	23	3322	70	3536	3370	68	O-H Stretch
24	3468	62	3634	3463	77	N-H Stretch	
<b>P2N</b>	2	115	50	163	156	56	
	8	846	34	866	825	35	
	13	1233	239	1282	1222	307	
	15	1369	37	1409	1343	40	$\text{CH}_3$ Bend
	18	1572	44	1627	1551	64	$\text{NH}_2$ Bend
	19	1761	190	1790	1706	204	C=O Stretch
	20	2951	1	3098	2953	1	Symmetric C-H Stretch ( $\text{CH}_3$ )
	23	3242	19	3407	3247	22	Symmetric N-H Stretch ( $\text{NH}_2$ )
24	3360	13	3543	3376	22	Asymmetric N-H Stretch	
<b>P5N</b>	2	191	104825	105	100	191	$\text{NH}_2$ Wagging
	5	248	34	450	428	68	
	12	1134	33	1137	1084	125	C-O Stretch
	15	1322	189	1363	1299	235	
	18	1598	56	1626	1550	83	$\text{NH}_2$ Bend
	19	1744	198	1767	1684	310	C=O Stretch
	20	2906	12	3050	2907	41	Symmetric C-H Stretch ( $\text{CH}_2$ )
	21	2944	17	3091	2946	14	Asymmetric C-H Stretch
	22	3434	32	3618	3448	64	Symmetric N-H Stretch ( $\text{NH}_2$ )
	23	3492	90	3667	3495	108	O-H Stretch
24	3547	42	2764	2634	65	Asymmetric N-H Stretch	
<b>P6N</b>	4	179	299	341	325	238	$\text{NH}_2$ Wagging
	11	1094	41	1136	1083	100	
	14	1327	303	1369	1305	390	
	15	1451	29	1489	1419	30	$\text{CH}_3$ Bend
	18	1572	94	1623	1547	120	$\text{NH}_2$ Bend
	19	1780	264	1811	1726	413	C=O Stretch
	20	2996	21	3094	2949	29	Symmetric C-H Stretch ( $\text{CH}_3$ )
	23	3498	50	3633	3462	55	Symmetric N-H Stretch ( $\text{NH}_2$ )
24	3626	68	3771	3594	69	Asymmetric N-H Stretch	
<b>P7N</b>	2	83	121	244	232	109	O-H Wagging
	7	619	25	636	607	37	
	10	925	34	937	893	69	
	13	1166	43	1233	1175	101,5	
	14	1275	102	1321	1259	168,6	
	19	1713	81	1729	1648	93	C=N Stretch
	20	2958	6	3087	2942	8	Symmetric C-H Stretch ( $\text{CH}_3$ )
23	3659	60	3706	3532	87	$\text{O}_6\text{-H}_7$ Stretch	

	24	3670	96	3849	3668	119	O <sub>9</sub> -H <sub>10</sub> Stretch
<b>P8N</b>	4	4055	-	439	418	168	
	10	1985	-	1126	1073	156	C-O Stretch
	13	1242	-	1213	1156	210	
	15	1423	-	1461	1392	21	CH <sub>3</sub> Bend
	18	1532	-	1590	1515	195	
	19	1779	-	1817	1731	471	C=O Stretch
	20	2986	-	3078	2933	48	Symmetric C-H Stretch (CH <sub>3</sub> )
	23	3514	-	3680	3507	58	N-H Stretch
	24	3615	-	3800	3621	106	O-H Stretch

### A.2 Complementary tables of acetohydroxamic acid

Complementary tables of the microwave study of acetohydroxamic acid. In Table A.2.1 we provide a complementary theoretical benchmark for all of the plausible structures in the conformational space of AHA. In Tables A.2.2–A.2.3 we list all the measured lines for the *A* torsional state of *Z*- and *E*-AHA

**Table A2.1** Theoretical spectroscopic parameters for the conformers of Acetohydroxamic acid.

<b>Z-amide</b>	<b>B3LYP</b> <sup>[a]</sup> 6-311++G(d,p)	<b>B3LYP</b> aug-cc-pVTZ	<b>MP2</b> 6-311++G(d,p)	<b>MP2</b> aug-cc-pVTZ	<b>B2PLYPD</b> 6-311++G(d,p)	<b>B2PLYPD</b> aug-cc-pVTZ
<i>A</i> <sup>[b]</sup>	10598.73	10650.27	10514.08	10599.33	10574.78	10634.12
<i>B</i>	4059.36	4080.92	4099.18	4139.48	4088.87	4114.34
<i>C</i>	999.19	3014.40	3019.00	3045.13	3015.74	3032.81
$ \mu_a / \mu_b / \mu_c $ <sup>[c]</sup>	1.6 / 2.9 / 0.7	1.5 / 2.9 / 0.7	1.8 / 2.6 / 0.9	1.7 / 2.7 / 0.8	1.8 / 2.8 / 0.8	1.6 / 2.8 / 0.7
<b>E-amide</b>	<b>B3LYP</b> 6-311++G(d,p)	<b>B3LYP</b> aug-cc-pVTZ	<b>MP2</b> 6-311++G(d,p)	<b>MP2</b> aug-cc-pVTZ	<b>B2PLYPD</b> 6-311++G(d,p)	<b>B2PLYPD</b> aug-cc-pVTZ
<i>A</i>	8950.26	8989.73	8922.46	8997.63	8941.77	8987.69
<i>B</i>	4295.01	4304.31	4317.89	4339.90	4308.18	4319.30
<i>C</i>	3009.84	3016.66	3026.60	3042.38	3019.25	3027.52
$ \mu_a / \mu_b / \mu_c $	2.9 / 1.3 / 1.1	2.9 / 1.3 / 0.9	2.7 / 1.1 / 0.8	2.8 / 1.2 / 0.8	2.9 / 1.2 / 0.9	2.9 / 1.2 / 0.8
<b>Z-imide</b>	<b>B3LYP</b> 6-311++G(d,p)	<b>B3LYP</b> aug-cc-pVTZ	<b>MP2</b> 6-311++G(d,p)	<b>MP2</b> aug-cc-pVTZ	<b>B2PLYPD</b> 6-311++G(d,p)	<b>B2PLYPD</b> aug-cc-pVTZ
<i>A</i>	10426.06	10446.96	10384.58	10435.74	10435.74	10472.94
<i>B</i>	4152.48	4160.32	4317.89	4195.28	4166.30	4178.05
<i>C</i>	3025.66	3031.24	3026.60	3050.76	3033.82	3042.78
$ \mu_a / \mu_b / \mu_c $	0.2 / 0.2 / 0.0	0.1 / 0.1 / 0	0.4 / 0.3 / 0.0	0.2 / 0.2 / 0.0	0.2 / 0.2 / 0.0	0.1 / 0.2 / 0.0
<b>E-imide</b>	<b>B3LYP</b> 6-311++G(d,p)	<b>B3LYP</b> aug-cc-pVTZ	<b>MP2</b> 6-311++G(d,p)	<b>MP2</b> aug-cc-pVTZ	<b>B2PLYPD</b> 6-311++G(d,p)	<b>B2PLYPD</b> aug-cc-pVTZ
<i>A</i>	9263.14	9165.23	9118.09	9131.47	9131.47	9171.30
<i>B</i>	4235.62	4266.1	4280.51	4277.77	4277.77	4284.49
<i>C</i>	2960.71	2964.3	2966.94	2966.77	2966.77	2973.79
$ \mu_a / \mu_b / \mu_c $	0.1 / 0.1 / 0.0	0.1 / 0.1 / 0.0	0.4 / 0.2 / 0.0	0.3 / 0.2 / 0.0	0.3 / 0.2 / 0.0	0.2 / 0.1 / 0.0

[a] Theoretical calculations at B3LYP, MP2 and B2PLYPD levels of theory, respectively [b] *A*, *B*, and *C* represent the rotational constants (in MHz); [c]  $\mu_a$ ,  $\mu_b$ , and  $\mu_c$  are the electric dipole moment components (in D).

**Table A2.2** Observed frequencies and residuals (in MHz) for the rotational transitions of the A state of Z-AHA.

$J'$	$K_a'$	$K_c'$	$F'$	$J''$	$K_a''$	$K_c''$	$F''$	$\nu_{obs}$	$\nu_{calc}$	$\nu_{obs} - \nu_{calc}$
1	0	1	0	0	0	0	1	7128.455	7128.452	0.003
1	0	1	2	0	0	0	1	7130.290	7130.291	-0.001
1	0	1	1	0	0	0	1	7131.519	7131.517	0.002
2	0	2	1	1	1	1	1	7599.650	7599.643	0.007
2	0	2	1	1	1	1	0	7600.465	7600.464	0.001
2	0	2	3	1	1	1	2	7600.778	7600.780	-0.002
2	0	2	2	1	1	1	1	7601.910	7601.907	0.003
2	0	2	2	1	1	1	2	7602.236	7602.235	0.001
1	1	0	1	1	0	1	1	7616.158	7616.161	-0.003
1	1	0	1	1	0	1	2	7617.385	7617.387	-0.001
1	1	0	2	1	0	1	1	7617.715	7617.715	-0.000
1	1	0	2	1	0	1	2	7618.949	7618.942	0.007
1	1	0	1	1	0	1	0	7619.227	7619.226	0.001
2	1	1	2	2	0	2	2	8822.689	8822.690	-0.001
2	1	1	3	2	0	2	2	8823.045	8823.042	0.002
2	1	1	1	2	0	2	2	8823.241	8823.238	0.003
2	1	1	2	2	0	2	3	8824.145	8824.146	-0.000
2	1	1	3	2	0	2	3	8824.497	8824.498	-0.001
2	1	1	2	2	0	2	1	8824.954	8824.954	-0.000
2	1	1	1	2	0	2	1	8825.504	8825.502	0.002
3	1	2	3	3	0	3	3	10847.624	10847.625	-0.001
3	1	2	4	3	0	3	4	10849.317	10849.318	-0.001
3	1	2	2	3	0	3	2	10849.913	10849.910	0.002
2	1	2	1	1	1	1	1	13177.289	13177.291	-0.002
2	1	2	1	1	1	1	0	13178.114	13178.113	0.001
2	1	2	3	1	1	1	2	13178.542	13178.545	-0.003
2	1	2	2	1	1	1	1	13179.886	13179.883	0.003
2	1	2	2	1	1	1	2	13180.211	13180.211	-0.000
1	1	1	0	0	0	0	1	13666.325	13666.328	-0.003
1	1	1	2	0	0	0	1	13666.824	13666.821	0.003
1	1	1	1	0	0	0	1	13667.150	13667.149	0.001
2	0	2	3	1	0	1	2	14137.313	14137.310	0.003
2	0	2	2	1	0	1	1	14137.538	14137.539	-0.001
2	0	2	1	1	0	1	0	14138.335	14138.341	-0.006
2	0	2	2	1	0	1	2	14138.764	14138.765	-0.001
3	0	3	2	2	1	2	2	15327.428	15327.440	-0.012
3	0	3	4	2	1	2	3	15329.709	15329.707	0.002
3	0	3	3	2	1	2	2	15329.760	15329.759	0.001
3	0	3	2	2	1	2	1	15330.037	15330.031	0.006
3	0	3	3	2	1	2	3	15331.425	15331.424	0.001
2	1	1	1	1	1	0	0	15340.735	15340.730	0.005
2	1	1	2	1	1	0	2	15342.514	15342.514	0.000
2	1	1	3	1	1	0	2	15342.866	15342.866	0.000
2	1	1	2	1	1	0	1	15344.070	15344.069	0.001
2	1	1	1	1	1	0	1	15344.600	15344.617	-0.017
2	1	2	3	1	0	1	2	19715.100	19715.076	0.024
2	1	2	2	1	0	1	2	19716.740	19716.742	-0.002
2	1	2	2	1	0	1	1	19715.540	19715.515	0.025
2	2	0	2	2	1	1	2	19731.760	19731.772	-0.012
2	2	0	3	2	1	1	3	19732.880	19732.875	0.005
2	2	0	1	2	1	1	1	19733.490	19733.488	0.002
3	0	3	4	2	0	2	3	20907.480	20907.472	0.008
3	0	3	3	2	0	2	2	20907.740	20907.735	0.005

3	1	2	4	2	1	1	3	22932.310	22932.292	0.018
3	1	2	3	2	1	1	2	22932.680	22932.670	0.010
3	1	3	4	2	0	2	3	25272.620	25272.612	0.008
3	1	3	2	2	0	2	1	25272.680	25272.681	-0.001
3	1	3	3	2	0	2	3	25274.740	25274.723	0.017
3	1	3	3	2	0	2	2	25273.280	25273.268	0.012
4	0	4	5	3	0	3	4	27372.420	27372.435	-0.015
2	2	1	2	1	1	0	2	34950.870	34950.851	0.019
2	2	1	3	1	1	0	2	34952.160	34952.165	-0.002
2	2	1	2	1	1	0	1	34952.400	34952.406	-0.006
2	2	0	2	1	1	1	1	36156.350	36156.369	-0.019
2	2	0	3	1	1	1	2	36158.140	36158.153	-0.013

Note.  $\nu_{obs}$  is the observed frequency and  $\nu_{obs} - \nu_{calc}$  is the residual.

**Table A2.3** Observed frequencies and residuals (in MHz) for the rotational transitions of the A state of E-AHA.

J'	K <sub>a</sub> '	K <sub>c</sub> '	F'	J''	K <sub>a</sub> ''	K <sub>c</sub> ''	F''	$\nu_{obs}$	$\nu_{calc}$	$\nu_{obs} - \nu_{calc}$
1	0	1	0	0	0	0	1	7341.489	7341.481	0.008
1	0	1	2	0	0	0	1	7343.274	7343.269	0.005
1	0	1	1	0	0	0	1	7344.469	7344.461	0.008
1	1	1	0	0	0	0	1	11973.404	11973.396	0.008
1	1	1	2	0	0	0	1	11974.209	11974.211	-0.002
1	1	1	1	0	0	0	1	11974.749	11974.755	-0.006
2	1	2	1	1	1	1	1	13387.141	13387.137	0.004
2	1	2	2	1	1	1	2	13390.558	13390.573	-0.015
2	1	2	2	1	1	1	1	13390.046	13390.029	0.015
2	1	2	3	1	1	1	2	13388.709	13388.713	-0.004
2	0	2	1	1	0	1	1	14448.111	14448.108	0.002
2	0	2	3	1	0	1	2	14450.160	14450.157	0.003
2	0	2	2	1	0	1	1	14450.499	14450.508	-0.009
2	0	2	1	1	0	1	0	14451.078	14451.088	-0.009
2	0	2	2	1	0	1	2	14451.701	14451.700	0.001
2	1	1	3	1	1	0	2	15984.608	15984.630	-0.022
2	1	1	2	1	1	0	1	15985.778	15985.783	-0.005
2	1	1	1	1	1	0	1	15986.714	15986.689	0.025

Note.  $\nu_{obs}$  is the observed frequency and  $\nu_{obs} - \nu_{calc}$  is the residual.

### A.3 Complementary tables for the study of glycolamide

Table A3.1 shows a comparison between experimental and theoretical spectroscopic constants for glycolamide and Table A3.2 lists the theoretical frequencies of the lowest vibrational modes. Table A.0 lists the noise levels and continuum levels measured around all glycolamide transitions with a lower level energy in temperature units below 10 K covered by the Effelsberg observations between 4 and 8 GHz and between 12 and 17 GHz.

**Table A3.1** Experimental (exp) and theoretical (theory) spectroscopic parameters for *syn*- and *anti*-glycolamide.

Parameters	<i>syn</i> (exp)	<i>anti</i> (exp)	<i>syn</i> (theory)	<i>anti</i> (theory)
A <sup>(a)</sup> (MHz)	10454.2591 (20) <sup>(f)</sup>	9852.6098 (20)	10448.9	9932.8
B (MHz)	4041.14440 (88)	4125.1538 (12)	4077.1	4149.0
C (MHz)	2972.08215 (78)	2967.57875 (70)	2987.5	2981.2
$\Delta_J$ (kHz)	0.678 (74)	-	-	-
$\Delta_{JK}$ (kHz)	9.87 (90)	-	-	-
$\chi_{aa}^{(b)}$ (MHz)	2.050 (4)	1.615 (5)	2.1	1.6
$\chi_{bb}$ (MHz)	1.926 (6)	2.005 (10)	1.9	2.0
$\chi_{cc}$ (MHz)	-3.976 (6)	-3.621 (10)	-4.0	-3.6
N <sup>(c)</sup>	55	21	-	-
$\sigma^{(d)}$ (kHz)	10.8	6.9	-	-
$\Delta E^{(e)}$ (cm <sup>-1</sup> )	-	-	0	218

**Notes.** Theoretical calculations at MP2/aug-cc-pVTZ level of theory. <sup>(a)</sup> A, B, and C represent the rotational constants. <sup>(b)</sup>  $\chi_{aa}$ ,  $\chi_{bb}$ ,  $\chi_{cc}$  are the diagonal elements of the 14N nuclear quadrupole coupling tensor. <sup>(c)</sup> N is the number of measured hyperfine components. <sup>(d)</sup>  $\sigma$  is the root mean square (rms) deviation of the fit. <sup>(e)</sup>  $\Delta E$  is the electronic energy relative to the global minimum, taking into account the zero-point vibrational energy (ZPE) calculated at the same level. <sup>(f)</sup> Standard error in parentheses in units of the last digit.

**Table A3.2** Theoretical frequencies of the lowest vibrational modes.

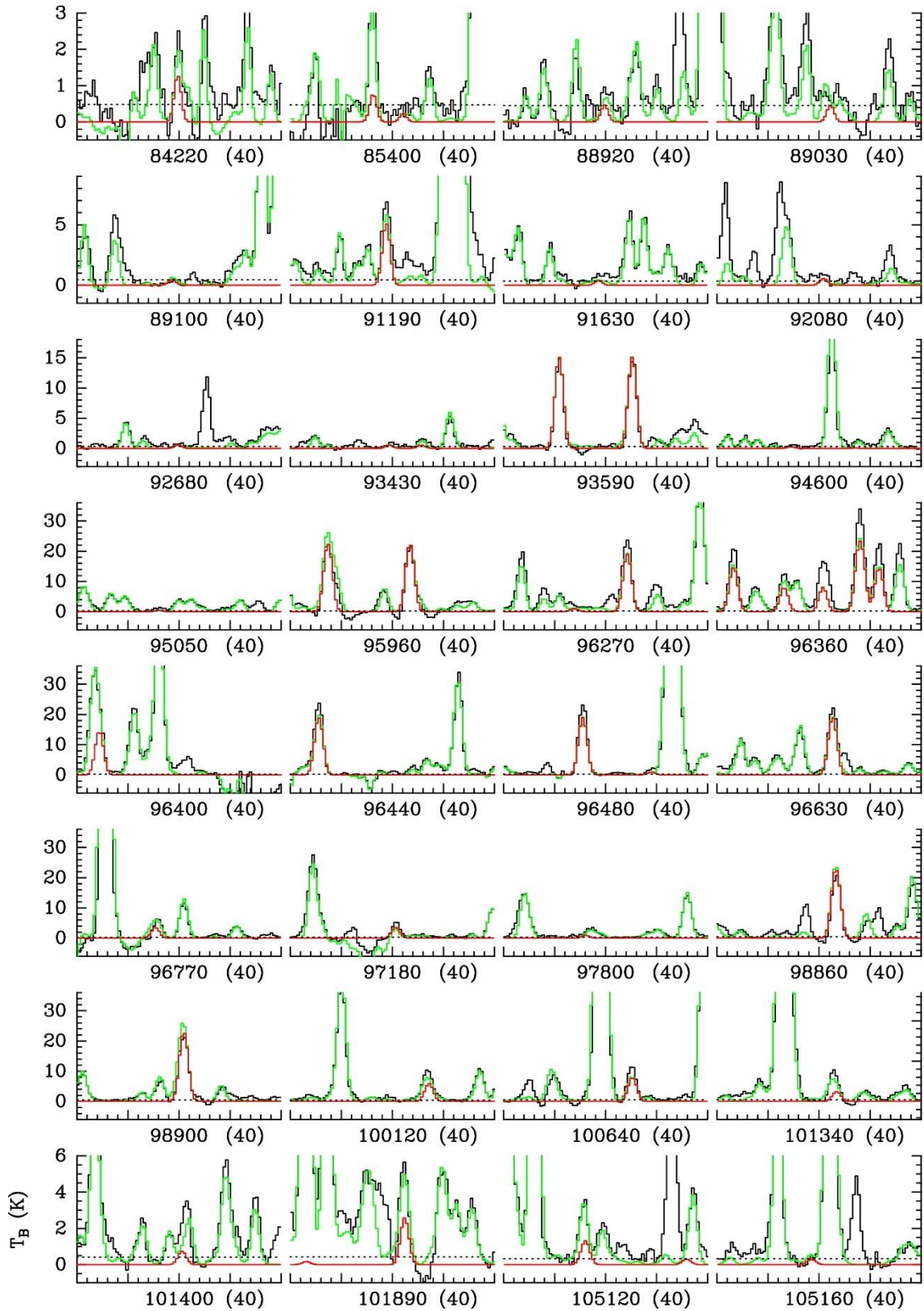
Vibrational mode	$\nu$ (cm <sup>-1</sup> )
1	78.4
2	183.6
3	293.0
4	444.5
5	447.7
6	500.9
7	634.3
8	645.3
9	827.5
10	1042.7

**Notes.**<sup>(a)</sup> Frequencies (in cm<sup>-1</sup>) calculated at the B2PLYPD3/aug-cc- pVTZ level of theory. They are necessary to estimate the vibrational contribution to the partition function in order to get a proper estimate of the total column density (upper limit) of the molecule.

#### *A.4 Complementary figures for the astronomical search of glycolamide*

Figures A.1–A.10 show the transitions of  $\text{CH}_3\text{CHO}$ ,  $\text{CH}_2(\text{OH})\text{CHO}$ , and some of their isotopologues or vibrationally excited states that are covered by the EMOCA survey and contribute significantly to the signal detected toward Sgr B2(N2). Figure A.11 shows examples of spectra obtained with Effelsberg toward Sgr B2(N) between 12 and 17 GHz. It illustrates the non-detection of glycolamide at these low frequencies.





**Fig. A.1.** Transitions of CH<sub>3</sub>CHO,  $\nu=0$  covered by our ALMA survey. The best-fit LTE synthetic spectrum of CH<sub>3</sub>CHO,  $\nu=0$  is displayed in red and overlaid on the  $\nu=0$  observed spectrum of Sgr B2(N2) shown in black. The green synthetic spectrum contains the contributions of all molecules identified in our survey so far, including the species shown in red. The central frequency and width are indicated in MHz below each panel. The y-axis is labeled in brightness temperature units (K). The dotted line indicates the  $3\sigma$  noise level.

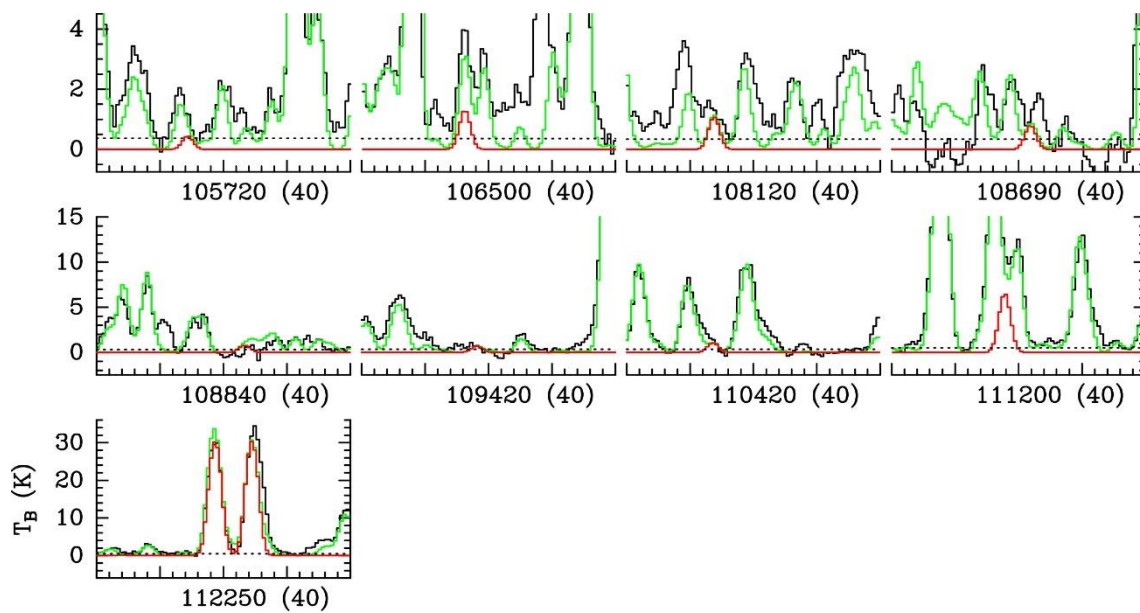
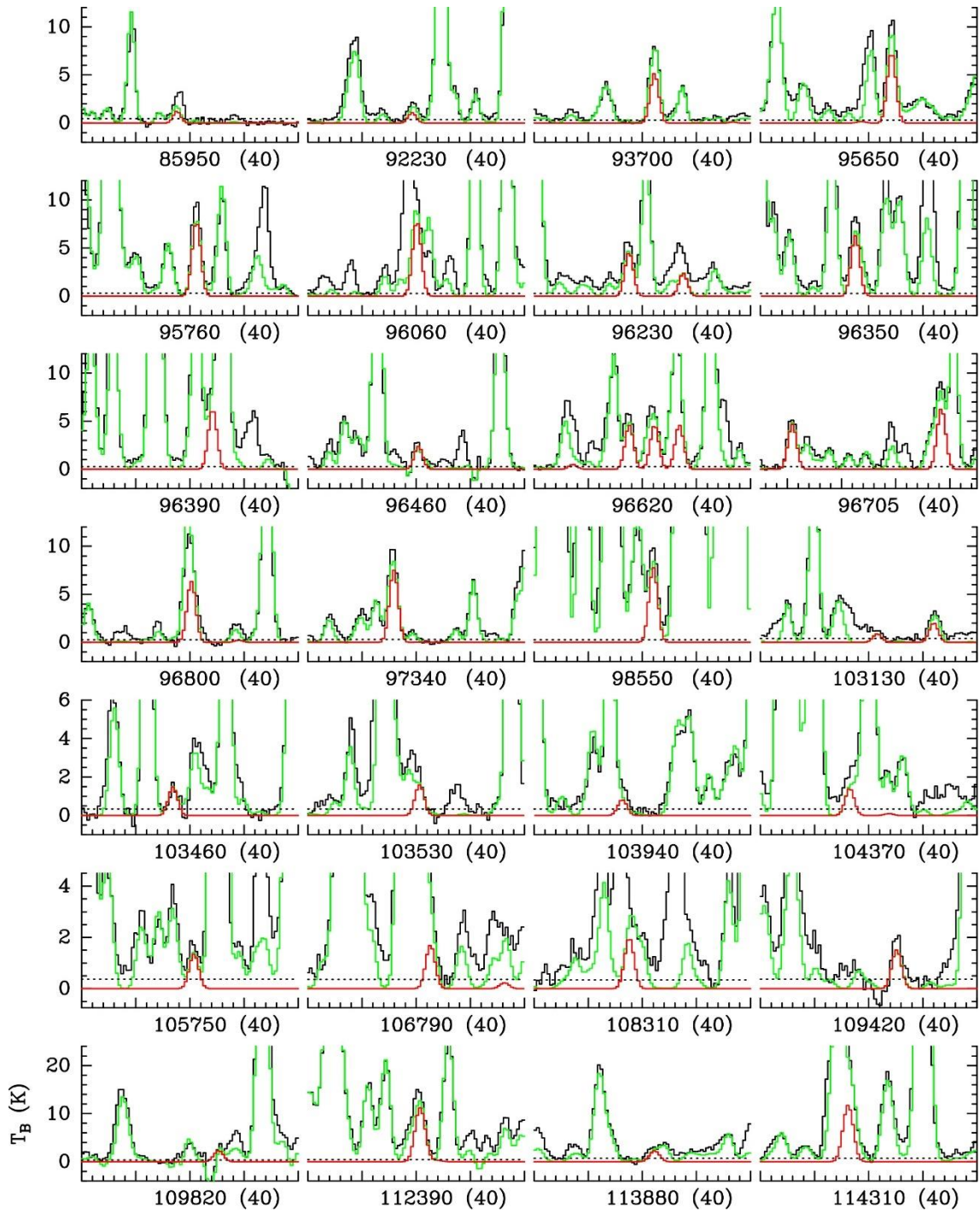
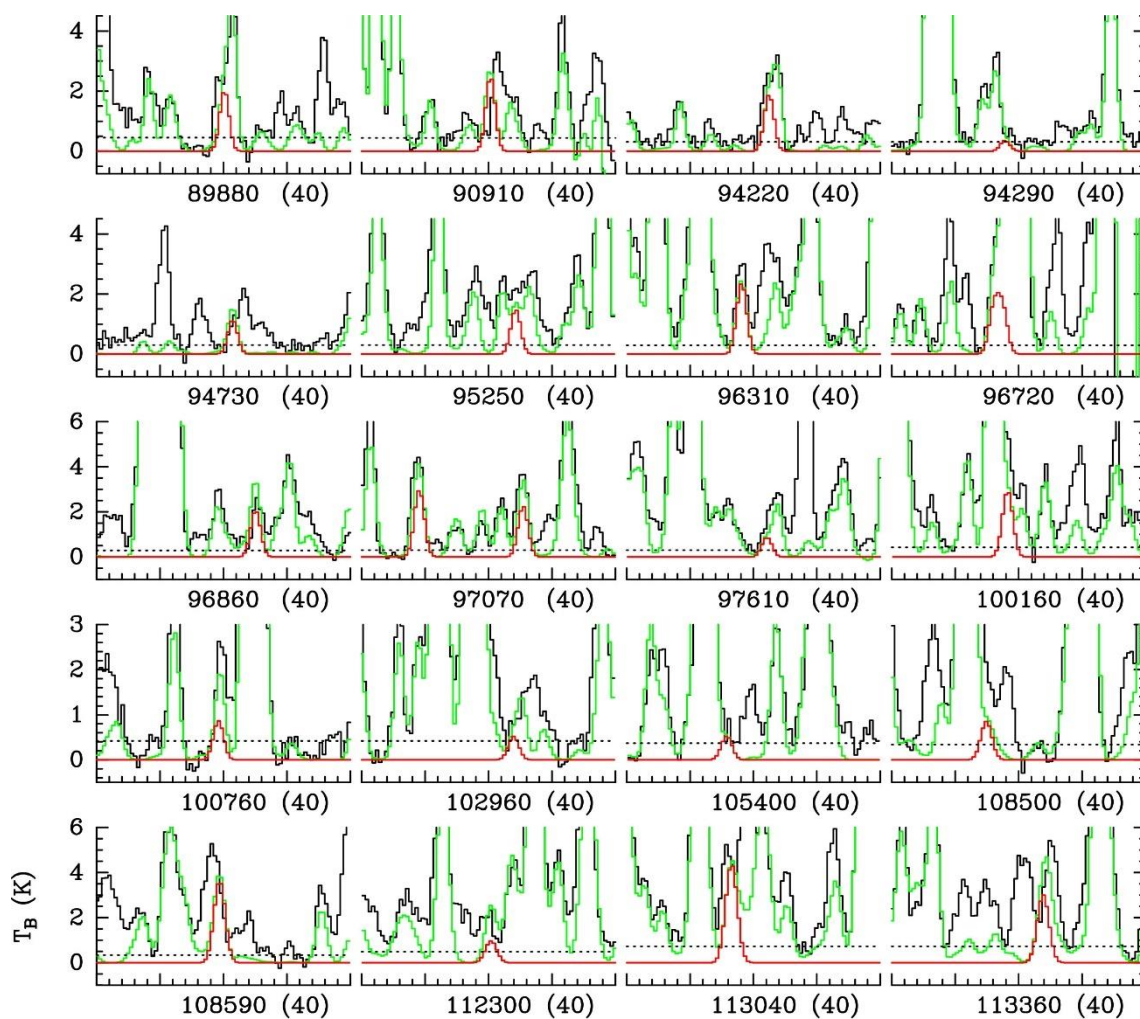


Fig. A.1. continued.



**Fig. A.2.** Same as Fig. A.1, but for  $\text{CH}_3\text{CHO}$ ,  $v_t = 1$ .



**Fig. A.3.** Same as Fig. A.1, but for CH<sub>3</sub>CHO,  $v_l = 2$

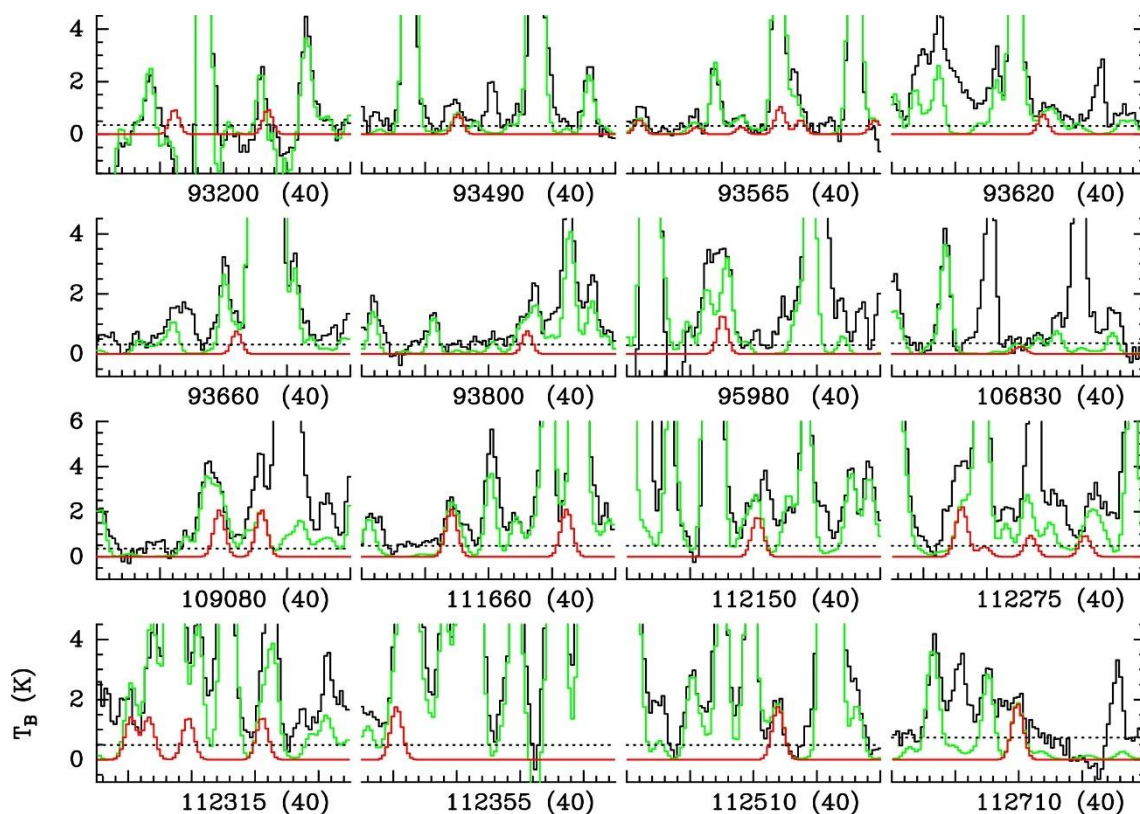


Fig. A.4. Same as Fig. A.1, but for  $^{13}\text{CH}_3\text{CHO}$ ,  $\nu = 0$ .

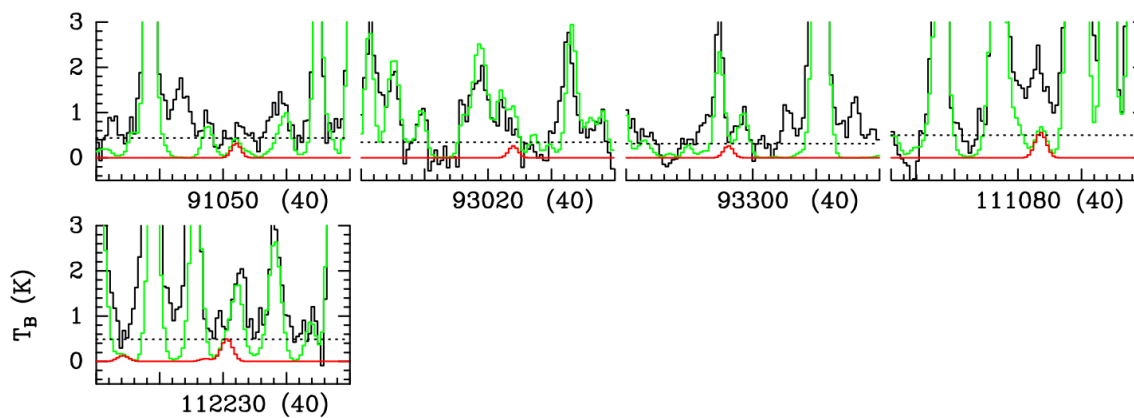
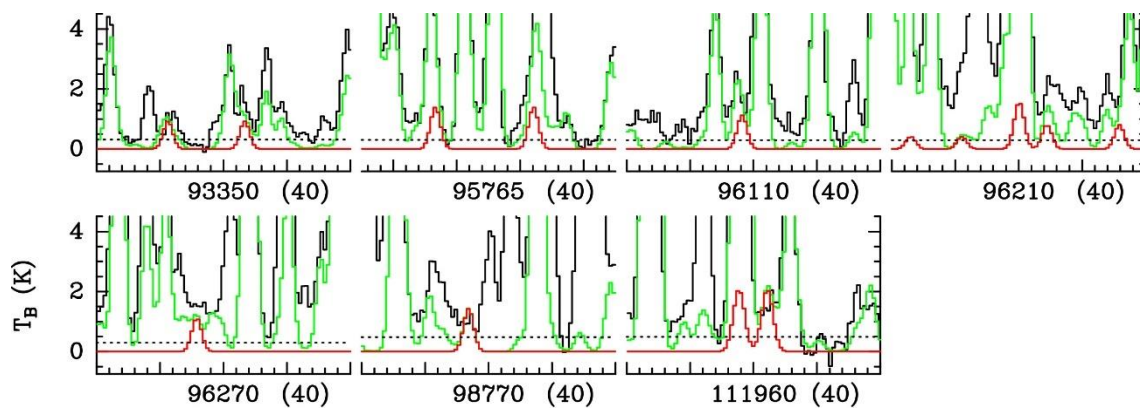
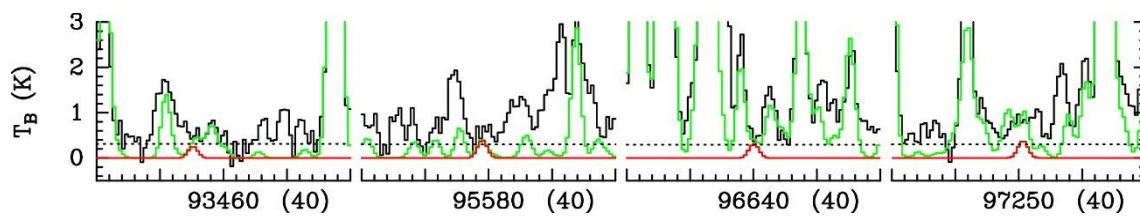


Fig. A.5. Same as Fig. A.1, but for  $^{13}\text{CH}_3\text{CHO}$ ,  $\nu = 1$ .



**Fig. A.6.** Same as Fig. A.1, but for  $\text{CH}_3^{13}\text{CHO}$ ,  $\nu = 0$ .



**Fig. A.7.** Same as Fig. A.1, but for  $\text{CH}_3^{13}\text{CHO}$ ,  $\nu = 1$

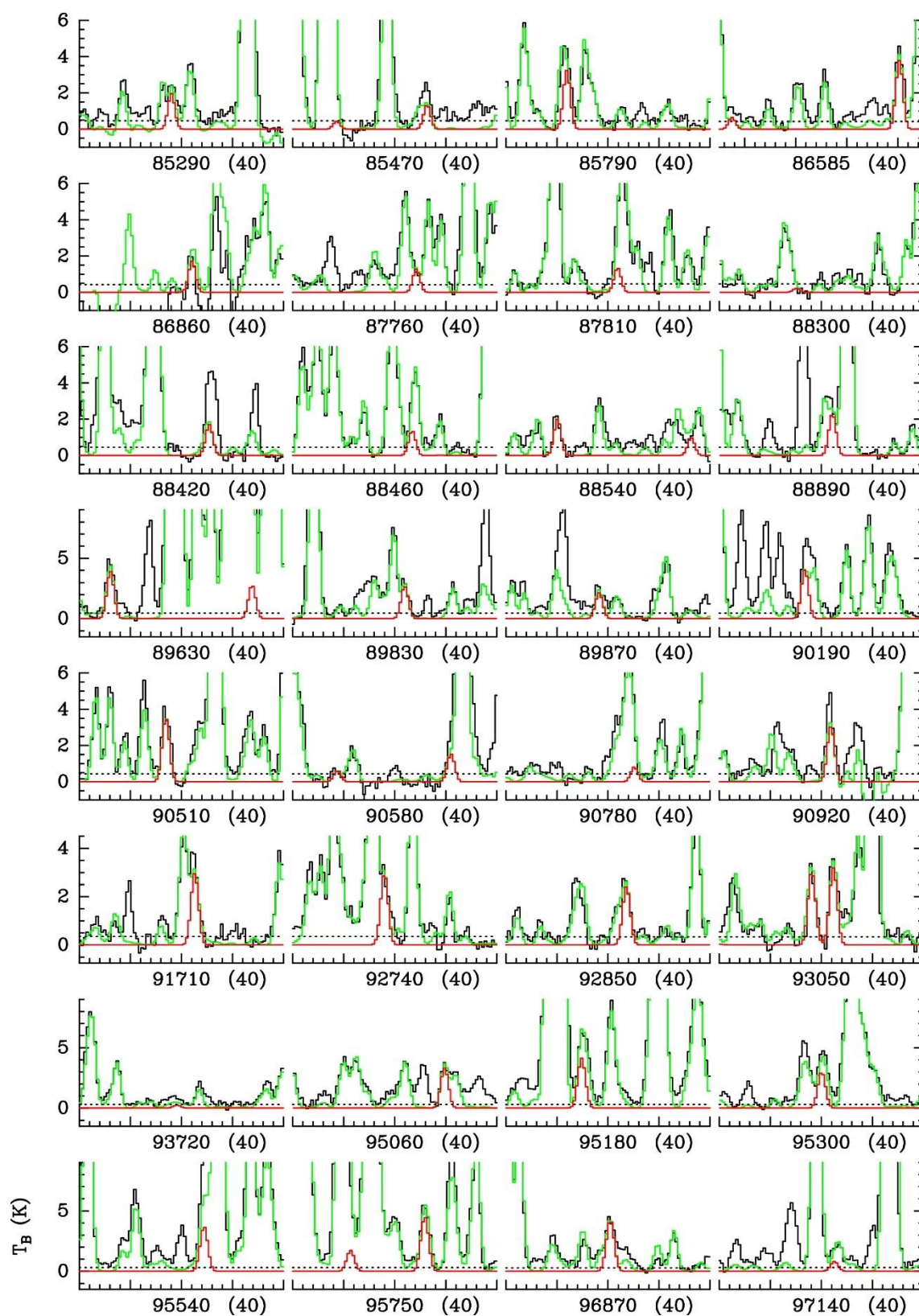


Fig. A.8. Same as Fig. A.1, but for  $\text{CH}_3(\text{OH})\text{CHO}$ ,  $\nu = 0$ .

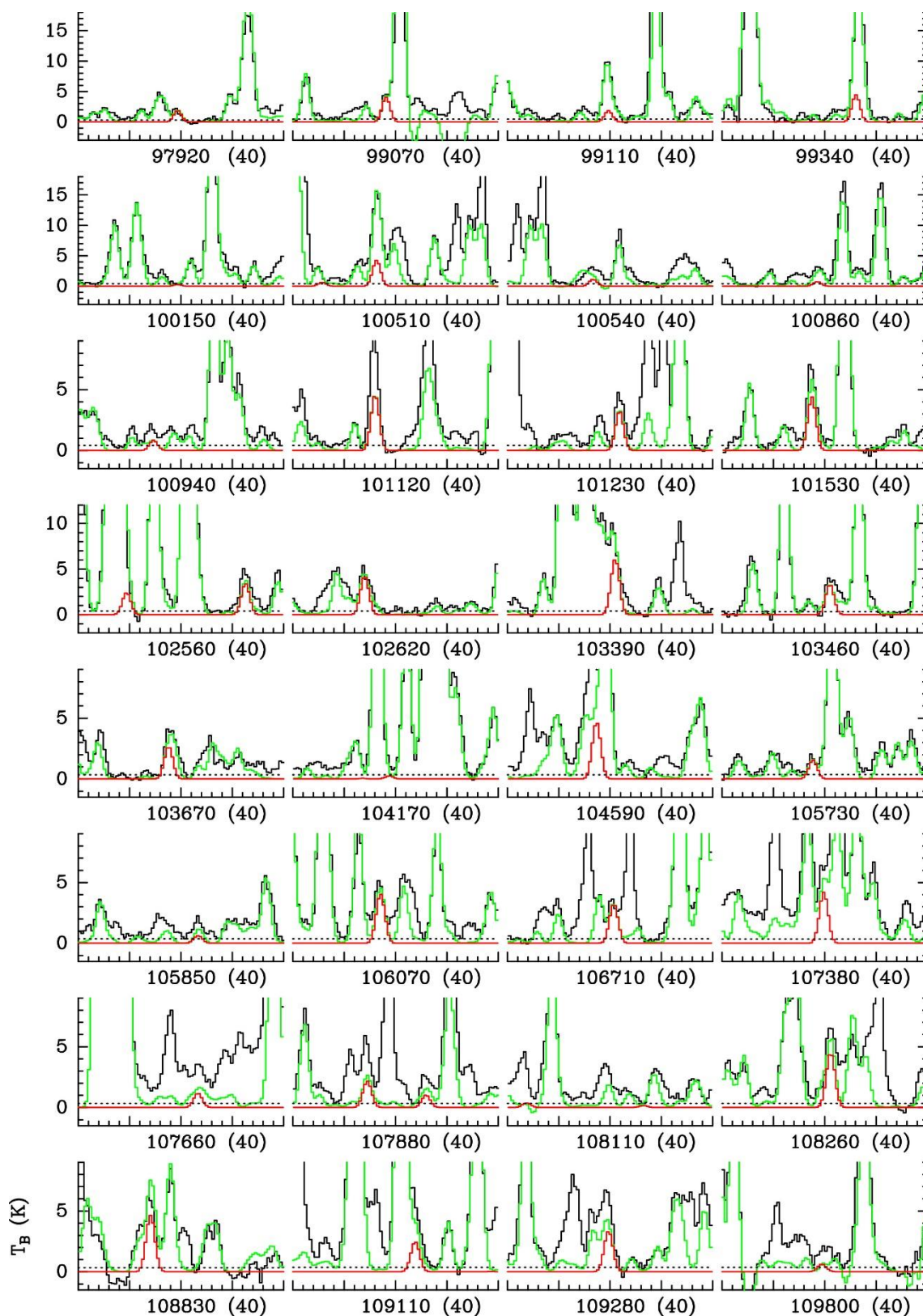


Fig. A.8. Continued.



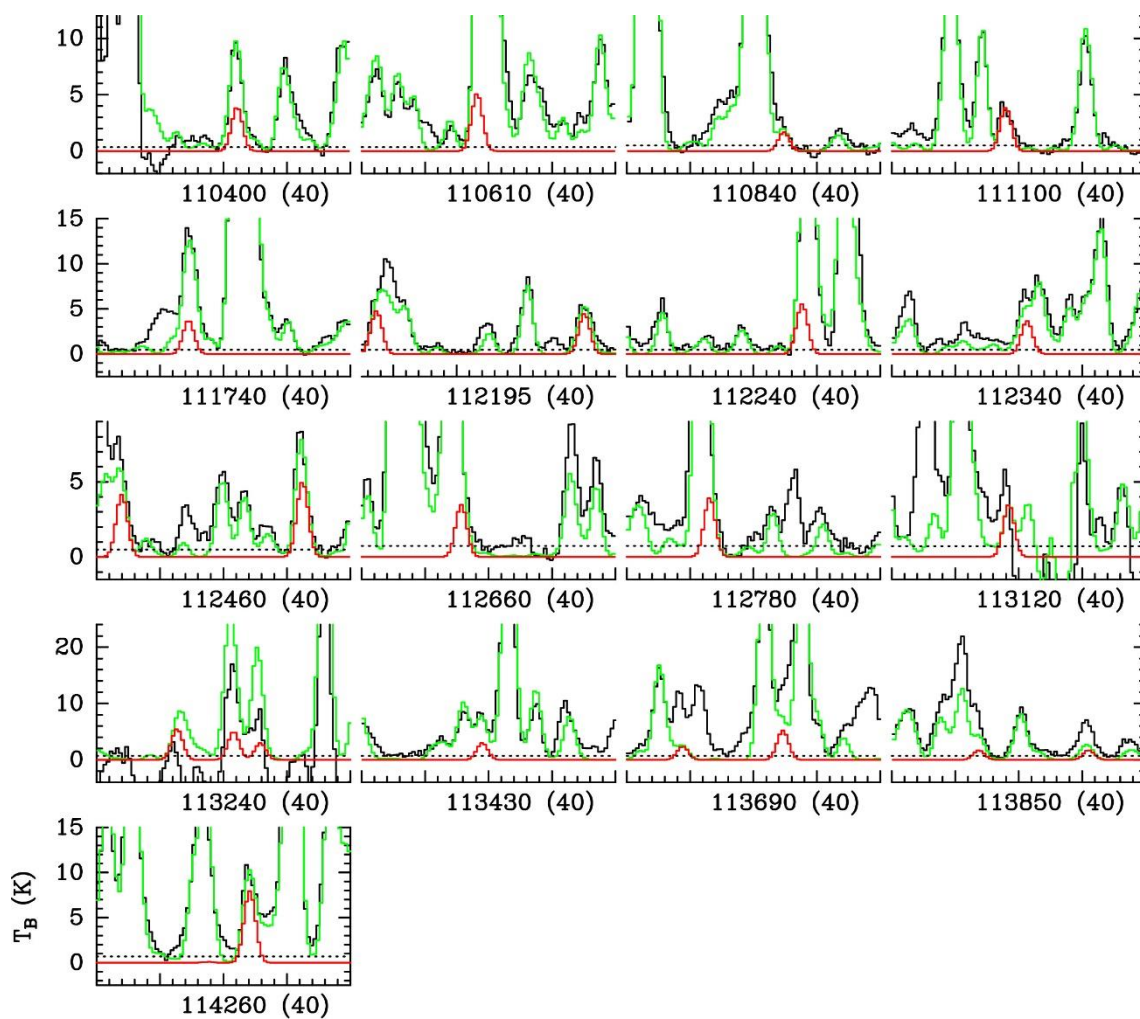


Fig. A.8. Continued.

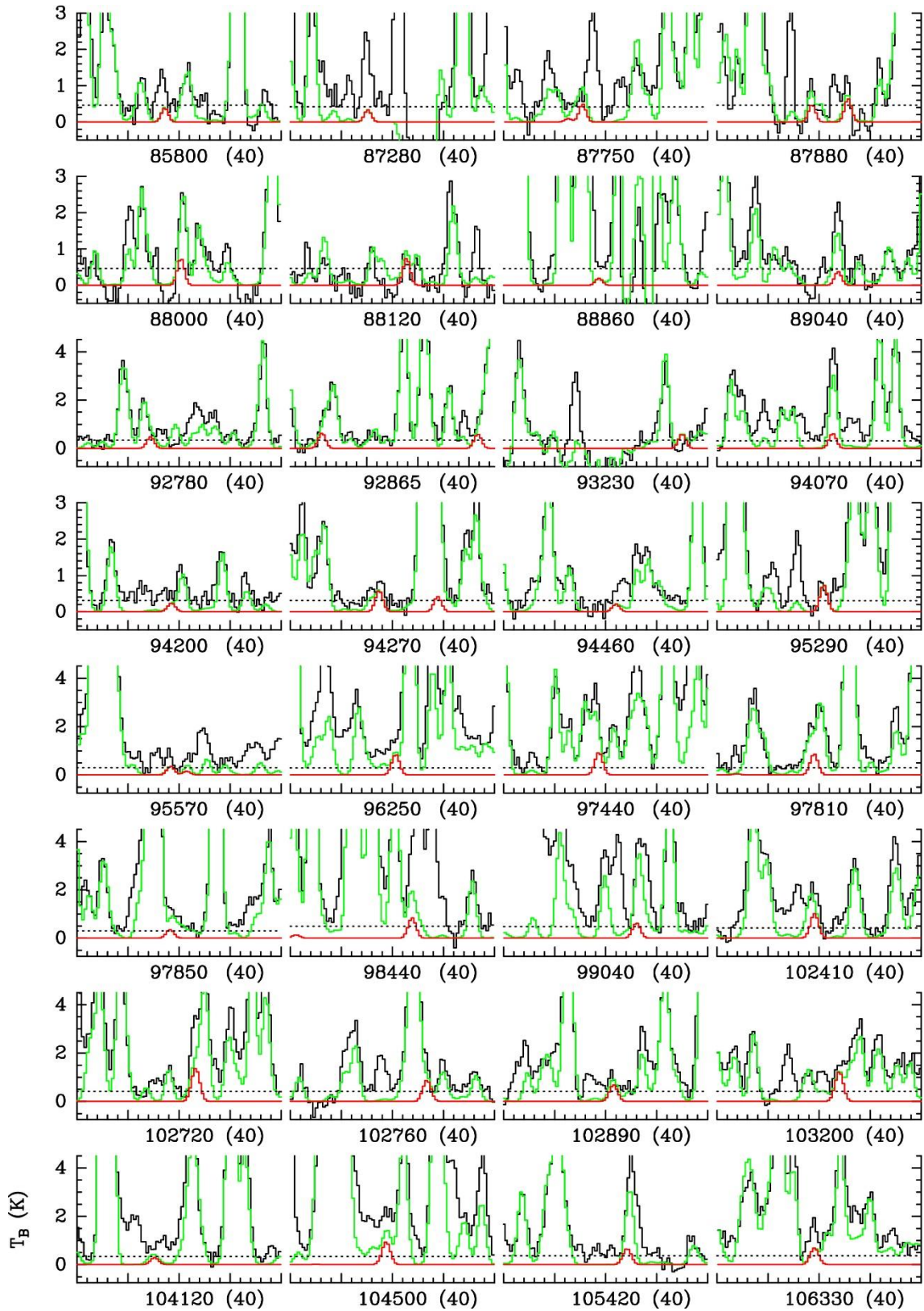


Fig. A.9. Same as Fig. A.1, but for  $\text{CH}_2(\text{OH})\text{CHO}$ ,  $\nu_7 = 1$ .

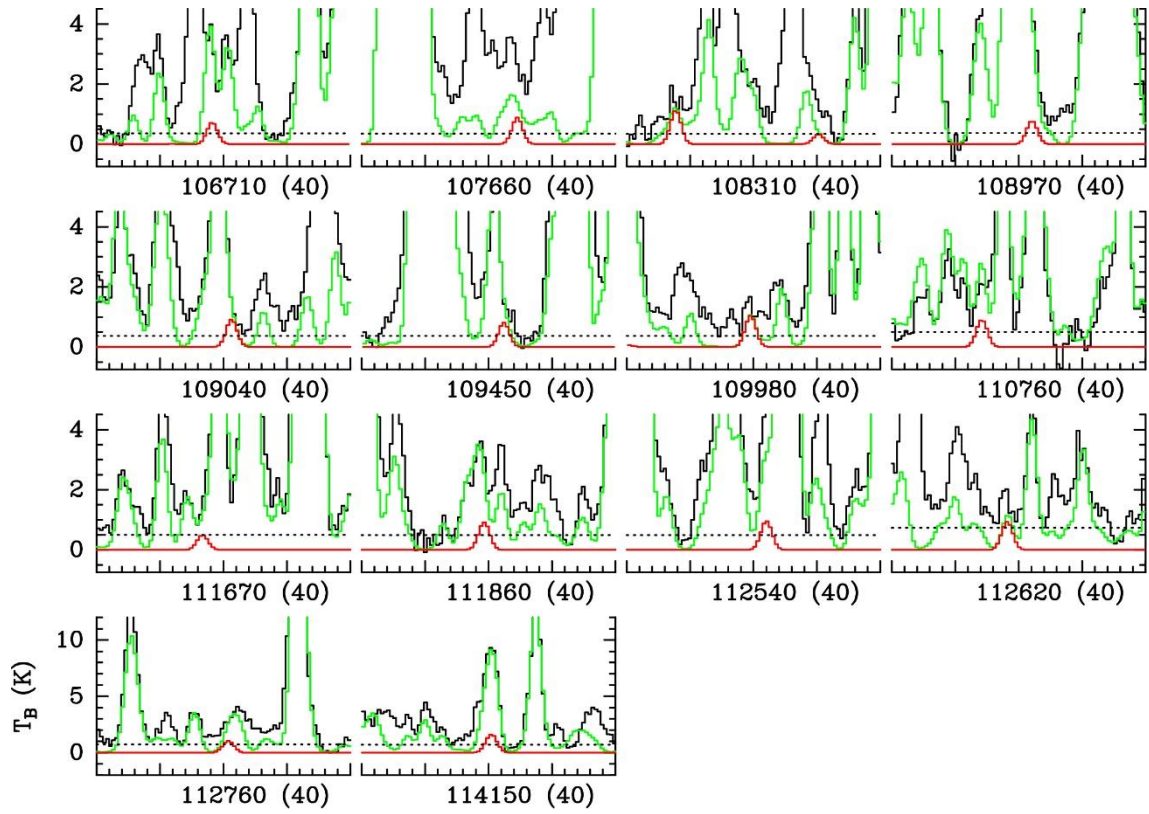


Fig. A.9. Continued.

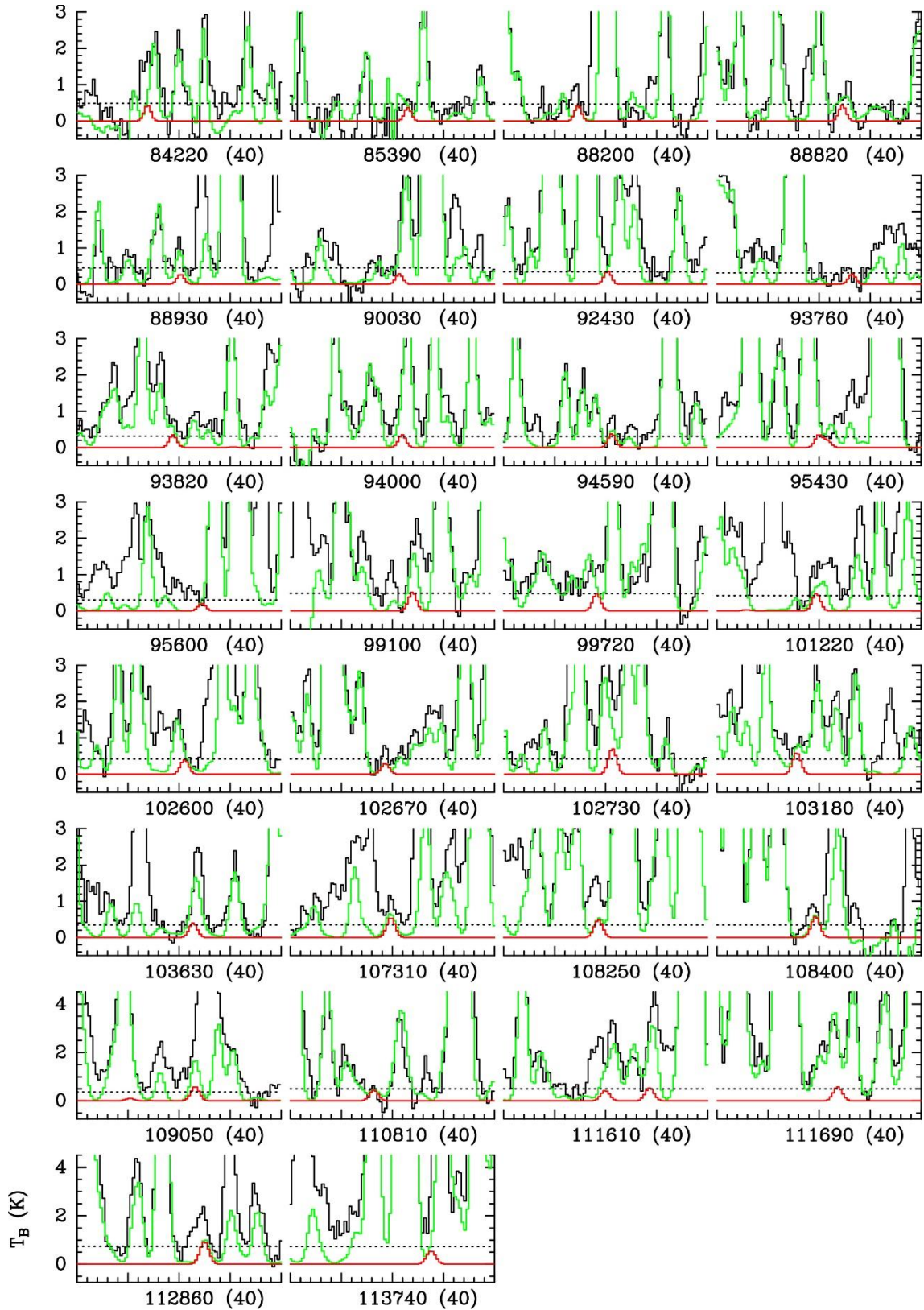


Fig. A.10. Same as Fig. A.1, but for  $\text{CH}_3(\text{OH})\text{CHO}$ ,  $v_l=2$

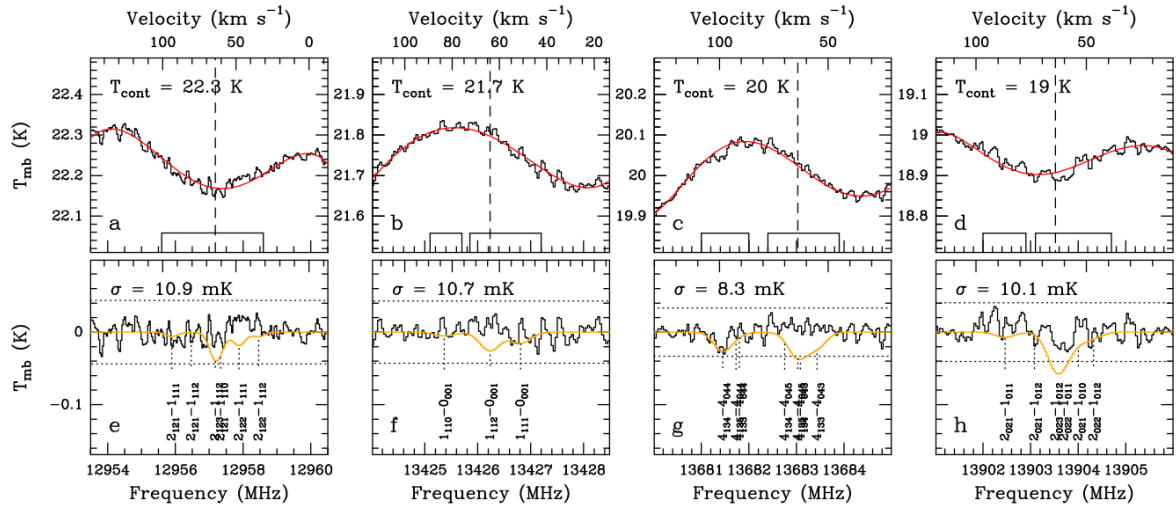
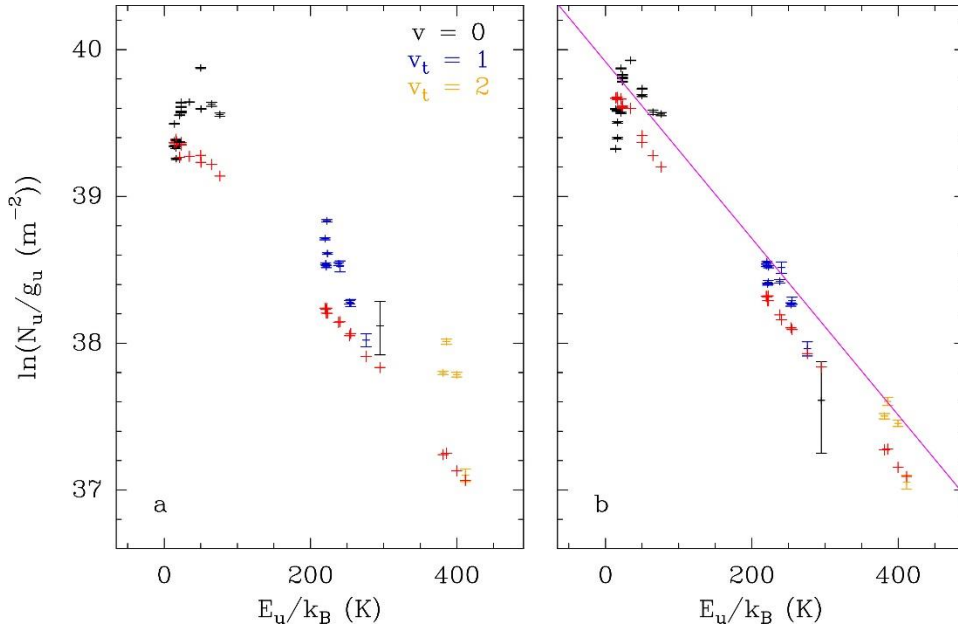


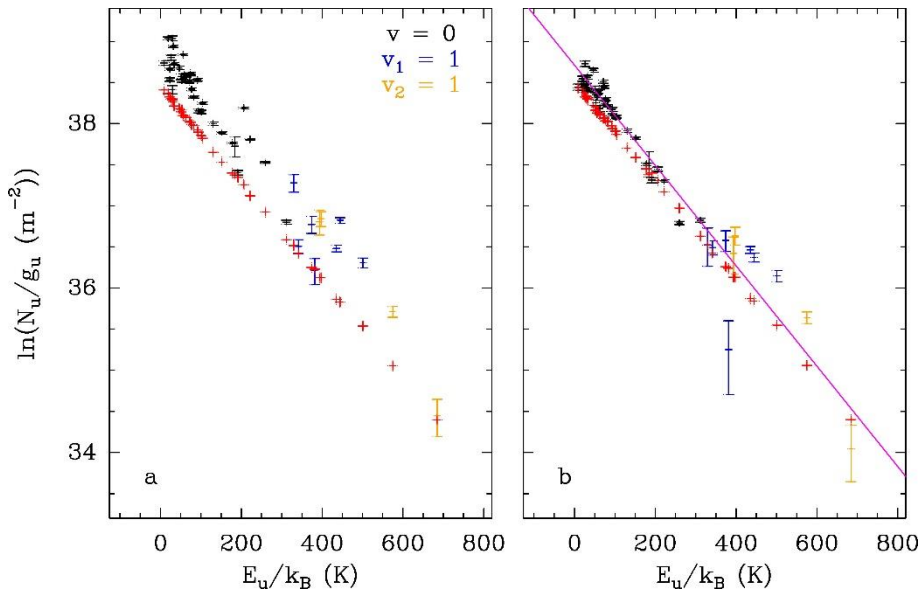
Fig. A.11. Same as Fig. 5, but for spectra obtained with Effelsberg between 12 and 17 GHz.

A.5 Complementary figures for the population diagrams of glycolamide

Figures A.12 and A.13 show the population diagrams of CH<sub>3</sub>CHO and CH<sub>2</sub>(OH)CHO toward Sgr B2(N2).



**Fig. A.12.** Population diagram of CH<sub>3</sub>CHO toward Sgr B2(N2). The observed datapoints are shown in various colors (but not red) as indicated in the upper right corner of panel a while the synthetic populations are shown in red. No correction is applied in panel a. In panel b, the optical depth correction has been applied to both the observed and synthetic populations and the contamination by all other species included in the full model has been removed from the observed datapoints. The purple line is a linear fit to the observed populations (in linear-logarithmic space).



**Fig. A.13.** Same as Fig. A.12, but for CH<sub>2</sub>(OH)CHO.

**Table A.0.** Spectral noise levels and continuum levels for glycolamide transitions with  $E_{\text{low}} < 10$  K observed with Effelsberg toward Sgr B2(N).

Frequency (MHz)	Transition								$E_l/k^a$ (K)	$A_{ul}^b$ ( $10^{-8} \text{ s}^{-1}$ )	$g_u^c$	rms <sup>d</sup> (mK)	$T_{\text{cont}}^e$ (K)	
	$J^u$	$K_a^u$	$K_c^u$	$F^u$	$J^l$	$K_a^l$	$K_c^l$	$F^l$						
4538.854	7	1	7	8	-	6	2	4	7	8.694	0.012	17	13.4	55.7
5043.746	6	1	6	5	-	5	2	3	4	6.548	0.024	11	13.2	51.0
5044.140	6	1	6	7	-	5	2	3	6	6.548	0.024	15		
5046.409	6	1	6	6	-	5	2	3	5	6.548	0.024	13		
5108.735	3	3	1	4	-	4	2	2	5	4.779	0.026	9	11.3	50.8
5550.175	6	2	5	5	-	5	3	2	4	8.092	0.062	11	6.0	46.5
5550.320	6	2	5	7	-	5	3	2	6	8.092	0.064	15		
5551.151	6	2	5	6	-	5	3	2	5	8.092	0.062	13		
6404.012	3	1	2	3	-	3	1	3	3	2.196	0.220	7	8.8	44.2
6405.972	3	1	2	4	-	3	1	3	4	2.195	0.246	9		
6406.658	3	1	2	2	-	3	1	3	2	2.195	0.233	5		
6861.504	3	3	0	3	-	4	2	3	4	4.695	0.060	7	8.3	38.3
6862.721	3	3	0	4	-	4	2	3	5	4.695	0.060	9		
6863.118	3	3	0	2	-	4	2	3	3	4.695	0.064	5		
7001.013	6	2	4	6	-	6	2	5	6	8.358	0.315	13	7.4	36.0
7002.061	6	2	4	7	-	6	2	5	7	8.358	0.324	15		
7002.238	6	2	4	5	-	6	2	5	5	8.358	0.321	11		
7012.199	1	0	1	0	-	0	0	0	1	0.000	1.370	1	6.8	36.3
7013.121	1	0	1	2	-	0	0	0	1	0.000	1.371	5		
7013.736	1	0	1	1	-	0	0	0	1	0.000	1.371	3		
7480.660	1	1	0	1	-	1	0	1	1	0.337	0.548	3	9.8	37.9
7481.275	1	1	0	1	-	1	0	1	2	0.337	0.914	3		
7481.853	1	1	0	2	-	1	0	1	1	0.337	0.548	5		
7482.198	1	1	0	1	-	1	0	1	0	0.337	0.731	3		
7482.468	1	1	0	2	-	1	0	1	2	0.337	1.646	5		
7489.391	2	0	2	1	-	1	1	1	1	0.644	0.227	3	9.6	37.6
7490.412	2	0	2	3	-	1	1	1	2	0.644	0.545	7		
7490.632	2	0	2	2	-	1	1	1	1	0.644	0.409	5		
7490.836	2	0	2	1	-	1	1	1	0	0.644	0.303	3		
7491.210	2	0	2	2	-	1	1	1	2	0.644	0.136	5		
12199.244	4	1	3	4	-	3	2	2	3	3.353	1.102	9	20.4	24.5
12199.244	4	1	3	4	-	3	2	2	4	3.353	0.074	9		
12199.522	4	1	3	5	-	3	2	2	4	3.353	1.175	11		
12199.593	4	1	3	3	-	3	2	2	3	3.353	0.094	7		
12199.593	4	1	3	3	-	3	2	2	2	3.353	1.079	7		
12786.796	4	4	1	4	-	5	3	2	5	8.092	0.252	9	8.1	23.6
12787.652	4	4	1	5	-	5	3	2	6	8.092	0.254	11		
12787.880	4	4	1	3	-	5	3	2	4	8.092	0.263	7		
12795.512	6	2	4	6	-	5	3	3	5	8.080	0.940	13	9.8	23.3
12795.816	6	2	4	7	-	5	3	3	6	8.080	0.967	15		
12795.865	6	2	4	5	-	5	3	3	4	8.080	0.935	11		
12955.894	2	1	2	1	-	1	1	1	1	0.644	3.241	3	10.9	22.3
12956.471	2	1	2	1	-	1	1	1	2	0.644	0.216	3		
12957.182	2	1	2	3	-	1	1	1	2	0.644	7.779	7		
12957.338	2	1	2	1	-	1	1	1	0	0.644	4.321	3		
12957.882	2	1	2	2	-	1	1	1	1	0.644	5.834	5		
12958.460	2	1	2	2	-	1	1	1	2	0.644	1.945	5		
13030.677	4	4	0	4	-	5	3	3	5	8.080	0.266	9	12.2	23.3
13031.620	4	4	0	5	-	5	3	3	6	8.080	0.268	11		
13031.866	4	4	0	3	-	5	3	3	4	8.080	0.278	7		
13425.365	1	1	1	0	-	0	0	0	1	0.000	8.450	1	10.7	21.7
13426.232	1	1	1	2	-	0	0	0	1	0.000	8.453	5		
13426.810	1	1	1	1	-	0	0	0	1	0.000	8.452	3		
13681.452	4	1	3	4	-	4	0	4	4	3.282	8.173	9	8.3	20.0
13681.730	4	1	3	5	-	4	0	4	4	3.282	0.363	11		
13681.801	4	1	3	3	-	4	0	4	4	3.282	0.566	7		
13682.749	4	1	3	4	-	4	0	4	5	3.282	0.443	9		
13683.027	4	1	3	5	-	4	0	4	5	3.282	8.697	11		

**Table A.0.** Continued.

Frequency	Transition								$E_l/k^a$	$A_{ul}^b$	$g_u^c$	rms <sup>d</sup>	$T_{\text{cont}}^e$	
13683.083	4	1	3	4	–	4	0	4	3	3.282	0.440	9		
13683.432	4	1	3	3	–	4	0	4	3	3.282	8.493	7		
13902.465	2	0	2	1	–	1	0	1	1	0.337	5.315	3	10.1	19.0
13903.080	2	0	2	1	–	1	0	1	2	0.337	0.354	3		
13903.523	2	0	2	3	–	1	0	1	2	0.337	12.760	7		
13903.706	2	0	2	2	–	1	0	1	1	0.337	9.569	5		
13904.002	2	0	2	1	–	1	0	1	0	0.337	7.088	3		
13904.321	2	0	2	2	–	1	0	1	2	0.337	3.190	5		
15090.186	3	0	3	2	–	2	1	2	2	1.266	1.107	5	30.5	18.5
15091.579	3	0	3	3	–	2	1	2	2	1.266	6.322	7		
15091.825	3	0	3	4	–	2	1	2	3	1.266	7.112	9		
15092.174	3	0	3	2	–	2	1	2	1	1.266	5.974	5		
15092.857	3	0	3	3	–	2	1	2	3	1.266	0.790	7		
15093.987	2	1	1	1	–	1	1	0	0	0.696	6.834	3	30.5	18.5
15094.813	2	1	1	2	–	1	1	0	2	0.696	3.075	5		
15095.432	2	1	1	3	–	1	1	0	2	0.696	12.301	7		
15095.777	2	1	1	1	–	1	1	0	2	0.696	0.342	3		
15096.006	2	1	1	2	–	1	1	0	1	0.696	9.225	5		
15096.969	2	1	1	1	–	1	1	0	1	0.696	5.126	3		
15791.081	5	1	4	5	–	5	1	5	5	4.961	1.550	11	12.6	14.9
15793.038	5	1	4	6	–	5	1	5	6	4.961	1.614	13		
15793.436	5	1	4	4	–	5	1	5	4	4.961	1.593	9		

**Notes.** <sup>(a)</sup> Energy of lower level in temperature units. <sup>(b)</sup> Einstein coefficient for spontaneous emission. <sup>(c)</sup> Degeneracy of upper energy level. <sup>(d)</sup> Noise level measured in baseline-removed spectrum at 38.1 kHz resolution. <sup>(e)</sup> Average continuum level measured around the transition or hyperfine multiplet.



## B.1 Complementary tables of cyanoacetamide

Table B.1 provides the centrifugal distortion parameters and Table B.2 lists the theoretical frequencies of the lowest vibrational modes.

**Table B.1** Quartic and sextic centrifugal distortion parameters for *trans* cyanoacetamide (*S*-reduction, *I*-representation).

Parameters	Theory <sup>(a)</sup>
$D_J$ (kHz)	0.49651
$D_K$ (kHz)	1.05766
$D_{JK}$ (kHz)	7.06844
$d_1$ (kHz)	-0.12301
$d_2$ (kHz)	-0.01721
$H_J$ (mHz)	0.32483
$H_{JK}$ (Hz)	-0.01259
$H_{KJ}$ (mHz)	5.29790
$H_K$ (Hz)	0.02284
$h_1$ (mHz)	0.13326
$h_2$ (mHz)	-0.00147
$h_3$ (mHz)	0.00865

**Notes.**<sup>(a)</sup> Theoretical computations at B2PLYPD3/augg-cc-pVTZ level of theory.

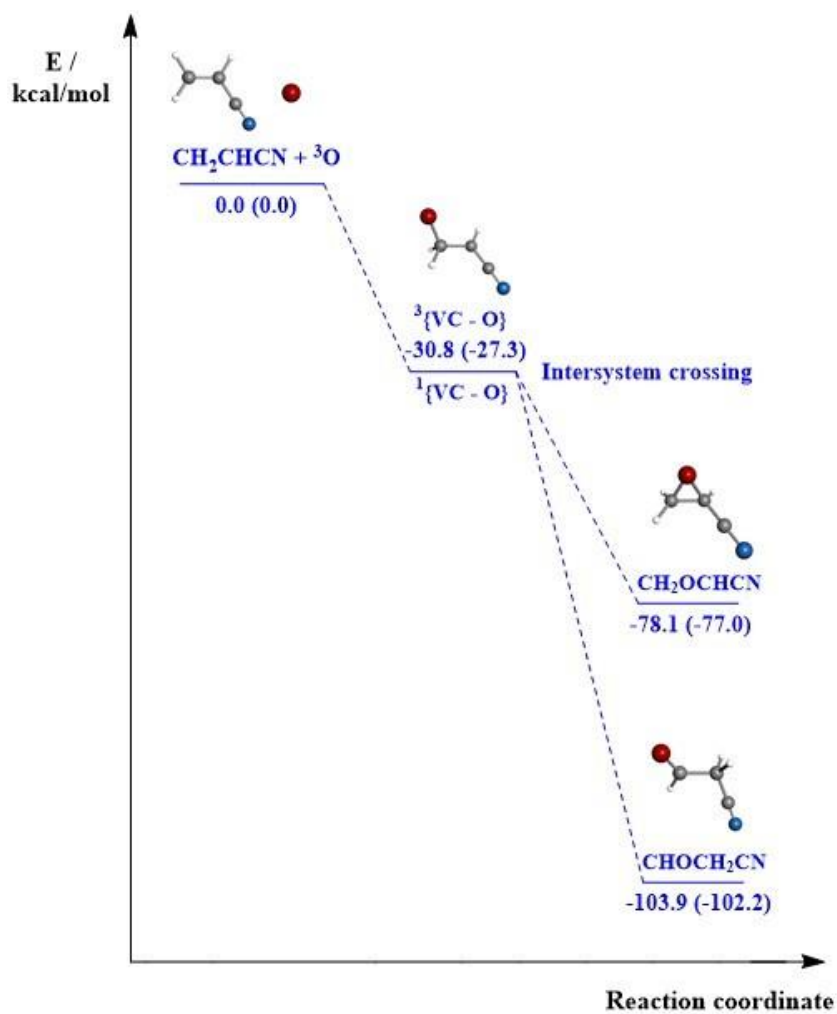
**Table B.2** Calculated harmonic and anharmonic frequencies together with the corresponding infrared intensities of the lowest vibrational modes.

Vibrational mode	$\omega^{(a)}$ (cm <sup>-1</sup> )	$I^{(b)}$ (km mol <sup>-1</sup> )	$\nu$ (cm <sup>-1</sup> )	$I$ (km mol <sup>-1</sup> )
1	54.2	16.0	54.1	27.7
2	159.6	17.7	166.3	16.9
3	324.7	178.8	447.8	112.8
4	354.2	8.4	360.3	13.6
5	400.7	5.2	398.6	4.6
6	498.8	7.9	491.5	6.3
7	510.1	4.7	499.3	10.0
8	583.4	9.6	579.0	16.0
9	672.5	7.2	648.9	7.7
10	852.6	8.6	831.5	7.5

**Notes.** <sup>(b)</sup> Harmonic ( $\omega$ ) and anharmonic ( $\nu$ ) frequencies (in cm<sup>-1</sup>) calculated at the B2PLYPD3/aug-cc-pVTZ level of theory. They are necessary to estimate the vibrational contribution to the partition function in order to obtain a proper estimate of the total column density (upper limit) of the molecule. <sup>(b)</sup> Infrared intensities (in km mol<sup>-1</sup>).

## B.2 Complementary reactions and figures

Figure B.2 shows the energy profile corresponding to the reaction between vinyl cyanide,  $\text{CH}_2\text{CHCN}$  (VC) and  $\text{O}(^3\text{P})$  radical to give cyanoacetaldehyde. This is an exothermic associative reaction that requires an intersystem spin crossing from the triplet state of the VC-O intermediate to the final singlet state of cyanoacetaldehyde. The reaction starts by forming a stable compound,  $^3[\text{VC-O}]$ , between the vinylcyanide molecule and the oxygen atom, in its triplet state. The formation of the triplet intermediate leads to a stabilization of 27.3 kcal/mol (at the CCSD(T)/aug-cc-pVTZ level).  $^3[\text{VC-O}]$  is a triplet intermediate from which two products, in their singlet state, can be generated by a spin crossing: (1) cyanoacetaldehyde, which can be formed by simple hydrogen migration, and (2) oxirane-2-carbonitrile, formed by a closure of the O-C-C 3-member ring. To obtain these two products, a spin transition is required. The change in spin multiplicity leads to a new intermediate, which is a singlet-state diradical  $^1[\text{VC-O}]$ . Depending on the value of the OCCH dihedral angle, this intermediate will finally evolve to oxirane-2-carbonitrile ( $\text{CH}_2\text{OCHCN}$ ) or to cyanoacetaldehyde ( $\text{CHOCH}_2\text{CN}$ ), which is the most stable product, and it is located 102.2 kcal/mol below reactants. From a purely thermodynamical point of view, cyanoacetaldehyde should be obtained. That the reaction needs an intersystem crossing to occur means that the reaction might be a slowly occurring but also a plausible process under astrophysical conditions. Furthermore, the singlet complex  $^1[\text{VC-O}]$  would require a further energetic stabilization by photon emission or even from the environment. A similar process was studied in the reaction of propylene and  $\text{O} (^3\text{P})$ , which led to propanal instead of the targeted propylene oxide (Bodo et al. 2019).

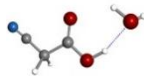
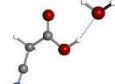


**Fig. B.2** Energy profile in kcal/mol for the reactions of vinyl cyanide with O ( ${}^3\text{P}$ ) radical producing cyanoacetaldehyde computed at the B2PLYPD3/aug-cc-pVTZ and CCSD(T)/aug-cc-pVTZ (in parentheses) levels of theory. The zero-point vibrational energy computed at the B2PLYPD3/aug-cc-pVTZ level is included.

### B.3 Cyanoacetic acid-water clusters

A systematic conformational search was carried out to predict cyanoacetic acid-water clusters using refined fast molecular mechanics methods (MMFFs). Subsequently, all the structures were geometrically optimized at higher levels of theory. All quantum mechanical computations were carried out with the GAUSSIAN 16 program package (Gaussian16 Revision A.03).

**Table B.3** Theoretical and experimental spectroscopic parameters for the cyanoacetic-acid water clusters.

Parameters	rotamer w-I	 <i>w-cis</i> <sup>[a]</sup>	 <i>w-gauche</i>
<i>A</i> <sup>[b]</sup>	7074.50 (33) <sup>[i]</sup>	7116.6	6108.6
<i>B</i>	1158.6515 (11)	1169.8	1227.2
<i>C</i>	1003.1681 (13)	1012.7	1056.1
$\Delta J$	-0.204 (9)	-	-
$\Delta JK$	4.68 (21)	-	-
<i>P<sub>c</sub></i> <sup>[c]</sup>	1.9162 (11)	1.997	8.001
$ \mu_a ,  \mu_b ,  \mu_c $ <sup>[d]</sup>	Yes / No / No	3.5 / 2.4 / 1.0	2.9 / 1.7 / 0.2
<i>N</i> <sup>[f]</sup>	14	-	-
$\sigma$ <sup>[g]</sup>	15.2	-	-
$\Delta E_{TOTAL}$ <sup>[h]</sup>	-	0.0	116.9

[a] Theoretical calculations at B2PLYPD3/aug-cc-pVTZ level of theory. [b] *A*, *B*, and *C* represent the rotational constants (in MHz); [c] *P<sub>c</sub>* is the planar inertial moment (in uÅ<sup>2</sup>), conversion factor: 505379.1 MHz-uÅ<sup>2</sup>. [d]  $\mu_a$ ,  $\mu_b$ , and  $\mu_c$  are the electric dipole moment components (in D). [e]  $\chi_{aa}$ ,  $\chi_{bb}$  and  $\chi_{cc}$  are the diagonal elements of the <sup>14</sup>N nuclear quadrupole coupling tensor (in MHz), which are fixed in this case to the theoretical values; [f] *N* is the number of measured transitions. [g]  $\sigma$  is the root mean square (rms) deviation of the fit (in kHz). [h]  $\Delta E_{TOTAL}$  is the electronic energy (in cm<sup>-1</sup>) relative to the global minimum calculated at the B2PLYPD3/aug-cc-pVTZ level of theory, taking into account the zero-point vibrational energy (ZPE) for the electronic energy ( $E_{TOTAL} = E + E_{ZPE}$ ) calculated at the same level [i] Standard error in parentheses in units of the last digit.

According to B2PLYPD3 double-hybrid computations, cyanoacetic acid-water clusters can adopt two different configurations labeled as *w-cis*- and *w-gauche*- forms (see Table B.3). In both structures, water molecules interact through the carboxyl group, simultaneously serving as a proton donor to the carboxyl oxygen atom and accepting a proton from the hydroxy group. Thus, these six-membered intermolecular networks benefit from two hydrogen bonds, with water acting as a bridge between the carboxyl O and the hydroxy moiety of cyanoacetic acid in a nearly planar structure and with the *w-cis*- form being the global minimum.

*B.4 Complementary tables of cyanoacetic acid*

In Tables B.4–B.6 we provide a list containing all measured lines for *w-cis*-cyanoacetic acid, *cis*-cyanoacetic acid, and *gauche*-cyanoacetic acid, respectively.

**Table B.4** List of the measured rotational transitions (frequencies and residuals, in MHz) of

J'	K <sub>a</sub> '	K <sub>c</sub> '	J''	K <sub>a</sub> ''	K <sub>c</sub> ''	$\nu_{obs}$ (MHZ) <sup>[a]</sup>	$\nu_{calc}$ (MHZ) <sup>[b]</sup>	$\nu_{obs} - \nu_{calc}$ (MHZ) <sup>[c]</sup>
8	1	8	7	1	7	16616.298	16616.283	0.015
7	1	6	6	1	5	15634.658	15634.674	-0.016
7	0	7	6	0	6	14966.697	14966.692	0.005
4	0	4	3	0	3	8617.053	8617.069	-0.016
5	0	5	4	0	4	10748.862	10748.827	0.034
6	1	6	5	1	5	12481.155	12481.183	-0.028
6	0	6	5	0	5	12866.130	12866.142	-0.012
6	1	5	5	1	4	13412.028	13412.017	0.011
5	2	3	4	2	2	10862.652	10862.640	0.011
6	2	5	5	2	4	12957.382	12957.386	-0.004
6	2	4	5	2	3	13061.990	13061.979	0.010
7	2	6	6	2	5	15109.168	15109.173	-0.005
7	2	5	6	2	4	15274.909	15274.913	-0.004
8	2	6	7	2	5	17502.370	17502.371	-0.001

**Note.** Upper and lower state quantum numbers are indicated by ' and '' , respectively. [a] Observed frequency. [b] Calculated frequency. [c] Observed minus calculated frequency.

**Table B.5.** List of the measured rotational transitions (frequencies and residuals, in MHz) of *cc*-CAA.

J'	K <sub>-l</sub> '	K <sub>+l</sub> '	J''	K <sub>-l</sub> ''	K <sub>+l</sub> ''	F'	F''	$\nu_{obs.}$	$\nu_{obs.} - \nu_{cal.}$
2	1	2	1	1	1	2	1	7898.197	0.001
2	1	2	1	1	1	2	2	7898.544	0.002
2	1	2	1	1	1	3	2	7899.204	0.001
2	1	2	1	1	1	1	2	7899.573	0.001
2	1	2	1	1	1	1	0	7900.091	0.001
2	0	2	1	0	1	2	2	8273.942	0.003
2	0	2	1	0	1	1	0	8274.088	0.000
2	0	2	1	0	1	2	1	8274.904	0.002
2	0	2	1	0	1	3	2	8274.984	0.002
2	0	2	1	0	1	1	1	8276.496	-0.001
1	1	0	1	0	1	1	0	8312.926	0.001
1	1	0	1	0	1	2	2	8313.732	0.000
1	1	0	1	0	1	2	1	8314.694	-0.001
1	1	0	1	0	1	1	1	8315.331	-0.002
2	1	1	1	1	0	2	1	8678.062	0.003
2	1	1	1	1	0	3	2	8679.071	0.003
2	1	1	1	1	0	1	0	8680.207	0.002
2	1	1	2	0	2	2	1	8716.899	0.003
2	1	1	2	0	2	3	3	8717.818	0.000
2	1	1	2	0	2	2	2	8718.489	-0.001
2	1	1	2	0	2	3	2	8718.862	0.001
2	1	1	2	0	2	1	2	8719.065	-0.002
3	1	2	3	0	3	2	2	9349.703	-0.003
3	1	2	3	0	3	4	4	9349.858	0.001
3	1	2	3	0	3	3	3	9350.291	0.001
4	0	4	3	1	3	4	3	9427.791	0.000
4	0	4	3	1	3	3	2	9427.910	-0.004
4	0	4	3	1	3	5	4	9427.962	0.004
4	1	3	4	0	4	3	3	10239.91	-0.001

4	1	3	4	0	4	5	5	10240.00	0.002
3	1	3	2	1	2	3	3	11838.99	-0.004
3	1	3	2	1	2	3	2	11839.66	-0.003
3	1	3	2	1	2	2	1	11839.92	0.001
3	1	3	2	1	2	4	3	11839.95	-0.002
3	1	3	2	1	2	2	2	11840.94	-0.004
1	1	1	0	0	0	0	1	12067.95	-0.004
1	1	1	0	0	0	2	1	12068.46	-0.002
1	1	1	0	0	0	1	1	12068.81	-0.003
3	0	3	2	0	2	3	3	12376.32	0.000
3	0	3	2	0	2	3	2	12377.37	0.003
3	0	3	2	0	2	4	3	12377.41	0.001
3	0	3	2	0	2	2	2	12378.83	0.000
3	2	2	2	2	1	2	2	12432.55	0.000
3	2	2	2	2	1	3	3	12433.58	-0.001
3	2	2	2	2	1	2	1	12434.16	0.001
3	2	1	2	2	0	3	2	12488.60	-0.003
3	2	1	2	2	0	2	3	12489.62	0.002
3	2	1	2	2	0	2	1	12490.19	-0.003
6	1	5	6	0	6	7	7	12946.29	0.001
6	1	5	6	0	6	6	6	12946.57	-0.001
3	1	2	2	1	1	3	2	13009.17	0.001
3	1	2	2	1	1	4	3	13009.45	0.001
3	1	2	2	1	1	2	1	13009.47	-0.004
5	0	5	4	1	4	5	4	14103.84	-0.02
5	0	5	4	1	4	6	5	14103.91	-0.01
4	1	4	3	1	3	3	2	15770.79	0.005
2	1	2	1	0	1	1	0	15821.93	-0.02
2	1	2	1	0	1	2	2	15822.36	-0.008
2	1	2	1	0	1	3	2	15823.02	-0.01
2	1	2	1	0	1	2	1	15823.31	-0.02
2	1	2	1	0	1	1	1	15824.35	-0.02
4	0	4	3	0	3	4	3	16438.56	0.05
4	2	3	3	2	2	4	3	16566.50	-0.02
4	2	3	3	2	2	5	4	16566.97	0.01
4	2	2	3	2	1	4	3	16705.82	-0.03
4	2	2	3	2	1	5	4	16706.26	-0.01
3	1	3	2	0	2	2	1	19387.81	0.03
3	1	3	2	0	2	4	3	19388.04	0.04
3	1	3	2	0	2	3	2	19388.10	0.01
3	1	3	2	0	2	2	2	19389.40	0.02
5	0	5	4	0	4	6	5	20446.80	-0.01
4	1	4	3	0	3	4	4	22780.34	-0.01
4	1	4	3	0	3	3	2	22781.32	-0.02
4	1	4	3	0	3	5	4	22781.44	0.004
4	1	4	3	0	3	5	4	22781.48	0.04
4	1	4	3	0	3	3	3	22782.84	0.04
2	2	1	2	1	2	3	3	24941.45	-0.002
2	2	1	2	1	2	3	2	24942.12	0.006
2	2	1	2	1	2	2	1	24942.16	0.04
2	2	1	2	1	2	2	2	24943.13	-0.02
5	1	5	4	0	4	4	3	26032.08	0.02
5	1	5	4	0	4	5	4	26032.08	0.008
5	1	5	4	0	4	6	5	26032.12	0.006
5	1	5	4	0	4	4	4	26033.48	0.005
2	2	1	1	1	0	3	2	32450.75	-0.001
2	2	1	1	1	0	1	0	32451.10	-0.008
2	2	0	1	1	1	1	1	32853.76	-0.01
2	2	0	1	1	1	1	0	32854.63	-0.009
2	2	0	1	1	1	3	2	32854.69	-0.001

**Note.** Upper and lower state quantum numbers are indicated by ' and ', respectively. [a] Observed frequency. [b] Calculated frequency. [c] Observed minus calculated frequency.

**Table B.6.** List of the measured rotational transitions (frequencies and residuals in MHz) of *gc*-CAA.

$J'$	$K'_{-1}$	$K'_{+1}$	$J''$	$K''_{-1}$	$K''_{+1}$	F'	F''	V <sub>obs.</sub>	V <sub>obs.</sub> -V <sub>cal.</sub>
2	1	2	1	1	1	2	1	8257.093	0.004
2	1	2	1	1	1	2	2	8257.403	-0.005
2	1	2	1	1	1	3	2	8257.978	0.005
2	1	2	1	1	1	1	0	8258.761	-0.004
2	0	2	1	0	1	2	2	8631.925	-0.000
2	0	2	1	0	1	1	0	8632.075	0.001
2	0	2	1	0	1	2	1	8632.772	0.001
2	0	2	1	0	1	3	2	8632.835	-0.001
2	0	2	1	0	1	1	1	8634.189	0.000
2	1	1	1	1	0	2	1	9039.183	0.001
2	1	1	1	1	0	2	2	9039.700	-0.008
2	1	1	1	1	0	1	1	9039.717	0.004
2	1	1	1	1	0	3	2	9040.050	-0.001
2	1	1	1	1	0	1	0	9041.036	0.005
3	1	3	2	1	2	3	2	12376.39	0.001
3	1	3	2	1	2	2	1	12376.62	0.001
3	1	3	2	1	2	4	3	12376.64	-0.001
3	0	3	2	0	2	3	3	12907.80	-0.002
3	0	3	2	0	2	2	1	12908.58	0.002
3	0	3	2	0	2	3	2	12908.71	0.001
3	0	3	2	0	2	4	3	12908.75	0.002
3	0	3	2	0	2	2	2	12909.99	-0.003
3	1	2	2	1	1	3	2	13549.10	-0.003
3	1	2	2	1	1	4	3	13549.35	-0.003
3	1	2	2	1	1	2	1	13549.38	0.005

**Note.** Upper and lower state quantum numbers are indicated by ' and ', respectively. [a] Observed frequency. [b] Calculated frequency. [c] Observed minus calculated frequency.





C.1 Complementary tables of *n*- and *i*-PrCHO along with the analysis of acetaldehyde.

In Table C.1 we provide a list with the theoretical frequencies of the lowest vibrational modes of *n*- and *i*-PrCHO. In Table C.2 we show the analysis of acetaldehyde towards the G+0.693-027 molecular cloud.

**Table C.1.** Theoretically predicted harmonic vibrational frequencies of the vibrational modes for the lowest energy conformers of *n*- and *i*-PrCHO.

vibrational mode	<i>n</i> -PrCHO				<i>i</i> -PrCHO	
	<i>cis-gauche</i>	<i>cis-trans</i>	<i>gauche-gauche</i>	<i>gauche-trans</i>	<i>gauche</i>	<i>trans</i>
1	109.5	82.5	66.3	67.2	95.3	63.4
2	162.8	172.6	116.6	98.3	204.9	208.4
3	201.0	196.2	222.3	236.1	238.6	236.2
4	283.6	248.2	304.4	255.5	275.4	331.0
5	364.9	347.9	402.2	385.7	341.6	331.3
6	664.5	682.0	515.5	518.9	399.9	352.4
7	721.2	703.8	771.2	753.4	649.9	551.8
8	802.2	799.0	795.6	821.9	816.5	858.4
9	863.3	873.7	917.4	917.5	926.9	938.4
10	959.9	960.1	961.8	964.7	946.8	946.9
11	980.2	979.0	993.3	1030.2	960.0	967.8
12	1073.5	1074.9	1087.3	1078.6	991.4	1006.7
13	1138.0	1155.4	1143.6	1158.0	1149.0	1163.5
14	1166.7	1174.0	1180.1	1182.6	1181.1	1203.3
15	1270.7	1272.4	1256.2	1270.7	1221.9	1206.3
16	1316.2	1331.7	1299.7	1305.3	1335.5	1339.3
17	1390.2	1341.2	1371.3	1344.2	1384.8	1365.3
18	1412.4	1424.1	1394.0	1406.3	1422.6	1417.8
19	1432.6	1436.5	1438.6	1434.6	1429.9	1431.9
20	1443.6	1442.3	1445.8	1447.3	1451.7	1444.9
21	1471.7	1476.6	1486.0	1489.4	1505.2	1507.3
22	1505.8	1510.7	1511.1	1513.4	1510.6	1512.5
23	1519.7	1521.4	1521.8	1520.2	1524.2	1524.7
24	1530.8	1527.9	1524.5	1527.9	1528.9	1530.6
25	1836.5	1836.7	1839.4	1839.0	1836.4	1840.1
26	2947.7	2947.3	2926.8	2934.1	2935.6	2916.4
27	3041.0	3037.6	3051.5	3046.4	3034.0	3050.7
28	3054.1	3049.4	3054.2	3053.2	3052.5	3053.0
29	3067.0	3065.9	3058.0	3053.6	3064.3	3098.7
30	3076.8	3076.6	3096.9	3086.1	3123.0	3121.2
31	3107.5	3102.8	3123.1	3121.8	3133.8	3128.2
32	3125.3	3126.3	3128.1	3130.1	3137.3	3130.8
33	3145.8	3129.0	3132.5	3131.4	3147.3	3133.7

**Notes.** <sup>(a)</sup> Frequencies (in  $\text{cm}^{-1}$ ) calculated at the CCSD/cc-pVTZ level of theory. They are necessary to estimate the vibrational contribution to the partition function in order to obtain a proper estimate of the total column density (upper limit) of the molecule. <sup>(b)</sup> Infrared intensities (in  $\text{km mol}^{-1}$ ).

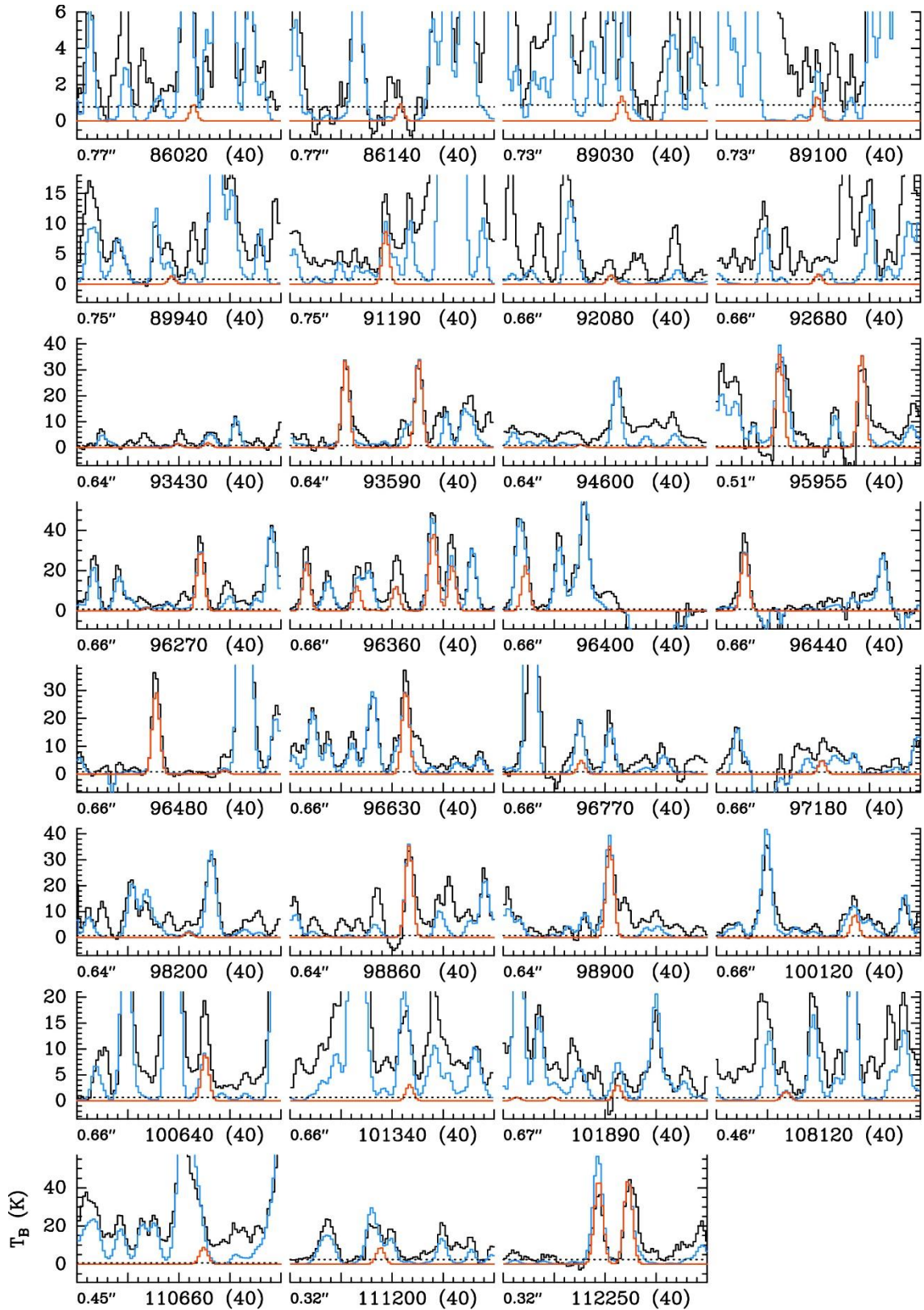
To perform the LTE analysis of CH<sub>3</sub>CHO we have selected transitions that are not blended with other molecular species, and that cover a broad range of energy levels, in order to constrain the excitation conditions. The transitions used in the analysis are listed in Table C.2. We have used the spectroscopic predictions of Smirnov et al. (2014), as in the analysis of Sgr B2(N1S). We used MADCUBA–SLIM, leaving as free parameters the molecular column density ( $N$ ), the rotational temperature  $T_{\text{rot}}$ , the velocity ( $V_{\text{LSR}}$ ) and linewidth  $\Delta V$ . The derived physical parameters are shown in Table 10, and the fits of the transitions are shown in Figure C.10 (following section).

**Table C.2.** Selected unblended transitions of CH<sub>3</sub>CHO detected towards G+0.693–0.027 used to perform the LTE analysis.

Frequency (MHz)	Transition	Log $I$ (nm <sup>2</sup> MHz)	$g_u$	$E_{l_0}$ (cm <sup>-1</sup> )
76866.4396	4 <sub>04</sub> ← 3 <sub>03</sub> E	-4.7276	9	3.9210
76878.9554	4 <sub>04</sub> ← 3 <sub>03</sub> A	-4.7273	9	3.8527
93580.9130	5 <sub>15</sub> ← 4 <sub>14</sub> A	-4.4871	11	7.8241
93595.2373	5 <sub>15</sub> ← 4 <sub>14</sub> E	-4.4871	11	7.8755
95947.4413	5 <sub>05</sub> ← 4 <sub>04</sub> E	-4.4444	11	6.4850
95963.4618	5 <sub>05</sub> ← 4 <sub>04</sub> A	-4.4441	11	6.4171
114940.1789	6 <sub>06</sub> ← 5 <sub>05</sub> E	-4.2159	13	9.6854
114959.9048	6 <sub>06</sub> ← 5 <sub>05</sub> A	-4.2156	13	9.6181
133830.4954	7 <sub>07</sub> ← 6 <sub>06</sub> E	-4.0256	15	13.5194
133854.1003	7 <sub>07</sub> ← 6 <sub>06</sub> A	-4.0253	15	13.4527
138285.0010	7 <sub>16</sub> ← 6 <sub>15</sub> E	-4.0094	15	15.4866
138319.6276	7 <sub>16</sub> ← 6 <sub>15</sub> A	-4.0100	15	15.4352
149505.1314	8 <sub>18</sub> ← 7 <sub>17</sub> E	-3.8898	17	19.1080
149507.4671	8 <sub>18</sub> ← 7 <sub>17</sub> A	-3.8897	17	19.0559
157937.7016	8 <sub>17</sub> ← 7 <sub>16</sub> E	-3.8445	17	20.0993
157974.5872	8 <sub>17</sub> ← 7 <sub>16</sub> A	-3.8442	17	20.0491
168008.6224	9 <sub>19</sub> ← 8 <sub>18</sub> E	-3.7465	19	24.0950
168093.4506	9 <sub>19</sub> ← 8 <sub>18</sub> A	-3.7464	19	24.0430
171265.4737	9 <sub>09</sub> ← 8 <sub>08</sub> E	-3.7236	19	23.0740
171296.9876	9 <sub>09</sub> ← 8 <sub>08</sub> A	-3.7233	19	23.0090

## C.2 Complementary figures: Spectra and population diagram of Sgr B2(N1S)

Figures C1–C3 show the transitions of CH<sub>3</sub>CHO,  $v_t = 0$ ,  $v_t = 1$ , and  $v_t = 2$  that are covered by the ReMoCA survey and contribute significantly to the signal detected toward Sgr B2(N1S). Figure C.4 shows the population diagram of CH<sub>3</sub>CHO for Sgr B2(N1S). Figures C.5 and C.6 illustrate the nondetection of *syn* and *gauche* C<sub>2</sub>H<sub>5</sub>CHO, respectively, Figs C.7 and C.8 the nondetection of the *cis-gauche* and *cis-trans* conformers of *normal*-C<sub>3</sub>H<sub>7</sub>CHO, respectively, and Fig. C.9 the nondetection of *gauche iso*-C<sub>3</sub>H<sub>7</sub>CHO, all toward Sgr B2(N1S). Finally, Fig. C.10 shows the fits of the transitions to illustrate the detection of acetaldehyde towards G+0693.



**Fig. C.1.** Transitions of  $\text{CH}_3\text{CHO}$ ,  $v=0$  covered by the ReMoCA survey. The best-fit LTE synthetic spectrum of  $\text{CH}_3\text{CHO}$ ,  $v=0$  is displayed in red and overlaid on the observed spectrum of Sgr B2(N1S) shown in black. The blue synthetic spectrum contains the contributions of all molecules identified in our survey so far, including the species shown in red. The central frequency and width are indicated in MHz below each panel as well as the half-power beam width. The y-axis is labeled in brightness temperature units (K). The dotted line indicates the  $3\sigma$  noise level.

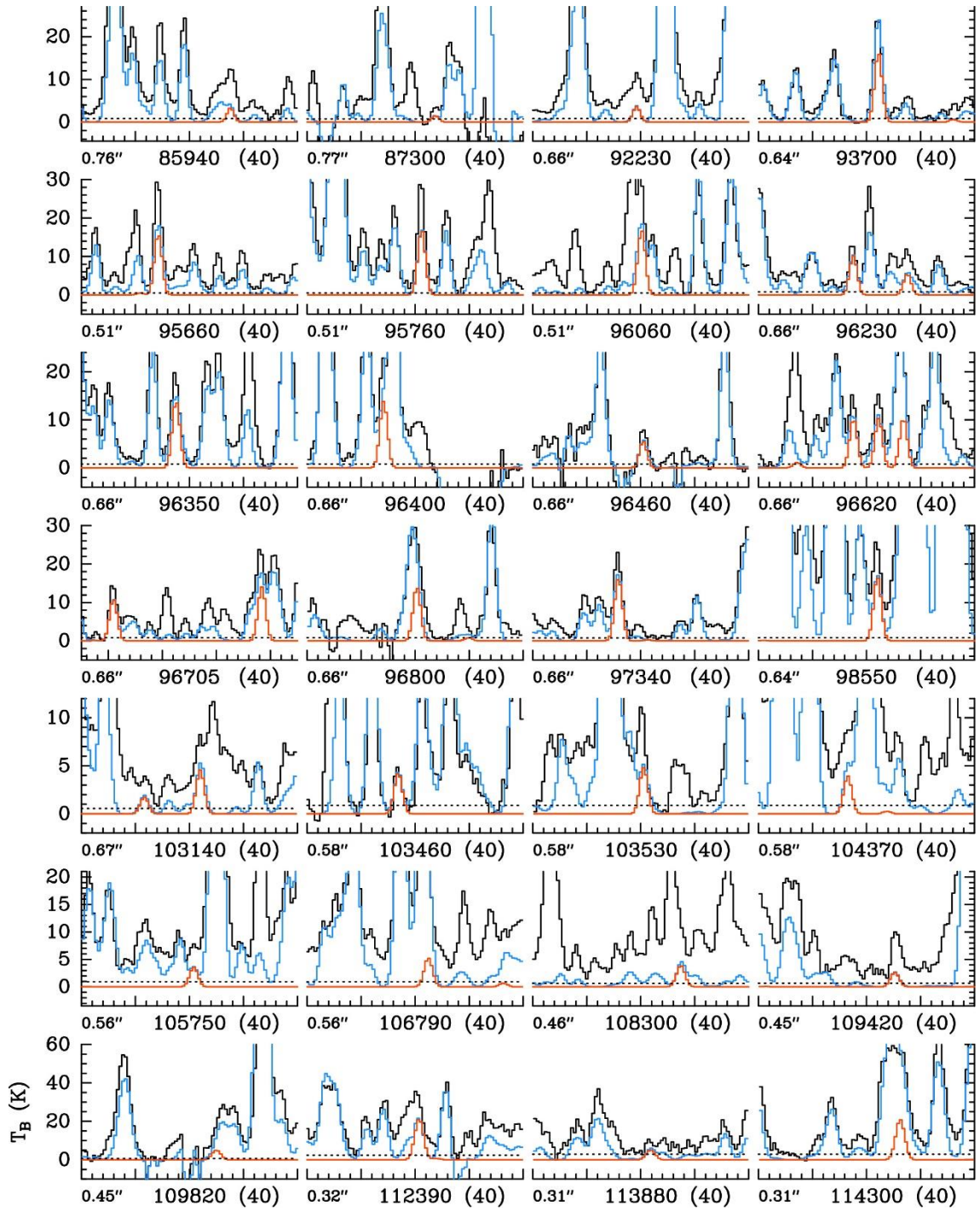


Fig. C.2. Same as Fig. B.1 but for  $\text{CH}_3\text{CHO}$ ,  $v=1$ .

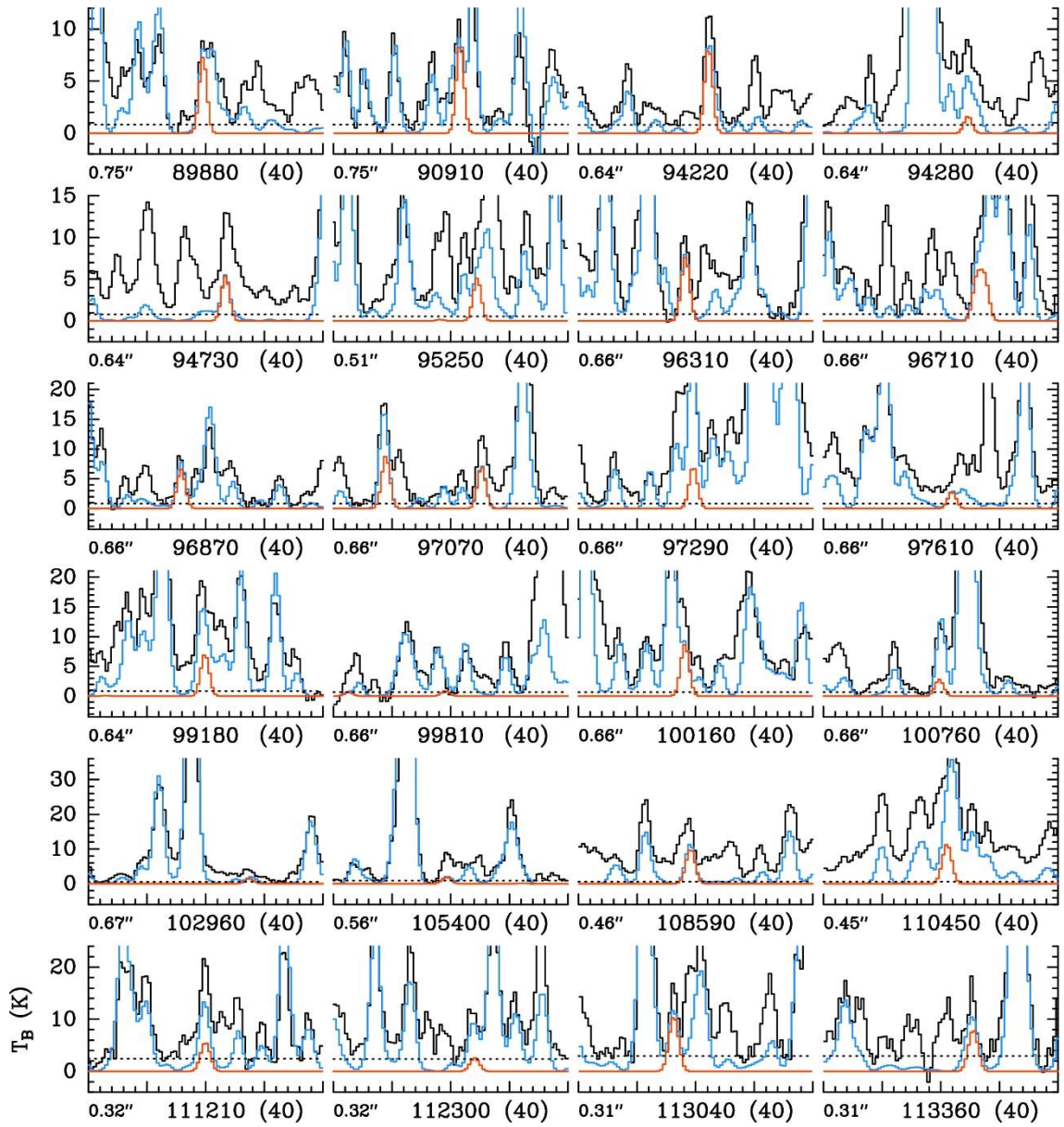
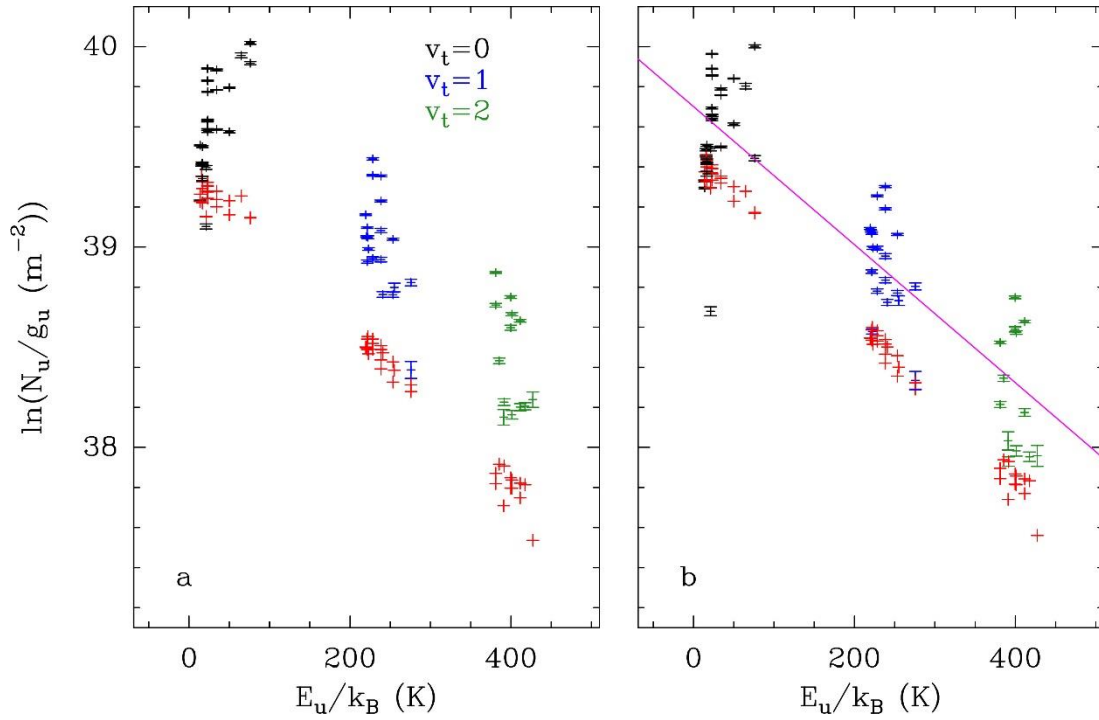
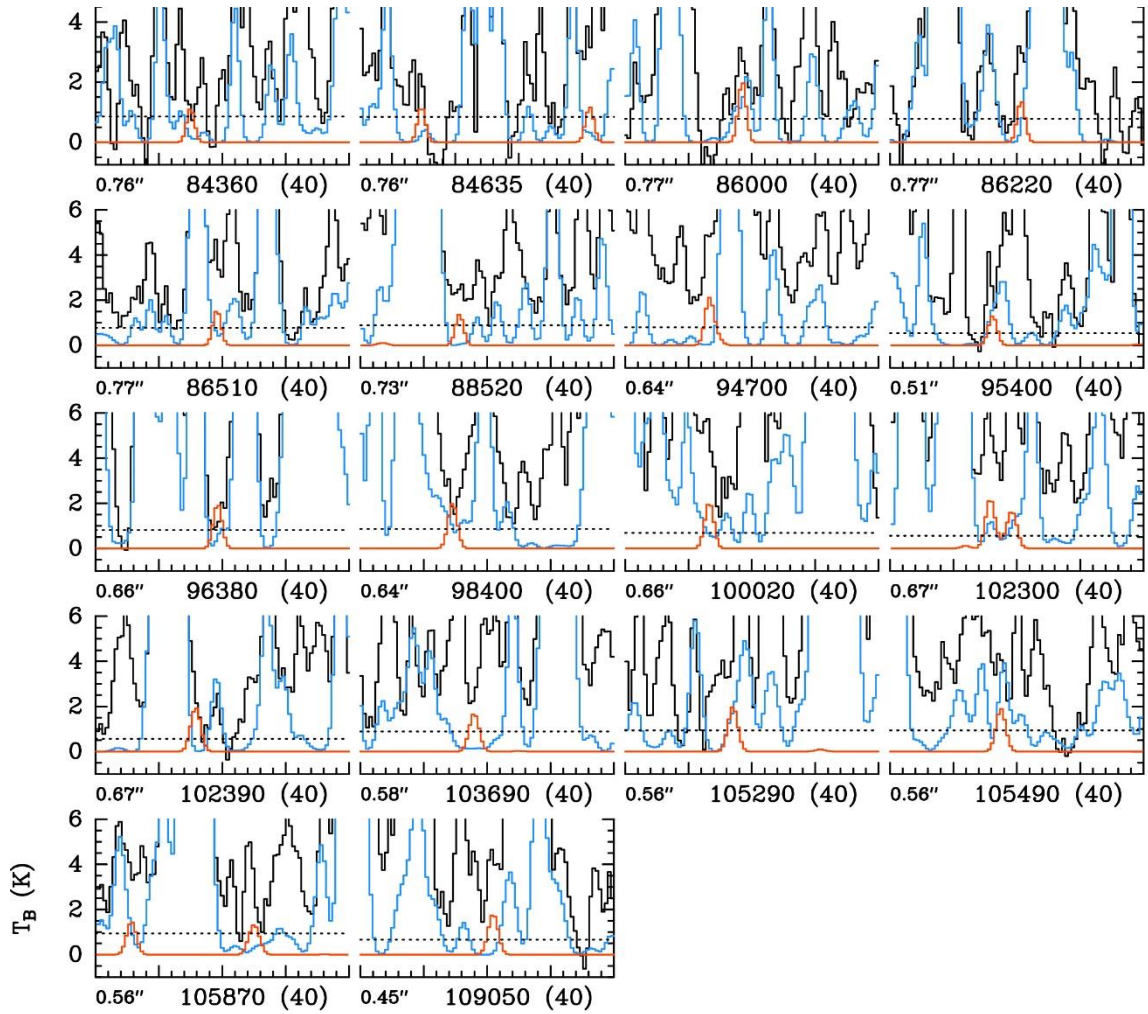


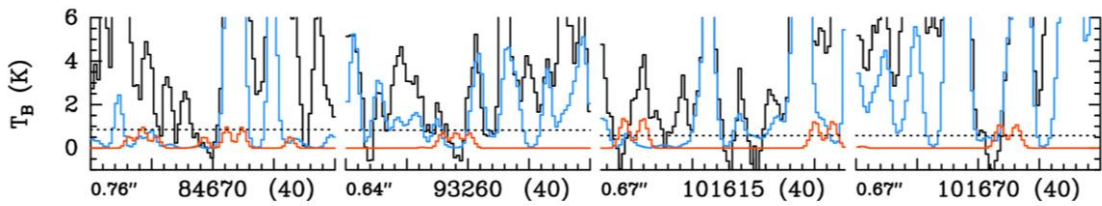
Fig. C.3. Same as Fig. B.1 but for  $\text{CH}_3\text{CHO}$ ,  $v=2$ .



**Fig. C.4.** Population diagram of CH<sub>3</sub>CHO toward Sgr B2(N1S). The observed data points are shown in various colors (but not red) as indicated in the upper right corner of panel **a** while the synthetic populations are shown in red. No correction is applied in panel **a**. In panel **b**, the optical depth correction has been applied to both the observed and synthetic populations and the contamination by all other species included in the full model has been removed from the observed data points. The purple line is a linear fit to the observed populations (in linear-logarithmic space).



**Fig. C.5.** Same as Fig. B.1 but the red spectrum shows the synthetic spectrum of *syn*-C<sub>2</sub>H<sub>5</sub>CHO,  $v=0$  used to derive the upper limit on its column density reported in Table 9. The blue spectrum does not contain the contribution of the species shown in red.



**Fig. C.6.** Same as Fig. B.5 but for *gauche*-C<sub>2</sub>H<sub>5</sub>CHO,  $v=0$ .

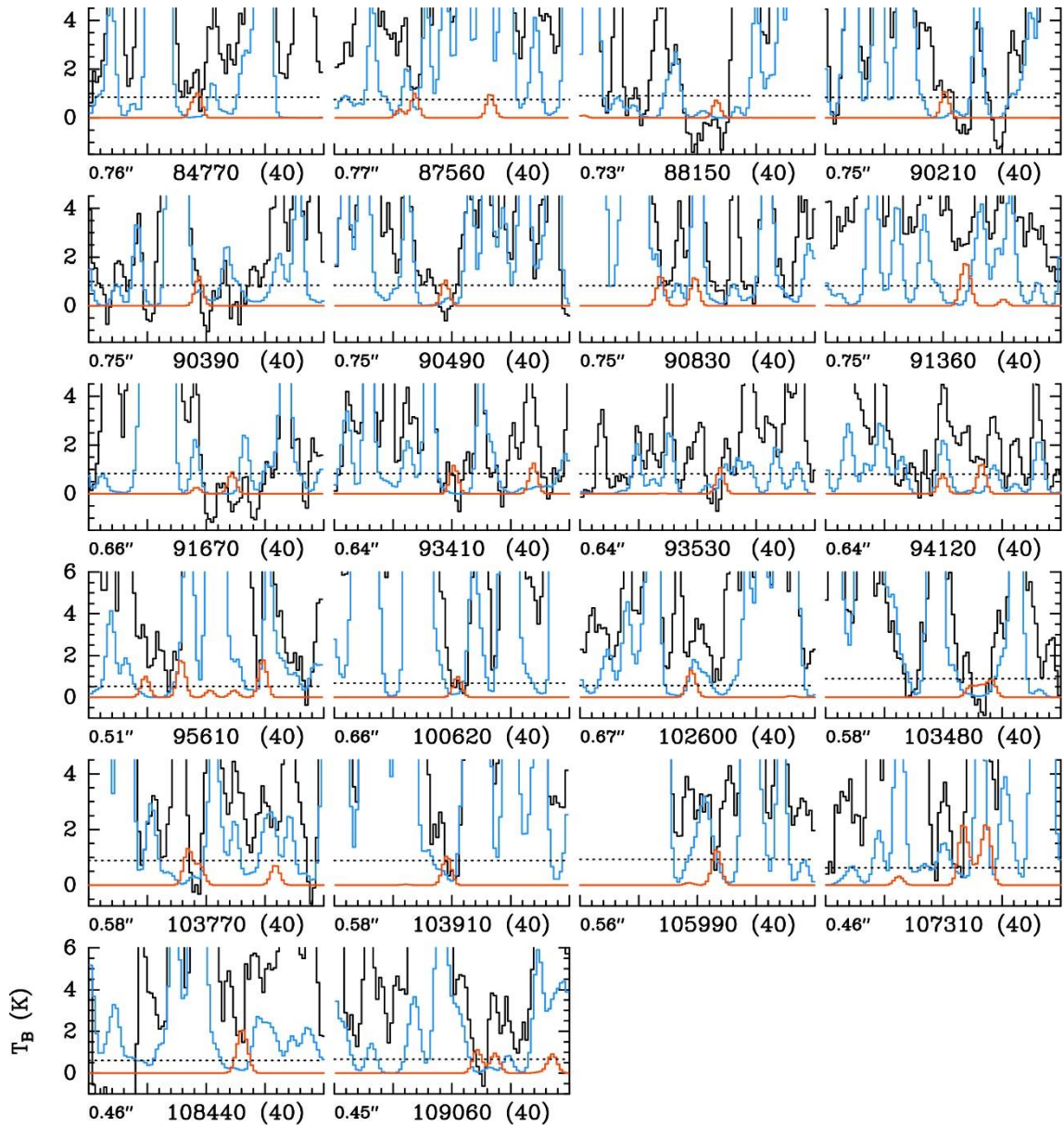


Fig. C.7. Same as Fig. B.5 but for *cis-gauche normal*- $C_3H_7CHO$ ,  $v=0$ .



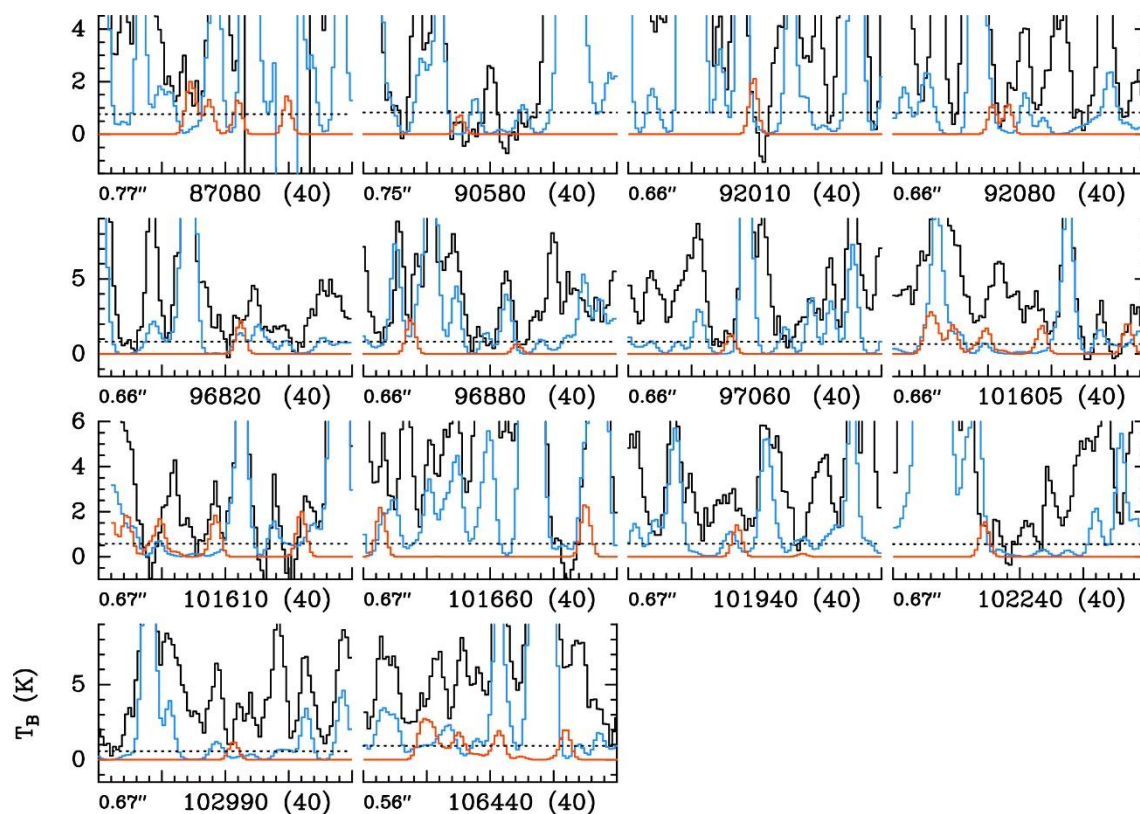


Fig. C.8. Same as Fig. B.5 but for *cis-trans normal*- $C_3H_7CHO$ ,  $v=0$ .

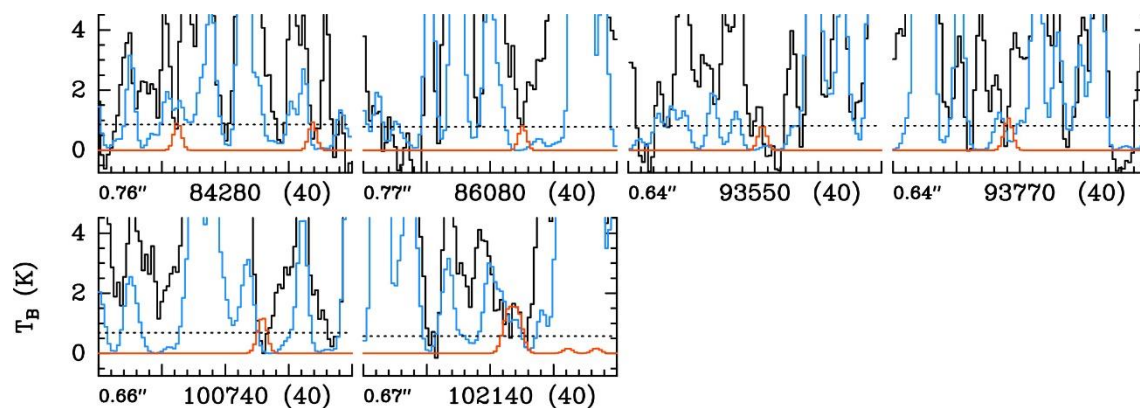
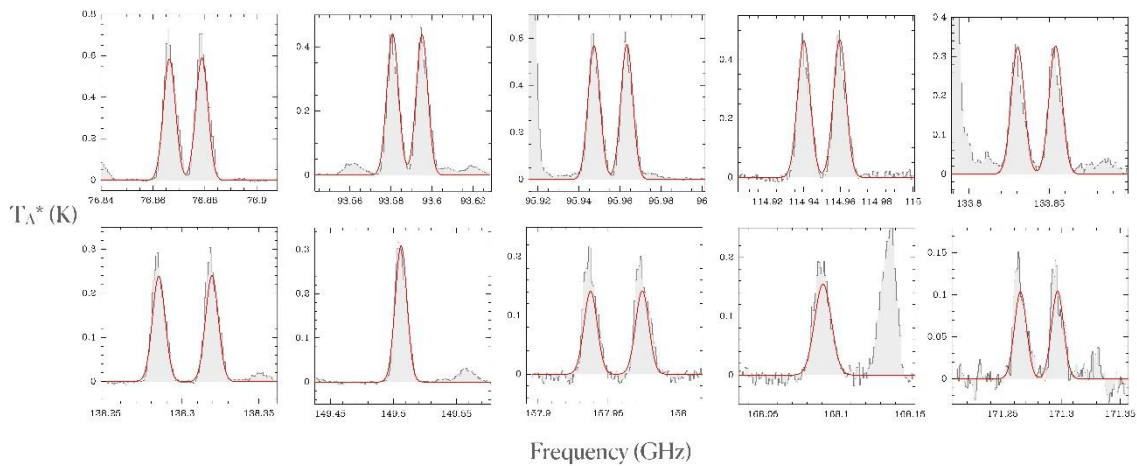


Fig. C.9. Same as Fig. B.9 but for *gauche iso*- $C_3H_7CHO$ ,  $v=0$ .



**Fig. C.10.** Selected unblended transitions of CH<sub>3</sub>CHO (see Table C.2) detected towards G+0.693–0.027 used to perform the LTE analysis. The gray histograms show the observed spectra, while the red curve is the best LTE fit obtained with MADCUBA.

*D.1 Theoretical benchmark*

A set of 12 candidate structures bellow an energy window of  $800 \text{ cm}^{-1}$  were identified by subsequent geometry optimizations of the MMFFs predicted structures. The density functional theory B3LYP method with the Becke's three-parameter exchange functional and the Lee-Yang-Parr correlation functional (Becke, 1988; Lee et al. 1988) was used. Also, Grimme's D3 dispersion correction factors were included (Grimme et al. 2010).

**Table D1.** Theoretical spectroscopic parameters for the low energy L-DOPA conformers.

Parameter <sup>[a]</sup>	GIa	GIIa	GIIb	GIIb	AIb	AIa	AIIb	AIIIb	AIVb	AIVa	AIIa	AIIIa
$A / \text{MHz}$ <sup>[b]</sup>	1195	1323	1203	1343	1743	1750	1800	1793	1720	1730	1804	1808
$B / \text{MHz}$	433	395	428	389	306	305	303	299	301	300	302	295
$C / \text{MHz}$	373	355	369	351	279	280	277	277	279	282	277	278
$ \mu_a /D$ <sup>[c]</sup>	0.3	2.6	2.6	2.1	3.6	5.7	5.7	5.6	3.3	5.6	4.0	3.9
$ \mu_b /D$	3.0	2.1	6.4	6.5	0.2	4.1	0.4	0.1	0.6	4.3	3.3	3.6
$ \mu_c /D$	0.3	0.5	1.3	0.0	2.1	1.9	2.7	2.5	1.7	1.6	1.3	0.8
$\Delta E / \text{cm}^{-1}$ <sup>[g]</sup>	0	276	495	582	508	537	546	635	652	661	657	795

[a] Calculated at the B3LYP/aug-cc-pVTZ level of the theory with Grimme D3 dispersion. [b]  $A$ ,  $B$ , and  $C$  are the rotational constants; [c]  $|\mu_a|$ ,  $|\mu_b|$  and  $|\mu_c|$  are the absolute values of the electric dipole moment components along the inertial axis  $a$ ,  $b$ ,  $c$ ; [d]  $\Delta E$  represents the relative energy (zero-point energy taken into account), with respect to the global minimum.

*D.2 Measured transitions for L-DOPA*

In Tables D2.1–D2.4 we provide a list containing all measured lines for the GIa, GIIa, AIb and AIIa conformers L-DOPA.

**Table D2.1** List of the measured rotational transitions (frequencies and residuals, in MHz) of GIa L-DOPA.

$J'$	$K_a'$	$K_c'$	$J''$	$K_a''$	$K_c''$	$\nu_{obs}(MHz)^{[a]}$	$\nu_{calc}(MHz)^{[b]}$	$\nu_{obs} - \nu_{calc}(MHz)^{[c]}$	weight
6	2	4	6	1	5	1981.217	1981.220	-0.003	
5	2	3	5	1	4	2034.990	2034.984	0.006	
3	1	3	2	0	2	3043.679	3043.683	-0.004	
5	0	5	4	1	4	3476.403	3476.373	0.031	
7	3	4	7	2	5	3664.242	3664.258	-0.016	
4	1	4	3	0	3	3741.790	3741.821	-0.031	
6	3	3	6	2	4	3768.218	3768.219	-0.001	
5	3	2	5	2	3	3842.695	3842.655	0.040	
4	3	1	4	2	2	3889.402	3889.373	0.029	
2	2	1	1	1	0	3954.316	3954.292	0.024	
6	3	4	6	2	5	3993.164	3993.147	0.017	
9	2	7	8	3	6	4080.072	4080.079	-0.008	
7	1	6	6	2	5	4108.932	4108.929	0.003	
8	3	6	8	2	7	4114.892	4114.911	-0.019	
9	3	7	9	2	8	4212.021	4212.038	-0.017	
5	1	5	4	0	4	4422.569	4422.583	-0.014	
11	3	9	11	2	10	4494.510	4494.548	-0.037	
3	2	2	2	1	1	4706.349	4706.391	-0.042	
6	1	6	5	0	5	5094.968	5094.992	-0.023	
11	4	7	11	3	8	5182.565	5182.593	-0.028	
7	0	7	6	1	6	5196.266	5196.265	0.001	
10	4	6	10	3	7	5292.090	5292.120	-0.030	
7	4	3	7	3	4	5457.506	5457.492	0.013	
6	4	2	6	3	3	5477.627	5477.634	-0.007	
9	4	6	9	3	7	5480.170	5480.231	-0.060	
8	4	5	8	3	6	5480.504	5480.488	0.017	
7	4	4	7	3	5	5483.948	5483.941	0.008	
10	4	7	10	3	8	5486.092	5486.119	-0.027	
6	4	3	6	3	4	5488.393	5488.413	-0.021	
7	1	7	6	0	6	5768.158	5768.188	-0.030	
8	0	8	7	1	7	6030.426	6030.432	-0.006	
9	1	8	8	2	7	6107.632	6107.635	-0.004	
5	2	4	4	1	3	6117.528	6117.527	0.002	
10	1	10	9	0	9	7849.698	7849.729	-0.031	
5	3	3	4	2	2	7967.294	7967.281	0.013	
5	3	2	4	2	3	8024.162	8024.153	0.009	
11	0	11	10	1	10	8430.105	8430.104	0.001	
11	1	11	10	0	10	8568.445	8568.446	-0.001	
6	3	4	5	2	3	8736.614	8736.611	0.004	
6	3	3	5	2	4	8869.262	8869.279	-0.017	
12	0	12	11	1	11	9205.201	9205.193	0.008	
12	1	12	11	0	11	9296.898	9296.908	-0.010	
7	3	5	6	2	4	9474.335	9474.351	-0.015	
5	4	2	4	3	1	9570.488	9570.424	0.064	
5	4	1	4	3	2	9571.444	9571.383	0.061	
7	3	4	6	2	5	9737.330	9737.344	-0.015	
13	0	13	12	1	12	9972.817	9972.810	0.007	
13	1	13	12	0	12	10032.750	10032.740	0.010	
6	4	3	5	3	2	10382.428	10382.339	0.059	
8	3	5	7	2	6	10638.501	10638.489	0.012	
14	0	14	13	1	13	10735.188	10735.176	0.012	
5	5	0	4	4	1	11142.063	11142.066	-0.004	
7	4	4	6	3	3	11190.091	11190.072	0.019	

7	4	3	6	3	4	11201.7293	11201.6897	0.0396	
9	3	6	8	2	7	11584.9228	11584.9677	-0.0449	
8	4	5	7	3	4	11989.9685	11990.0064	-0.0379	
4	4	1	3	3	0	8756.5464	8756.5526	-0.0063	0.50
4	4	0	3	3	1	8756.5464	8756.5526	-0.0063	0.50

**Note.** Upper and lower state quantum numbers are indicated by ' and ", respectively. [a] Observed frequency. [b] Calculated frequency. [c] Observed minus calculated frequency.

**Table D2.2** List of the measured rotational transitions (frequencies and residuals, in MHz) of GIIa *L-DOPA*.

$J'$	$K_a'$	$K_c'$	$J''$	$K_a''$	$K_c''$	$\nu_{obs}(MHz)^{[a]}$	$\nu_{calc}(MHz)^{[b]}$	$\nu_{obs} - \nu_{calc}(MHz)^{[c]}$
5	1	4	4	1	3	3866.784	3866.764	0.020
6	1	6	5	1	5	4402.949	4402.910	0.039
6	2	5	5	2	4	4523.947	4523.925	0.021
6	2	4	5	2	3	4565.137	4565.133	0.004
6	1	5	5	1	4	4636.399	4636.412	-0.013
3	2	2	2	1	1	5046.165	5046.163	0.002
7	1	7	6	1	6	5132.830	5132.810	0.021
7	2	5	6	2	4	5339.232	5339.220	0.012
7	1	6	6	1	5	5403.804	5403.820	-0.016
8	1	8	7	1	7	5861.183	5861.192	-0.009
8	0	8	7	0	7	5945.754	5945.804	-0.050
8	2	7	7	2	6	6024.405	6024.395	0.009
8	3	6	7	3	5	6050.902	6050.903	-0.001
8	3	5	7	3	4	6056.741	6056.743	-0.002
8	1	7	7	1	6	6168.421	6168.435	-0.013

**Note.** Upper and lower state quantum numbers are indicated by ' and ", respectively. [a] Observed frequency. [b] Calculated frequency. [c] Observed minus calculated frequency.

**Table D2.3** List of the measured rotational transitions (frequencies and residuals, in MHz) of AIIb *L-DOPA*.

$J'$	$K_a'$	$K_c'$	$J''$	$K_a''$	$K_c''$	$\nu_{obs}(MHz)^{[a]}$	$\nu_{calc}(MHz)^{[b]}$	$\nu_{obs} - \nu_{calc}(MHz)^{[c]}$
6	1	6	5	1	5	3440.950	3440.936	0.014
7	1	7	6	1	6	4013.018	4013.017	0.002
7	0	7	6	0	6	4092.584	4092.638	-0.054
7	2	5	6	2	4	4133.744	4133.770	-0.025
7	1	6	6	1	5	4206.023	4206.001	0.022
8	1	8	7	1	7	4584.512	4584.496	0.016
8	2	7	7	2	6	4697.721	4697.699	0.022
8	2	6	7	2	5	4730.569	4730.596	-0.027
8	1	7	7	1	6	4804.654	4804.661	-0.007
9	1	9	8	1	8	5155.334	5155.320	0.014
9	0	9	8	0	8	5243.947	5243.954	-0.007
9	2	8	8	2	7	5283.224	5283.217	0.007
9	1	8	8	1	7	5402.369	5402.365	0.004
10	1	10	9	1	9	5725.457	5725.449	0.008
10	0	10	9	0	9	5815.356	5815.376	-0.020
10	2	8	9	2	7	5930.942	5930.946	-0.005
10	1	9	9	1	8	5998.977	5998.948	0.029

**Note.** Upper and lower state quantum numbers are indicated by ' and ", respectively. [a] Observed frequency. [b] Calculated frequency. [c] Observed minus calculated frequency.

**Table D2.4** List of the measured rotational transitions (frequencies and residuals, in MHz) of A1a *L-DOPA*.

$J'$	$K_a'$	$K_c'$	$J''$	$K_a''$	$K_c''$	$\nu_{obs}(\text{MHZ})^{[a]}$	$\nu_{calc}(\text{MHZ})^{[b]}$	$\nu_{obs} - \nu_{calc}(\text{MHZ})^{[c]}$	weight
3	3	1	2	2	0	9045.121	9045.197	-0.076	0.5
3	3	0	2	2	1	9045.121	9045.197	-0.076	0.5
16	1	16	15	1	15	9145.543	9145.578	-0.035	
15	1	15	14	1	14	8578.190	8578.141	0.049	
4	3	2	3	2	1	9631.300	9631.255	0.045	
4	3	1	3	2	2	9632.932	9632.887	0.044	
5	3	3	4	2	2	10216.261	10216.237	0.023	
5	3	2	4	2	3	10221.172	10221.149	0.022	
6	3	4	5	2	3	10799.043	10799.092	-0.049	
6	3	3	5	2	4	10810.579	10810.590	-0.011	
11	0	11	10	0	10	6388.454	6388.445	0.009	
11	1	11	10	1	10	6302.151	6302.174	-0.023	
9	1	8	8	1	7	5384.632	5384.627	0.005	
9	2	7	8	2	6	5313.780	5313.795	-0.015	
8	1	8	7	1	7	4588.694	4588.681	0.012	
7	1	6	6	1	5	4191.372	4191.365	0.006	
7	1	7	6	1	6	4016.375	4016.389	-0.013	

**Note.** Upper and lower state quantum numbers are indicated by ' and ', respectively. [a] Observed frequency. [b] Calculated frequency. [c] Observed minus calculated frequency.

**Table D2.5** List of the measured hyperfine components (frequencies and residuals, in MHz) of G1a *L-DOPA* using the LA-MB-FTMW technique.

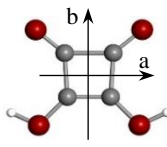
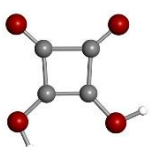
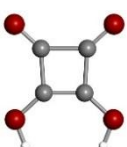
$J'$	$K_a'$	$K_c'$	$F'$	$J''$	$K_a''$	$K_c''$	$F''$	$\nu_{obs}(\text{MHZ})^{[a]}$	$\nu_{calc}(\text{MHZ})^{[b]}$	$\nu_{obs} - \nu_{calc}(\text{MHZ})^{[c]}$
6	1	6	7	5	0	5	6	5094.947	5094.953	-0.006
6	1	6	6	5	0	5	5	5095.059	5095.059	0.000
3	2	2	3	2	1	1	2	4706.437	4706.442	-0.005
3	2	2	4	2	1	1	3	4706.384	4706.383	0.001
5	1	5	6	4	0	4	5	4422.539	4422.538	0.002
5	1	5	5	4	0	4	4	4422.662	4422.659	0.003
4	1	4	5	3	0	3	4	3741.780	3741.781	-0.001
4	1	4	4	3	0	3	3	3741.913	3741.909	0.004
2	2	1	3	1	1	0	2	3954.291	3954.291	-0.000
2	2	1	2	1	1	0	2	3954.071	3954.069	0.002
2	2	1	1	1	1	0	0	3954.022	3954.021	0.001

**Note.** Upper and lower state quantum numbers are indicated by ' and ', respectively. [a] Observed frequency. [b] Calculated frequency. [c] Observed minus calculated frequency.

### D.3 Theoretical benchmark of squaric acid

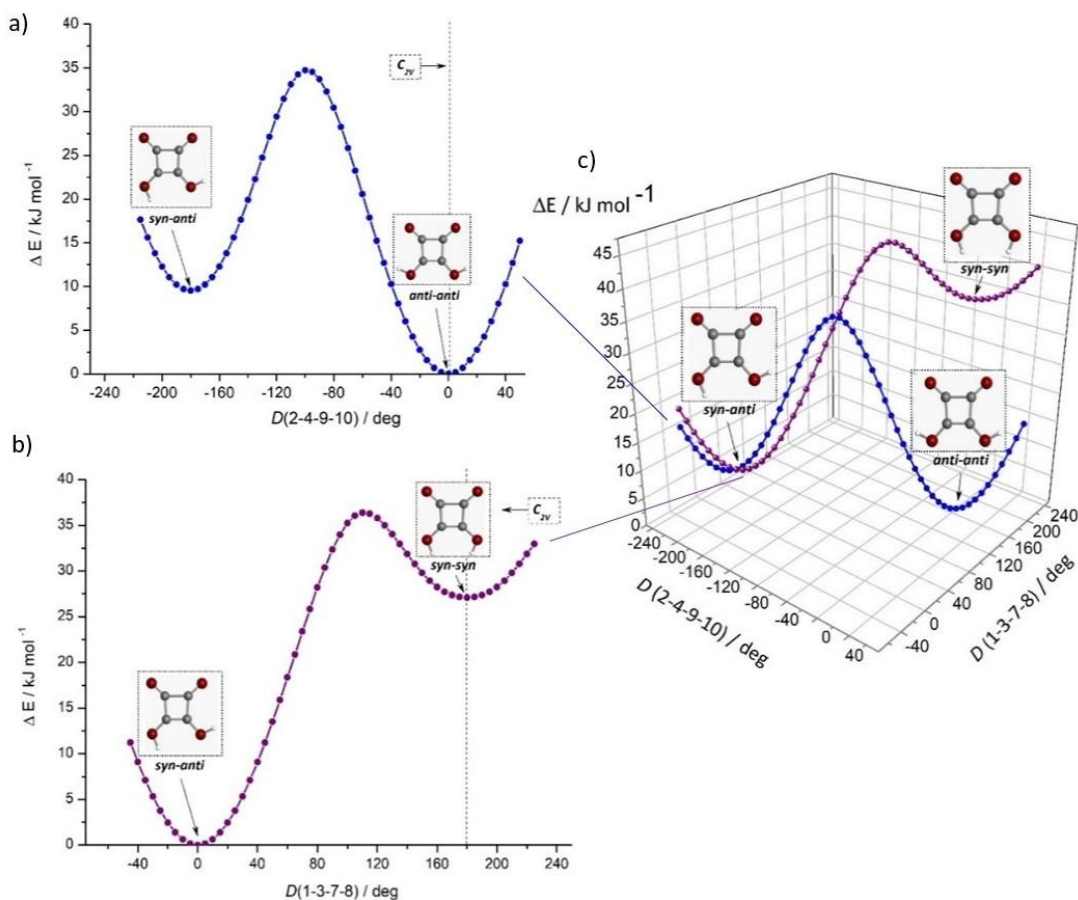
A conformational search was carried out on all plausible configurations of SA and its water clusters. A set of three candidate structures was identified by MMFFs force field (Halgren, 1996) for squaric acid and six different conformers in case of the SA-water clusters. These structures were then optimized employing the density functional theory B3LYP method with the Becke's three-parameter exchange functional and the Lee-Yang-Parr correlation functional (Becke, 1988; Lee et al. 1988), where also Grimme's D2 dispersion interactions were included (Grimme et al. 2010). Both Pople's 6-311++G (d,p) and Dunning's aug-cc-pVTZ basis set were applied in these calculations. At the respective optimized geometries, harmonic vibrational frequencies were calculated in order to apply zero-point-energy corrections to the electronic energy. Furthermore, two relaxed potential-energy scans were calculated at the B3LYP/6-311++G(d, p) level, choosing CCOH *syn*-OH and CCOH *anti*-OH torsion, respectively, as the driving coordinate (Figure S1). All the structural parameters were freely optimized, except both torsion angles, which were explored with a step of 5°.

**Table D3.1** Theoretical relative electronic ( $E$ ), zero-point vibrational ( $E_{ZPE}$ ), and total energies ( $\Delta E_{TOTAL} = E + E_{ZPE}$ ) for Squaric Acid conformers.

			
	<i>anti-anti</i>	<i>syn-anti</i>	<i>syn-syn</i>
<sup>[a]</sup> $\Delta E/\Delta E_{TOTAL}^{B3LYP/6-311++G(d,p)}$	0/0	9.6/9.2	36.7/33.9
$\Delta E/\Delta E_{TOTAL}^{B3LYP/aug-cc-pVTZ}$	0/0	10.1/9.4	34.9/32.3
$\Delta E/\Delta E_{TOTAL}^{MP2/6-311++G(d,p)}$	0/0	8.5/9.0	35.0/32.3
$\Delta E/\Delta E_{TOTAL}^{MP2/aug-cc-pVTZ}$	0/0	9.8/9.2	34.7/32.1
$\Delta E/\Delta E_{TOTAL}^{B2PLYPD/6-311++G(d,p)}$	0/0	9.2/8.9	36.2/32.2
$\Delta E/\Delta E_{TOTAL}^{B2PLYPD/aug-cc-pVTZ}$	0/0	10.1/9.4	34.9/32.2
<sup>[b]</sup> $\Delta E/\Delta E_{TOTAL}^{MP4/6-311++G(d,p)}$	0/0	6.9/6.4	31.6/29.3
$\Delta E/\Delta E_{TOTAL}^{CCSD/6-311++G(d,p)}$	0/0	6.9/6.4	31.9/29.6
$\Delta E/\Delta E_{TOTAL}^{QCISD/6-311++G(d,p)}$	0/0	7.1/6.6	32.1/29.8

[a] Electronic ( $E$ ) and zero-point vibrational ( $E_{ZPE}$ ) energies (in kJ/mol) calculated using the corresponding theoretical methods at geometries optimized at the same level of theory; [b] Electronic ( $E$ ) energies calculated elsewhere (Schwabe et al. 2007) using the corresponding theoretical methods at geometries optimized at the same level of theory and zero-point vibrational ( $E_{ZPE}$ ) calculated only at the B3LYP/6-311++G(d,p) level for all computations. Principal axes (a and b) are illustrated in for the *anti-anti* SA.

Finally, the geometries of the three SA conformers were reoptimized at higher levels of theory, including *ab initio* calculations at the MP2 level (Møller–Plesset second order perturbation theory) [7] and an advanced DFT method based on a double-hybrid density functional (B2PLYPD) with long-range dispersion corrections, [8] where both 6-311++G (d, p) and aug-cc-pVTZ basis sets were used (see Table S1.1). All quantum-mechanical calculations were carried out with the GAUSSIAN 16 program. [9]



**Figure D1.** The isomerization potentials between the *syn-anti* / *anti-anti* and *syn-anti* / *syn-syn* SA conformers, respectively B3LYP/6-311++G(d, p). (a) Results of  $D(2-4-9-10)$  scan which corresponds to the *syn*-OH torsion isomerization path. (b) Results of  $D(1-3-7-8)$  scan which corresponds to the *anti*-OH torsion isomerization path. (c) Three dimensional potential curves of (a) and (b) along the  $D(2-4-9-10)$  and  $D(1-3-7-8)$  coordinates.



**Table D3.2.** Theoretical spectroscopic parameters for the conformers of Squaric Acid.

<b>anti-anti</b>	<b>B3LYP</b> <sup>[a]</sup> <i>6-311++G(d,p)</i>	<b>B3LYP</b> <i>aug-cc-pVTZ</i>	<b>MP2</b> <i>6-311++G(d,p)</i>	<b>MP2</b> <i>aug-cc-pVTZ</i>	<b>B2PLYPD</b> <i>6-311++G(d,p)</i>	<b>B2PLYPD</b> <i>aug-cc-pVTZ</i>
<i>A</i> <sup>[b]</sup>	2583.49	2595.24	2564.10	2588.44	2583.23	2598.76
<i>B</i>	2383.70	2384.00	2375.97	2375.57	2377.97	2379.63
<i>C</i>	1239.79	1242.57	1233.23	1238.72	1238.18	1242.19
$ \mu_a / \mu_b / \mu_c $ <sup>[c]</sup>	0.0 / 3.5 / 0.0	0.0 / 3.7 / 0.0	0.0 / 3.1 / 0.0	0.0 / 3.4 / 0.0	0.0 / 3.4 / 0.0	0.0 / 3.6 / 0.0
<b>syn-anti</b>	<b>B3LYP</b> <i>6-311++G(d,p)</i>	<b>B3LYP</b> <i>aug-cc-pVTZ</i>	<b>MP2</b> <i>6-311++G(d,p)</i>	<b>MP2</b> <i>aug-cc-pVTZ</i>	<b>B2PLYPD</b> <i>6-311++G(d,p)</i>	<b>B2PLYPD</b> <i>aug-cc-pVTZ</i>
<i>A</i>	2500.50	2512.94	2490.00	2507.94	2501.21	2512.91
<i>B</i>	2444.73	2444.64	2429.86	2433.65	2437.17	2442.01
<i>C</i>	1236.15	1239.16	1229.78	1235.12	1234.39	1238.48
$ \mu_a / \mu_b / \mu_c $	2.0 / 5.6 / 0.0	1.8 / 5.7 / 0.0	2.1 / 5.2 / 0.0	1.9 / 5.4 / 0.0	2.0 / 5.5 / 0.0	1.8 / 5.6 / 0.0
<b>syn-syn</b>	<b>B3LYP</b> <i>6-311++G(d,p)</i>	<b>B3LYP</b> <i>aug-cc-pVTZ</i>	<b>MP2</b> <i>6-311++G(d,p)</i>	<b>MP2</b> <i>aug-cc-pVTZ</i>	<b>B2PLYPD</b> <i>6-311++G(d,p)</i>	<b>B2PLYPD</b> <i>aug-cc-pVTZ</i>
<i>A</i>	2495.21	2495.93	2499.10	2501.84	2496.05	2496.16
<i>B</i>	2433.50	2444.44	2404.95	2423.23	2426.37	2440.29
<i>C</i>	1231.99	1239.16	1225.56	1230.95	1230.36	1233.95
$ \mu_a / \mu_b / \mu_c $	0.0 / 7.7 / 0.0	0.0 / 7.5 / 0.0	0.0 / 7.3 / 0.0	0.0 / 7.4 / 0.0	0.0 / 7.6 / 0.0	0.0 / 7.5 / 0.0

[a] Theoretical calculations at B3LYP, MP2 and B2PLYPD levels of theory, respectively [b] *A*, *B*, and *C* represent the rotational constants (in MHz); [c]  $\mu_a$ ,  $\mu_b$ , and  $\mu_c$  are the electric dipole moment components (in D).

#### D.4 Topological Analysis of Chemical Bonding [SA and (SA-water) clusters]

Bader's Theory, also known as Atoms in Molecules (AIM) theory is the physical-mathematical formalism based on matrix algebra which allows to describe concepts like atoms (in a molecule), chemical bonds, and the molecular structure and stability by a topological analysis of the electron density  $\rho(r)$ . Different topological properties of  $\rho(r)$  can be explained looking at the number and class of the critical points, points where its gradient becomes zero ( $\nabla\rho(r) = 0$ ). These can be classified by the pair  $(\omega, \sigma)$ , being  $\omega$  the rank (number of curvatures of different from zero in the critical point) and  $\sigma$  the signature (the difference between the positive and negative curvatures). The most important points, from a chemical point of view, are: (a) (3, -3): Local maximum of the electro density related to the nuclear positions; (b) (3, -1): Minimum along an axis and maximum along two axes and are known as bond critical points (BCPs), which mathematically characterize the existence of a chemical bond between two atoms; (c) (3, +1): Minimum along two axes and maximum along an axis and are known as ring critical points (RCPs), which are related to the presence of a cyclic structure; (d) (3, +3): Local minimum known as cage critical points (CCPs).

In parallel, this theory allows us to classify the distinct interactions and possible bonds between atoms by analyzing several topological parameters which evaluate different properties in the BCPs: the Laplacian [ $\nabla^2\rho(\mathbf{r})$ ], the relationship between the potential energy density  $V(\mathbf{r})$  and the lagrangian form of kinetic energy  $G(\mathbf{r})$  and the total energy density  $H(\mathbf{r})$ .

**Table D4.1** Local topological properties <sup>[a]</sup> (in a.u.) of the electronic charge density distribution calculated at the position of the bond critical points for the different Squaric Acid conformers. The level of calculation is B3LYP/6-311++G(d,p).

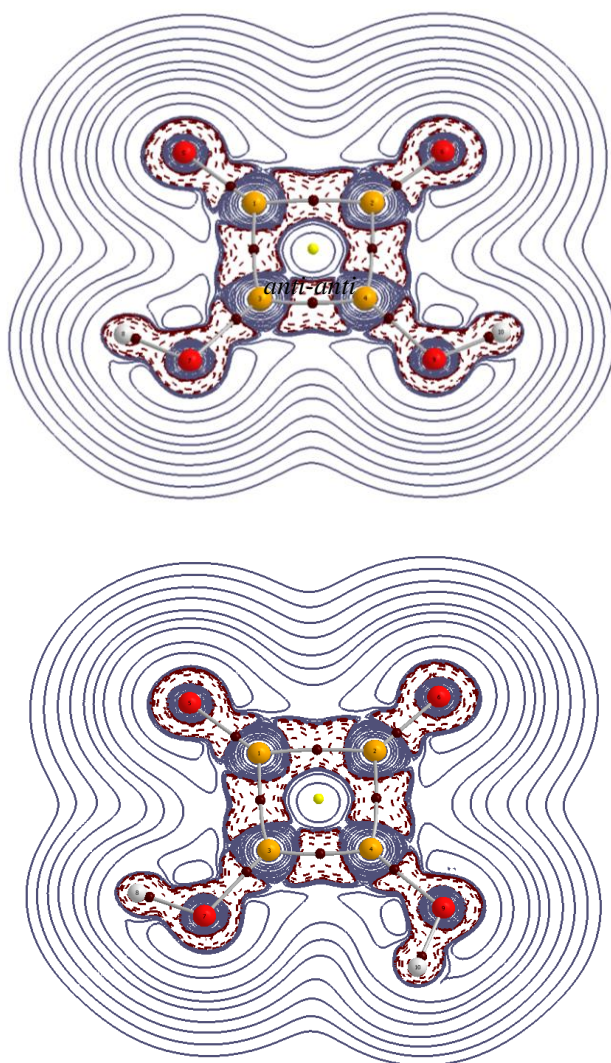
Specie	Bond	$\rho(\mathbf{r})$	$\nabla^2\rho(\mathbf{r})$	$ V(\mathbf{r}) /G(\mathbf{r})$	$H(\mathbf{r})$
<i>anti-anti</i> ( <sup>1</sup> A')	C <sub>1</sub> – C <sub>2</sub>	0.243	-0.551	3.919	-0.210
	C <sub>1</sub> – C <sub>3</sub>	0.275	-0.749	4.101	-0.276
	C <sub>2</sub> – C <sub>4</sub>	0.275	-0.749	4.101	-0.276
	C <sub>3</sub> – C <sub>4</sub>	0.330	-1.080	4.096	-0.399
	C <sub>1</sub> – O <sub>5</sub>	0.413	-0.123	2.043	-0.740
	C <sub>2</sub> – O <sub>6</sub>	0.413	-0.123	2.043	-0.740
	C <sub>3</sub> – O <sub>7</sub>	0.316	-0.365	2.217	0.512
	C <sub>4</sub> – O <sub>9</sub>	0.316	-0.365	2.217	0.512
	O <sub>7</sub> – H <sub>8</sub>	0.354	-2.643	10.870	-0.735
	O <sub>9</sub> – H <sub>10</sub>	0.354	-2.643	10.870	-0.735
<i>syn-anti</i> ( <sup>1</sup> A')	C <sub>1</sub> – C <sub>2</sub>	0.240	-0.532	3.904	-0.203
	C <sub>1</sub> – C <sub>3</sub>	0.274	-0.739	4.065	-0.274
	C <sub>2</sub> – C <sub>4</sub>	0.267	-0.703	4.072	-0.261
	C <sub>3</sub> – C <sub>4</sub>	0.333	-1.097	4.064	-0.407
	C <sub>1</sub> – O <sub>5</sub>	0.420	-0,0522	2.017	-0.756
	C <sub>2</sub> – O <sub>6</sub>	0.425	-0.0179	2.006	-0.770
	C <sub>3</sub> – O <sub>7</sub>	0.309	-0.340	2.206	-0.498
	C <sub>4</sub> – O <sub>9</sub>	0.314	-0.337	2.199	-0.508
	O <sub>7</sub> – H <sub>8</sub>	0.354	-2.658	10.928	-0.739
	O <sub>9</sub> – H <sub>10</sub>	0.356	-2.655	10.703	-0.740

[a] The electronic charge density [ $\rho(\mathbf{r})$ ], the Laplacian [ $\nabla^2\rho(\mathbf{r})$ ], the relationship between the potential energy density  $V(\mathbf{r})$  and the lagrangian form of kinetic energy  $G(\mathbf{r})$  and the total energy density  $H(\mathbf{r})$ .

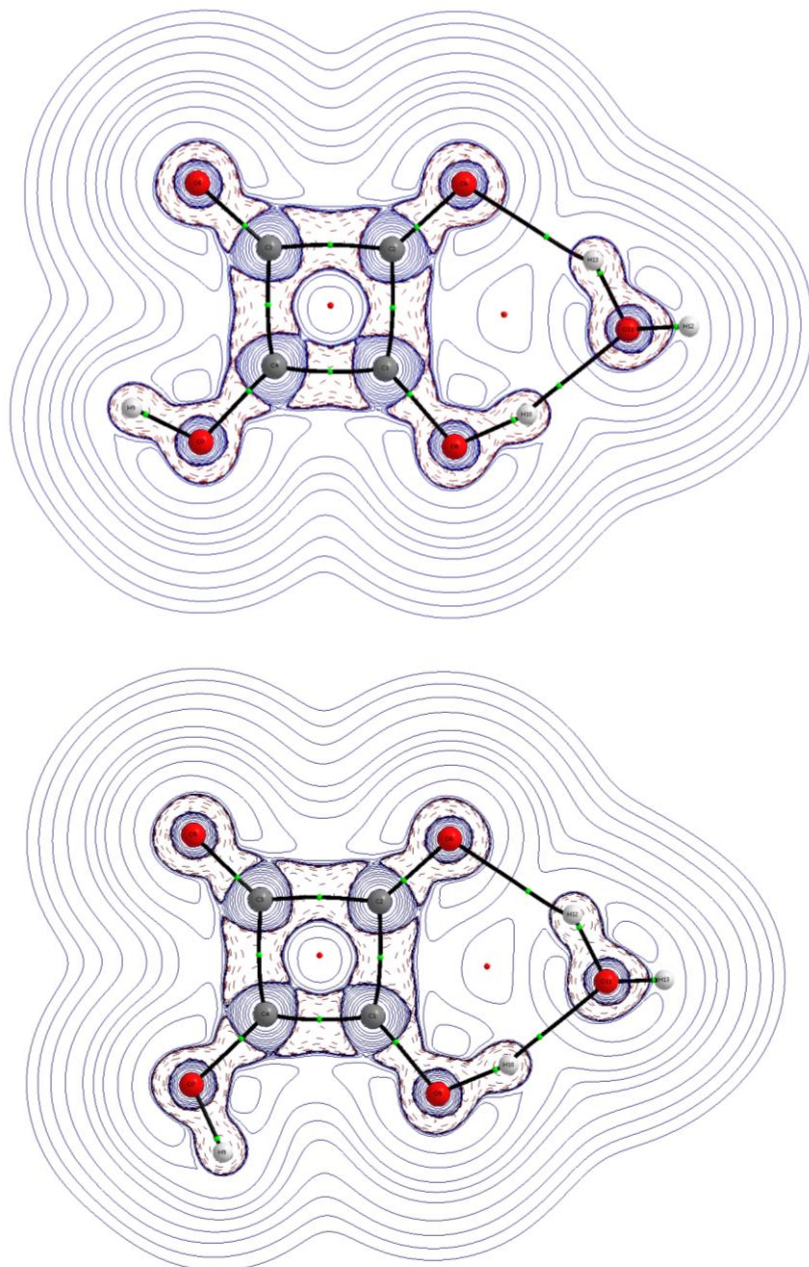
**Table D4.2.** Local topological properties <sup>[a]</sup> (in a.u.) of the electronic charge density distribution calculated at the position of the bond critical points for the SA-H<sub>2</sub>O clusters. The level of calculation is DFT(B3LYP)/6-311++G(d,p).

Specie	Bond	$\rho(\mathbf{r})$	$\nabla^2\rho(\mathbf{r})$	$ V(\mathbf{r}) /G(\mathbf{r})$	H( $\mathbf{r}$ )
<i>w-anti-anti</i> ( <sup>1</sup> A')	C <sub>1</sub> – C <sub>2</sub>	0.263	-0.590	4.023	-0.227
	C <sub>2</sub> – C <sub>3</sub>	0.245	-0.761	4.131	-0.196
	C <sub>1</sub> – C <sub>4</sub>	0.313	-0.858	4.013	-0.321
	C <sub>3</sub> – C <sub>4</sub>	0.270	-0.649	4.049	-0.241
	C <sub>3</sub> – O <sub>5</sub>	0.407	-0.133	2.051	-0.682
	C <sub>2</sub> – O <sub>6</sub>	0.399	-0.143	1.946	-0.663
	C <sub>4</sub> – O <sub>7</sub>	0.310	-0.300	2.195	0.458
	C <sub>1</sub> – O <sub>8</sub>	0.325	-0.371	2.233	0.490
	O <sub>7</sub> – H <sub>9</sub>	0.359	-2.543	11.780	-0.766
	O <sub>8</sub> – H <sub>10</sub>	0.325	-2.284	10.829	-0.635
	O <sub>11</sub> – H <sub>10</sub>	0.0494	0.142	1.145	-0.00607
	O <sub>11</sub> – H <sub>13</sub>	0.347	-2.456	11.218	-0.680
	O <sub>6</sub> – H <sub>13</sub>	0.0296	0.101	0.955	0.00107
O <sub>11</sub> – H <sub>12</sub>	0.365	-2.531	10.825	-0.704	
<i>w-syn-anti</i> ( <sup>1</sup> A')	C <sub>1</sub> – C <sub>2</sub>	0.265	-0.615	4.000	-0.231
	C <sub>2</sub> – C <sub>3</sub>	0.242	-0.522	4.124	-0.192
	C <sub>1</sub> – C <sub>4</sub>	0.314	-0.857	3.956	-0.324
	C <sub>3</sub> – C <sub>4</sub>	0.266	-0.633	4.091	-0.234
	C <sub>3</sub> – O <sub>5</sub>	0.412	-0.0980	2.037	-0.693
	C <sub>2</sub> – O <sub>6</sub>	0.399	-0.140	2.056	-0.663
	O <sub>7</sub> – H <sub>9</sub>	0.360	-2.542	11.608	-0.702
	C <sub>4</sub> – O <sub>7</sub>	0.309	-0.296	2.194	-0.456
	C <sub>1</sub> – O <sub>8</sub>	0.319	-0.371	2.241	-0.477
	O <sub>8</sub> – H <sub>10</sub>	0.323	-2.278	10.822	-0.634
	O <sub>11</sub> – H <sub>10</sub>	0.0493	0.143	1.146	-0.00609
	O <sub>11</sub> – H <sub>13</sub>	0.365	-2.532	10.830	-0.704
	O <sub>6</sub> – H <sub>12</sub>	0.0300	0.102	0.961	0.000955
O <sub>11</sub> – H <sub>12</sub>	0.346	-2.446	11.202	0.680	

[a] The electronic charge density [ $\rho(\mathbf{r})$ ], the Laplacian [ $\nabla^2\rho(\mathbf{r})$ ], the relationship between the potential energy density  $V(\mathbf{r})$  and the lagrangian form of kinetic energy  $G(\mathbf{r})$  and the total energy density  $H(\mathbf{r})$ .



**Figure D4.1** Molecular graphs and contour maps of the laplacian of the electronic charge density,  $\nabla^2\rho$ , for SA conformers. Red lines correspond to negative values of the laplacian and blue lines to positive ones. The contour maps were calculated at the MP2/aug-cc-pVTZ level of theory. **Color code:** orange / C, white / H, red / O, blue / N, black / bond critical points, Brown / BCPs, yellow / RCPs.



**Figure D4.2** Molecular graphs and contour maps of the laplacian of the electronic charge density,  $\nabla^2\rho$ , for SA-water clusters. Red lines correspond to negative values of the laplacian and blue lines to positive ones. The contour maps were calculated at the MP2/aug-cc-pVTZ level of theory. **Color code:** grey / C, white / H, red / O, green / BCPs, orange/ RCPs.

---

*D.5 Measured transitions in the rotational spectrum of squaric acid*

**Table D5.1** Observed frequencies and residuals (in MHz) for the rotational transitions of *anti-anti* SA.

$J'$	$K_a'$	$K_c'$	$J''$	$K_a''$	$K_c''$	$\nu_{obs}$	$\nu_{obs} - \nu_{calc}$
4	1	3	4	0	4	8657.895	0.001
3	0	3	2	1	2	8665.259	0.003
3	1	3	2	0	2	8692.070	0.002
2	2	1	1	1	0	8997.170	-0.002
3	1	2	2	2	1	10743.016	0.008
2	2	0	1	1	1	11103.431	0.018
4	0	4	3	1	3	11156.197	-0.000
4	1	4	3	0	3	11158.240	0.002
3	2	2	2	1	1	11476.332	0.004
5	1	4	5	0	5	11147.456	-0.007
5	2	4	5	1	5	11148.551	-0.005
4	1	3	3	2	2	13577.278	-0.025
4	2	3	3	1	2	13712.714	0.002

**Note.**  $\nu_{obs}$  is the observed frequency and  $\nu_{obs} - \nu_{calc}$  is the residual.

**Table D5.2** Observed frequencies and residuals (in MHz) for the rotational transitions of *syn-anti* SA.

$J'$	$K_a'$	$K_c'$	$J''$	$K_a''$	$K_c''$	$\nu_{obs}$	$\nu_{obs} - \nu_{calc}$
2	1	1	1	1	0	8566.547	-0.0113
3	0	3	2	1	2	8651.461	-0.006
3	1	3	2	0	2	8653.614	0.012
2	2	1	1	1	0	8742.640	0.008
3	1	2	2	2	1	11028.083	-0.025
2	2	0	1	1	1	11101.103	-0.004
4	0	4	3	1	3	11124.400	-0.024
3	1	2	2	1	1	11204.167	-0.014
3	2	2	2	1	1	11214.622	0.019
5	0	5	4	1	4	13596.408	0.005
5	1	5	4	0	4	13596.408	0.004
4	2	3	3	1	2	13603.644	0.008
3	3	1	2	2	0	13775.593	-0.010

**Note.**  $\nu_{obs}$  is the observed frequency and  $\nu_{obs} - \nu_{calc}$  is the residual.

**Table D5.3** Observed frequencies and residuals (in MHz) for the rotational transitions of *anti-anti*  $^{13}\text{C}=\text{O}$  SA.

$J'$	$K_a'$	$K_c'$	$J''$	$K_a''$	$K_c''$	$\nu_{obs}$	$\nu_{obs} - \nu_{calc}$
3	0	3	2	1	2	8638.983	-0.031
2	2	1	1	1	0	8968.560	0.007
3	1	2	2	2	1	10712.196	0.003
3	2	2	2	1	1	11440.174	0.005
5	0	5	4	1	4	13594.704	0.017

**Note.**  $\nu_{obs}$  is the observed frequency and  $\nu_{obs} - \nu_{calc}$  is the residual.

**Table D5.4** Observed frequencies and residuals (in MHz) for the rotational transitions of *anti-anti* <sup>13</sup>C-OH SA.

J'	K <sub>a</sub> '	K <sub>c</sub> '	J''	K <sub>a</sub> ''	K <sub>c</sub> ''	<i>v</i> <sub>obs</sub>	<i>v</i> <sub>obs</sub> - <i>v</i> <sub>calc</sub>
3	0	3	2	1	2	8646.193	0.014
3	1	3	2	0	2	8672.992	0.016
3	1	2	2	2	1	10718.869	0.000
2	2	0	1	1	1	11078.767	-0.000
4	0	4	3	1	3	11131.653	-0.025
4	1	4	3	0	3	11133.696	-0.024

**Note.** *v*<sub>obs</sub> is the observed frequency and *v*<sub>obs</sub> - *v*<sub>calc</sub> is the residual.

**Table D5.5** Observed frequencies and residuals (in MHz) for the rotational transitions of *anti-anti* C-<sup>18</sup>OH SA.

J'	K <sub>a</sub> '	K <sub>c</sub> '	J''	K <sub>a</sub> ''	K <sub>c</sub> ''	<i>v</i> <sub>obs</sub>	<i>v</i> <sub>obs</sub> - <i>v</i> <sub>calc</sub>
3	0	3	2	1	2	8448.675	0,003
3	1	3	2	0	2	8482.854	0,003
4	0	4	3	1	3	10881.771	0,025
4	1	4	3	0	3	10884.781	0,025

**Note.** *v*<sub>obs</sub> is the observed frequency and *v*<sub>obs</sub> - *v*<sub>calc</sub> is the residual.

**Table D5.6** Observed frequencies and residuals (in MHz) for the rotational transitions of *anti-anti* C=<sup>18</sup>O SA.

J'	K <sub>a</sub> '	K <sub>c</sub> '	J''	K <sub>a</sub> ''	K <sub>c</sub> ''	<i>v</i> <sub>obs</sub>	<i>v</i> <sub>obs</sub> - <i>v</i> <sub>calc</sub>
3	0	3	2	1	2	8439.618	0.000
3	1	3	2	0	2	8476.785	0.000
4	0	4	3	1	3	10871.834	0.000
4	1	4	3	0	3	10875.186	0.000

**Note.** *v*<sub>obs</sub> is the observed frequency and *v*<sub>obs</sub> - *v*<sub>calc</sub> is the residual.

**Table D5.5** Observed frequencies and residuals (in MHz) for the rotational transitions of *w-anti-anti* SA.

J'	K <sub>a</sub> '	K <sub>c</sub> '	J''	K <sub>a</sub> ''	K <sub>c</sub> ''	<i>v</i> <sub>obs</sub>	<i>v</i> <sub>obs</sub> - <i>v</i> <sub>calc</sub>
4	1	4	3	1	3	7023.480	0.010
5	1	4	4	2	3	7415.826	0.026
5	2	4	5	1	5	7584.227	-0.009
4	1	4	3	0	3	7749.889	0.013
2	2	1	1	1	0	8413.281	0.005
5	0	5	4	1	4	8466.161	0.001
2	2	0	1	1	1	8844.063	0.006
5	1	5	4	0	4	9133.062	-0.000
5	2	4	4	2	3	9690.716	-0.005

3	2	2	2	1	1	10013.281	0.010
6	0	6	5	1	5	10230.157	-0.002
6	1	6	5	0	5	10579.890	-0.013
5	2	3	4	2	2	10618.650	-0.022
4	2	3	3	1	2	11423.953	0.001
7	0	7	6	1	6	11918.175	-0.013
7	1	7	6	0	6	12089.965	-0.001
5	2	4	4	1	3	12666.841	0.008

**Note.**  $\nu_{obs}$  is the observed frequency and  $\nu_{obs} - \nu_{calc}$  is the residual.

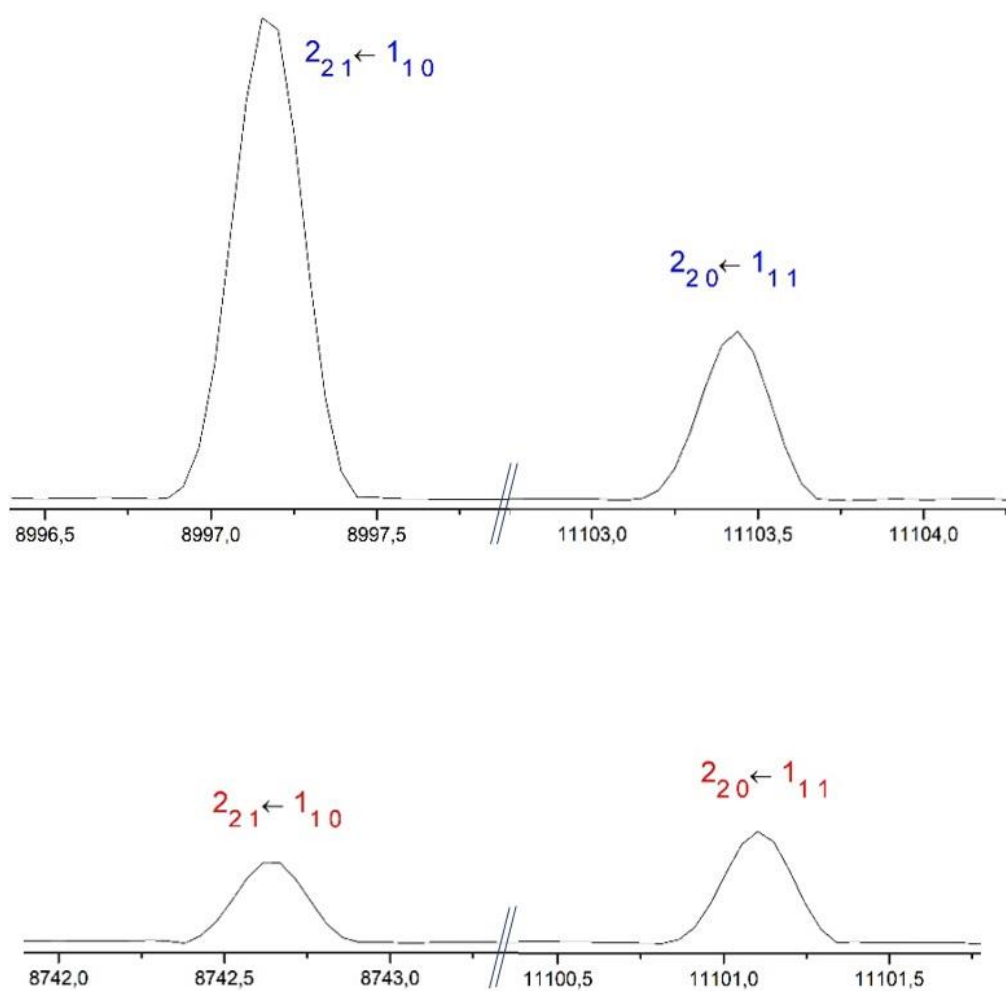
**Table D5.6** Observed frequencies and residuals (in MHz) for the rotational transitions of *w-syn-anti* SA.

$J'$	$K_a'$	$K_c'$	$J''$	$K_a''$	$K_c''$	$\nu_{obs}$	$\nu_{obs} - \nu_{calc}$
4	1	4	3	1	3	7078.268	0.021
4	0	4	3	0	3	7348.835	0.031
4	1	4	3	0	3	7692.119	-0.011
2	2	1	1	1	0	8174.371	-0.007
4	2	2	3	2	1	8559.595	-0.020
4	1	3	3	1	2	8577.340	-0.017
5	0	5	4	1	4	8584.827	0.018
2	2	0	1	1	1	8640.051	-0.013
5	1	5	4	1	4	8757.076	-0.005
5	0	5	4	0	4	8928.143	0.008
3	2	2	2	1	1	9781.672	0.007
6	0	6	5	1	5	10325.957	0.029
6	1	6	5	0	5	10578.447	-0.025
3	2	1	2	1	2	11314.365	0.023
7	1	7	6	0	6	12115.874	-0.014
6	1	5	5	1	4	12295.325	-0.004
6	1	6	5	1	5	10406.188	-0.012

**Note.**  $\nu_{obs}$  is the observed frequency and  $\nu_{obs} - \nu_{calc}$  is the residual.



*D.6 Nuclear statistics*



**Figure D6.** Comparison of the statistical weights for the  $2_{20} \leftarrow 1_{11}$  (*ee-oo*) and  $2_{21} \leftarrow 1_{10}$  (*oe-oo*) transitions of *anti-anti* (blue) and *syn-anti* (red) SA, respectively.

*D.7 Measured transitions for the barbaralone study*

In Tables D7.1-D7.2 the complete list of measured transitions for 1- and 5-phe-BA is given.

**Table D7.1** Observed frequencies and residuals (in MHz) for the rotational transitions of 1-phe-BA.

$J'$	$K_a'$	$K_c'$	$J''$	$K_a''$	$K_c''$	$\nu_{obs}$	$\nu_{calc}$	$\nu_{obs} - \nu_{calc}$
2	0	2	1	0	1	1535.373	1535.397	-0.023
3	1	3	2	1	2	2290.397	2290.398	-0.001
3	0	3	2	0	2	2302.931	2302.943	-0.012
3	2	2	2	2	1	2303.186	2303.186	0.000
3	2	1	2	2	0	2303.437	2303.430	0.006
3	1	2	2	1	1	2315.901	2315.898	0.003
4	1	4	3	1	3	3053.794	3053.792	0.001
4	0	4	3	0	3	3070.307	3070.305	0.002
4	2	3	3	2	2	3070.835	3070.864	-0.029
4	3	1	3	3	0	3071.064	3071.036	0.027
4	2	2	3	2	1	3071.475	3071.476	-0.001
4	1	3	3	1	2	3087.792	3087.791	0.002
5	1	5	4	1	4	3817.128	3817.126	0.002
5	0	5	4	0	4	3837.424	3837.423	0.001
5	2	4	4	2	3	3838.498	3838.501	-0.003
5	3	3	4	3	2	3838.843	3838.846	-0.002
5	3	2	4	3	1	3838.843	3838.852	-0.008
5	2	3	4	2	2	3839.725	3839.723	0.002
5	1	4	4	1	3	3859.622	3859.620	0.002
6	1	6	5	1	5	4580.391	4580.385	0.007
6	0	6	5	0	5	4604.239	4604.237	0.001
6	2	5	5	2	4	4606.085	4606.084	0.000
6	2	4	5	2	3	4608.223	4608.222	0.001
6	1	5	5	1	4	4631.369	4631.369	0.000
7	1	7	6	1	6	5343.554	5343.555	-0.001
7	0	7	6	0	6	5370.688	5370.689	-0.001
7	2	6	6	2	5	5373.595	5373.604	-0.009
7	2	5	6	2	4	5377.013	5377.019	-0.006
7	1	6	6	1	5	5403.023	5403.019	0.003
8	1	8	7	1	7	6106.623	6106.624	-0.001
8	0	8	7	0	7	6136.723	6136.721	0.002
8	2	7	7	2	6	6141.049	6141.049	0.000
8	2	6	7	2	5	6146.161	6146.161	-0.000
8	1	7	7	1	6	6174.556	6174.554	0.002
1	1	0	0	0	0	1662.479	1662.471	0.008
8	2	7	7	3	5	1686.471	1686.462	0.009
8	2	6	7	3	4	1699.205	1699.207	-0.002
6	1	6	5	2	4	1844.917	1844.925	-0.007
6	1	5	5	2	3	2021.266	2021.262	0.004
2	1	1	1	0	1	2438.701	2438.700	0.001
9	2	8	8	3	6	2452.365	2452.370	-0.005
9	2	7	8	3	5	2472.339	2472.325	0.014
7	1	7	6	2	5	2582.383	2582.395	-0.012
7	1	6	6	2	4	2816.062	2816.060	0.002
5	0	5	4	1	3	2903.685	2903.679	0.006
3	1	2	2	0	2	3219.203	3219.201	0.002
10	2	8	9	3	6	3247.363	3247.363	0.000
8	1	8	7	2	6	3315.411	3315.415	-0.004
8	1	7	7	2	5	3613.644	3613.595	0.049

6 0 6 5 1 4	3648.299	3648.296	0.003
4 1 3 3 0 3	4004.053	4004.049	0.004
11 2 9 10 3 7	4024.695	4024.692	0.003
9 1 9 8 2 7	4043.939	4043.946	-0.007
2 2 0 1 1 0	4202.748	4202.744	0.004
2 2 1 1 1 1	4211.186	4211.183	0.003
7 0 7 6 1 5	4387.611	4387.616	-0.005
9 1 8 8 2 6	4413.398	4413.386	0.012
12 2 11 11 3 9	4745.202	4745.209	-0.007
10 1 10 9 2 8	4767.943	4767.950	-0.007
5 1 4 4 0 4	4793.371	4793.364	0.007
12 2 10 11 3 8	4804.684	4804.691	-0.007
3 2 1 2 1 1	4962.224	4962.217	0.007
3 2 2 2 1 2	4987.415	4987.411	0.004
8 0 8 7 1 6	5121.318	5121.317	0.001
10 1 9 9 2 7	5214.887	5214.890	-0.003
11 1 11 10 2 9	5487.377	5487.387	-0.011
4 2 2 3 1 2	5717.798	5717.795	0.003
4 2 3 3 1 3	5767.877	5767.877	-0.000
9 0 9 8 1 7	5849.040	5849.045	-0.005
11 1 10 10 2 8	6017.497	6017.504	-0.007
12 1 12 11 2 10	6202.204	6202.224	-0.020
7 1 6 6 0 6	6386.098	6386.092	0.006
5 2 3 4 1 3	6469.731	6469.727	0.004
5 2 4 4 1 4	6552.594	6552.586	0.008
12 2 11 12 1 11	2346.023	2346.040	-0.017
11 2 10 11 1 10	2395.278	2395.280	-0.002
10 2 9 10 1 9	2440.696	2440.702	-0.006
9 2 8 9 1 8	2482.220	2482.220	-0.001
8 2 7 8 1 7	2519.756	2519.762	-0.007
7 2 6 7 1 6	2553.264	2553.267	-0.003
6 2 5 6 1 5	2582.686	2582.682	0.004
5 2 4 5 1 4	2607.964	2607.968	-0.004
4 2 3 4 1 3	2629.087	2629.087	0.000
3 2 2 3 1 2	2646.013	2646.013	0.000
2 2 1 2 1 1	2658.727	2658.725	0.002
2 2 0 2 1 2	2684.291	2684.286	0.005
3 2 1 3 1 3	2697.319	2697.318	0.001
4 2 2 4 1 4	2715.006	2715.002	0.004
5 2 3 5 1 5	2737.602	2737.600	0.002
6 2 4 6 1 6	2765.438	2765.436	0.002
7 2 5 7 1 7	2798.898	2798.901	-0.002
7 1 6 6 2 4	2816.064	2816.060	0.004
8 2 6 8 1 8	2838.438	2838.438	0.000
9 2 7 9 1 9	2884.546	2884.547	-0.001
10 2 8 10 1 10	2937.778	2937.775	0.003
11 2 9 11 1 11	2998.711	2998.711	0.000
12 2 10 12 1 12	3067.964	3067.968	-0.004
13 2 11 13 1 13	3146.182	3146.182	0.000
14 2 12 14 1 14	3233.983	3233.989	-0.006
15 2 13 15 1 15	3332.014	3332.014	-0.000
16 2 14 16 1 16	3440.822	3440.856	-0.034
17 2 15 17 1 17	3561.060	3561.068	-0.008
8 1 7 7 2 5	3613.644	3613.595	0.049
18 2 16 18 1 18	3693.152	3693.151	0.001
21 3 19 21 2 19	4108.911	4108.897	0.015
18 3 16 18 2 16	4254.207	4254.227	-0.020
17 3 15 17 2 15	4291.692	4291.700	-0.008

16 3 14 16 2 14	4324.056	4324.054	0.002
15 3 13 15 2 13	4351.616	4351.601	0.015
14 3 12 14 2 12	4374.692	4374.700	-0.008
13 3 11 13 2 11	4393.741	4393.750	-0.009
12 3 10 12 2 10	4409.167	4409.173	-0.006
11 3 9 11 2 9	4421.397	4421.404	-0.007
10 3 8 10 2 8	4430.874	4430.879	-0.005
9 3 7 9 2 7	4438.014	4438.022	-0.008
8 3 6 8 2 6	4443.233	4443.236	-0.003
7 3 5 7 2 5	4446.895	4446.895	0.000
6 3 4 6 2 4	4449.344	4449.340	0.004
5 3 3 5 2 3	4450.877	4450.870	0.006
4 3 2 4 2 2	4451.752	4451.748	0.004
3 3 1 3 2 1	4452.183	4452.189	-0.006
3 3 0 3 2 2	4452.499	4452.495	0.003
4 3 1 4 2 3	4452.680	4452.667	0.013
5 3 2 5 2 4	4453.011	4453.018	-0.007
6 3 3 6 2 5	4453.621	4453.640	-0.019
7 3 4 7 2 6	4454.642	4454.646	-0.003
8 3 5 8 2 7	4456.158	4456.168	-0.010
9 3 6 9 2 8	4458.357	4458.362	-0.005
10 3 7 10 2 9	4461.406	4461.406	0.000
11 3 8 11 2 10	4465.450	4465.501	-0.001
12 3 9 12 2 11	4470.868	4470.875	-0.007
13 3 10 13 2 12	4477.774	4477.783	-0.009
14 3 11 14 2 13	4486.505	4486.508	-0.003
15 3 12 15 2 14	4497.397	4497.363	0.034
16 3 13 16 2 15	4510.693	4510.694	-0.001
14 3 12 13 4 10	4521.512	4521.533	-0.021
17 3 14 17 2 16	4526.886	4526.877	0.008
18 3 15 18 2 17	4546.327	4546.328	-0.001
19 4 16 19 3 16	6201.231	6201.234	-0.003
17 4 14 17 3 14	6213.947	6213.947	0.000
18 4 14 18 3 16	6223.477	6223.471	0.006
16 4 12 16 3 14	6226.127	6226.152	-0.025
14 4 10 14 3 12	6228.580	6228.560	0.020
12 4 9 12 3 9	6229.073	6229.072	0.001
13 4 9 13 3 11	6229.591	6229.585	0.006
10 4 6 10 3 8	6231.842	6231.830	0.014
8 4 5 8 3 5	6232.572	6232.551	0.021

**Note.**  $v_{obs}$  is the observed frequency and  $v_{obs} - v_{calc}$  is the residual

**Table D7.2** Observed frequencies and residuals (in MHz) for the rotational transitions of 5-phe-BA.

$J'$	$K_a'$	$K_c'$	$J''$	$K_a''$	$K_c''$	$\nu_{obs}$	$\nu_{calc}$	$\nu_{obs} - \nu_{calc}$
4	1	4	3	1	3	3079.967	3079.960	0.007
4	0	4	3	0	3	3103.769	3103.759	0.010
4	2	3	3	2	2	3104.876	3104.883	-0.007
4	2	2	3	2	1	3106.109	3106.102	0.007
4	1	3	3	1	2	3129.488	3129.507	-0.019
5	1	5	4	1	4	3849.727	3849.727	-0.000
5	0	5	4	0	4	3878.786	3878.786	-0.000
5	2	4	4	2	3	3880.951	3880.951	0.000
5	2	3	4	2	2	3883.387	3883.387	-0.000
5	1	4	4	1	3	3911.649	3911.651	-0.002
6	1	6	5	1	5	4619.345	4619.348	-0.003
6	0	6	5	0	5	4653.208	4653.211	-0.002
6	2	5	5	2	4	4656.920	4656.918	0.002
6	2	4	5	2	3	4661.171	4661.174	-0.003
6	1	5	5	1	4	4693.640	4693.634	0.006
7	1	7	6	1	6	5388.806	5388.799	0.007
7	0	7	6	0	6	5426.914	5426.919	-0.005
7	2	6	6	2	5	5432.757	5432.763	-0.006
7	2	5	6	2	4	5439.549	5439.555	-0.006
7	1	6	6	1	5	5475.416	5475.420	-0.004
8	1	8	7	1	7	6158.050	6158.058	-0.008
8	0	8	7	0	7	6199.796	6199.807	-0.011
8	2	7	7	2	6	6208.471	6208.465	0.006
8	3	6	7	3	5	6211.292	6211.326	-0.034
8	3	5	7	3	4	6211.532	6211.515	0.017
8	2	6	7	2	5	6218.655	6218.612	0.043
8	1	7	7	1	6	6256.970	6256.972	-0.002
2	1	1	1	0	1	2514.587	2514.599	-0.012
5	0	5	4	1	3	2871.815	2871.812	0.002
3	1	2	2	0	2	3309.474	3309.471	0.003
4	1	3	3	0	3	4110.734	4110.733	0.001
7	0	7	6	1	5	4346.662	4346.658	0.004
2	2	0	1	1	0	4377.002	4377.002	0.000
2	2	1	1	1	1	4389.267	4389.268	-0.001
6	3	4	6	2	4	4711.375	4711.414	-0.039
5	3	3	5	2	3	4714.464	4714.471	-0.007
4	3	2	4	2	2	4716.220	4716.225	-0.005
3	3	1	3	2	1	4717.156	4717.107	0.048
3	3	0	3	2	2	4717.710	4717.718	-0.008
4	3	1	4	2	3	4718.053	4718.059	-0.006
5	3	2	5	2	4	4718.774	4718.757	0.017
5	1	4	4	0	4	4918.624	4918.624	-0.000
3	2	1	2	1	1	5141.351	5141.346	0.005

**Note.**  $\nu_{obs}$  is the observed frequency and  $\nu_{obs} - \nu_{calc}$  is the residual

*D.8 <sup>13</sup>C isotopomers of 1-phe-BA*

During the analysis, five different <sup>13</sup>C isotopomers corresponding to 1-phe-BA have been unequivocally identified in the rotational spectrum, which correspond to the equivalent carbons (the C<sub>s</sub> symmetry) shown in Table 2. The rotational spectroscopic parameters derived from a semirigid rotor analysis are collected in Table S1.

**Table D8.1** Experimental spectroscopic constants of the <sup>13</sup>C isotopomers of 1-phe-BA (A-reduction, Ir-Representation).

Parameter	<b>1-phe-BA</b>	<b><sup>13</sup>C-2</b>	<b><sup>13</sup>C-3</b>	<b><sup>13</sup>C-4</b>	<b><sup>13</sup>C-8 (phe)</b>	<b><sup>13</sup>C-9 (phe)</b>
<i>A</i> <sup>[a]</sup> / MHz	1274.35677 (46) <sup>[e]</sup>	1267.2799 (12)	1264.3088 (25)	1269.6921 (27)	1269.6995 (31)	1269.7595 (23)
<i>B</i> / MHz	388.10352 (12)	387.37379 (56)	386.48315 (70)	385.63582 (53)	386.9353 (16)	384.7298 (14)
<i>C</i> / MHz	379.62560 (12)	379.19506 (42)	377.62267 (55)	376.86932 (46)	378.1018 (12)	376.0010 (15)
Δ <sub>JK</sub> / kHz	0.0742 (22)	0.0742 (fixed)	0.0742 (fixed)	0.0742 (fixed)	0.0742 (fixed)	0.0742 (fixed)
δ <sub>K</sub> / kHz	-5.519 (28)	-5.519 (fixed)	-5.519 (fixed)	-5.519 (fixed)	-5.519 (fixed)	-5.519 (fixed)
N	147	23	15	24	15	15
σ <sub>RMS</sub> <sup>[b]</sup>	10.7	15.8	15.0	15.2	15.0	15.5

<sup>[a]</sup> *A*, *B*, and *C* represent the rotational constants (in MHz);  $\mu_a$ ,  $\mu_b$ , and  $\mu_c$  are the electric dipole moment components (in D); <sup>[b]</sup> Number of measured transitions. <sup>[c]</sup> σ<sub>RMS</sub> is the root mean square deviation of the fit (in kHz). <sup>[d]</sup> Standard error in parentheses in units of the last digit. Note. The different carbon atoms are labelled following Figure 1.

**Table D8.2** Observed frequencies and residuals (in MHz) for the rotational transitions of the C<sub>2</sub><sup>13</sup>C isotopomer of 1-phe-BA.

J' K <sub>a</sub> ' K <sub>c</sub> '	J'' K <sub>a</sub> '' K <sub>c</sub> ''	<i>v</i> <sub>obs</sub>	<i>v</i> <sub>calc</sub>	<i>v</i> <sub>obs</sub> - <i>v</i> <sub>calc</sub>
7 0 7	6 0 6	5333.877	5333.878	-0.001
7 2 6	6 2 5	5336.969	5337.001	-0.033
7 2 5	6 2 4	5340.670	5340.659	0.010
8 0 8	7 0 7	6094.553	6094.567	-0.013
8 2 7	7 2 6	6099.216	6099.203	0.012
8 2 6	7 2 5	6104.659	6104.677	-0.019
6 2 4	5 2 3	4576.999	4577.010	-0.011
5 0 5	4 0 4	3811.235	3811.216	0.019
5 2 4	4 2 3	3812.405	3812.371	0.034
5 3 3	4 3 2	3812.743	3812.741	0.002
5 2 3	4 2 2	3813.688	3813.681	0.007
4 2 2	3 1 2	5689.946	5689.940	0.006
4 2 3	3 1 3	5741.706	5741.688	0.018
6 1 5	5 0 5	5557.027	5557.029	-0.002
7 1 7	6 1 6	5305.930	5305.921	0.009
3 1 2	2 0 2	3202.344	3202.344	0.001
6 0 6	5 1 4	3616.572	3616.567	0.004
5 1 5	4 1 4	3790.249	3790.266	-0.017
5 1 4	4 1 3	3834.202	3834.202	0.000
4 1 3	3 0 3	3982.556	3982.553	0.002
2 2 1	1 1 1	4194.688	4194.722	-0.034
7 0 7	6 1 5	4349.591	4349.591	0.000
5 1 4	4 0 4	4767.391	4767.390	0.001
3 2 2	2 1 2	4966.024	4966.014	0.010

**Note.** *v*<sub>obs</sub> is the observed frequency and *v*<sub>obs</sub> - *v*<sub>calc</sub> is the residual

**Table D8.3** Observed frequencies and residuals (in MHz) for the rotational transitions of the C<sub>3</sub><sup>13</sup>C isotopomer of 1-phe-BA.

J' K <sub>a</sub> ' K <sub>c</sub> '	J'' K <sub>a</sub> '' K <sub>c</sub> ''	<i>v</i> <sub>obs</sub>	<i>v</i> <sub>calc</sub>	<i>v</i> <sub>obs</sub> - <i>v</i> <sub>calc</sub>
4 2 3	3 1 3	5729.822	5729.822	-0.000
4 2 2	3 1 2	5677.544	5677.538	0.006
3 2 1	2 1 1	4926.096	4926.095	0.001
3 2 2	2 1 2	4952.398	4952.406	-0.008
2 2 1	1 1 1	4179.419	4179.419	0.000
4 1 4	3 1 3	3038.521	3038.522	-0.001
5 1 5	4 1 4	3798.038	3798.026	0.012
5 0 5	4 0 4	3819.176	3819.182	-0.006
5 2 4	4 2 3	3820.386	3820.370	0.016
5 1 4	4 1 3	3842.427	3842.432	-0.005
6 1 6	5 1 5	4557.399	4557.448	-0.049
5 1 4	4 0 4	4769.928	4769.929	-0.001
8 0 8	7 1 6	5094.107	5094.092	0.015
6 1 5	5 0 5	5561.487	5561.472	0.015
3 1 2	2 0 2	3201.250	3201.245	0.005

**Note.** *v*<sub>obs</sub> is the observed frequency and *v*<sub>obs</sub> - *v*<sub>calc</sub> is the residual

**Table D8.4** Observed frequencies and residuals (in MHz) for the rotational transitions of the C<sub>4</sub> <sup>13</sup>C isotopomer of 1-phe-BA.

J' K <sub>a</sub> ' K <sub>c</sub> '	J'' K <sub>a</sub> '' K <sub>c</sub> ''	<i>v<sub>obs</sub></i>	<i>v<sub>calc</sub></i>	<i>v<sub>obs</sub> - v<sub>calc</sub></i>
5 1 5	4 1 4	3812.064	3812.088	-0.024
5 0 5	4 0 4	3831.679	3831.696	-0.017
5 2 4	4 2 3	3832.718	3832.706	0.011
5 2 3	4 2 2	3833.853	3833.853	0.000
5 1 4	4 1 3	3853.075	3853.085	-0.010
6 1 5	5 1 4	4623.542	4623.538	0.004
7 0 7	6 0 6	5362.771	5362.776	-0.005
8 0 8	7 0 7	6127.782	6127.753	0.029
8 2 7	7 2 6	6131.809	6131.812	-0.003
8 1 8	7 1 7	6098.584	6098.611	-0.027
2 2 0	1 1 0	4181.074	4181.087	-0.013
2 2 1	1 1 1	4189.222	4189.230	-0.008
3 2 1	2 1 1	4939.665	4939.683	-0.018
3 2 2	2 1 2	4964.010	4963.998	0.012
6 1 5	5 0 5	5570.665	5570.632	0.033
4 2 2	3 1 2	5694.490	5694.513	-0.022
4 2 3	3 1 3	5742.850	5742.855	-0.005
9 0 9	8 1 7	5851.562	5851.540	0.022
5 3 2	5 2 4	4420.500	4420.513	-0.013
4 3 1	4 2 3	4420.174	4420.184	-0.010
3 3 1	3 2 1	4419.702	4419.736	-0.034
4 3 2	4 2 2	4419.323	4419.322	0.001
5 3 3	5 2 3	4418.495	4418.499	-0.004

**Note.** *v<sub>obs</sub>* is the observed frequency and *v<sub>obs</sub> - v<sub>calc</sub>* is the residual

**Table D8.5** Observed frequencies and residuals (in MHz) for the rotational transitions of the C<sub>8</sub> <sup>13</sup>C isotopomer of 1-phe-BA.

J' K <sub>a</sub> ' K <sub>c</sub> '	J'' K <sub>a</sub> '' K <sub>c</sub> ''	<i>v<sub>obs</sub></i>	<i>v<sub>calc</sub></i>	<i>v<sub>obs</sub> - v<sub>calc</sub></i>
5 0 5	4 0 4	3823.852	3823.854	-0.002
5 2 4	4 2 3	3825.038	3825.029	0.009
2 2 0	1 1 0	4187.246	4187.254	-0.008
2 2 1	1 1 1	4196.049	4196.043	0.006
5 1 5	4 1 4	3802.726	3802.753	-0.027
6 0 6	5 0 5	4587.930	4587.895	0.034
6 2 4	5 2 3	4592.215	4592.235	-0.020
3 2 1	2 1 1	4943.703	4943.701	0.002
7 1 7	6 1 6	5323.401	5323.397	0.004
7 0 7	6 0 6	5351.542	5351.541	0.001
7 2 5	6 2 4	5358.414	5358.435	-0.021
8 1 8	7 1 7	6083.560	6083.561	-0.001
8 0 8	7 0 7	6114.737	6114.732	0.005
8 2 7	7 2 6	6119.462	6119.445	0.016
8 2 6	7 2 5	6125.005	6125.010	-0.005

**Note.** *v<sub>obs</sub>* is the observed frequency and *v<sub>obs</sub> - v<sub>calc</sub>* is the residual



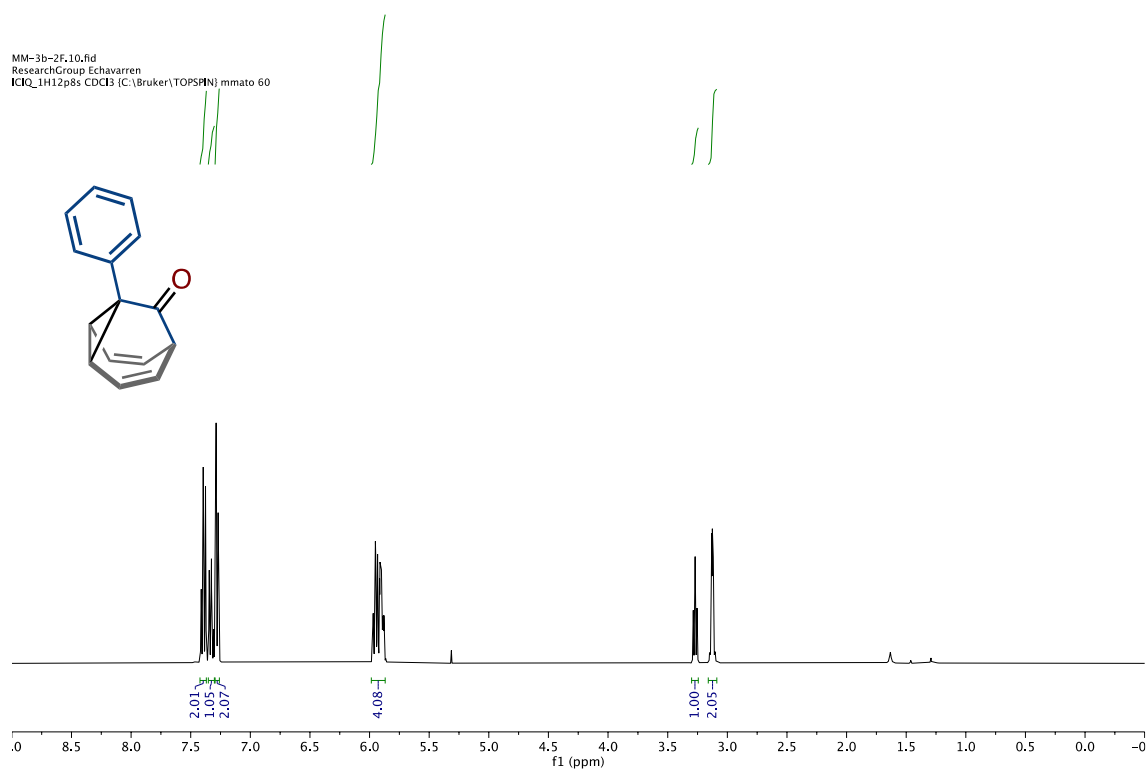
**Table D8.6** Observed frequencies and residuals (in MHz) for the rotational transitions of the C<sub>9</sub> <sup>13</sup>C isotopomer of 1-phe-BA.

J'	K <sub>a</sub> '	K <sub>c</sub> '	J''	K <sub>a</sub> ''	K <sub>c</sub> ''	<i>v</i> <sub>obs</sub>	<i>v</i> <sub>calc</sub>	<i>v</i> <sub>obs</sub> - <i>v</i> <sub>calc</sub>
5	0	5	4	0	4	3802.333	3802.357	-0.024
5	2	4	4	2	3	3803.514	3803.501	0.013
5	2	3	4	2	2	3804.818	3804.798	0.020
2	2	0	1	1	0	4185.319	4185.332	-0.013
2	2	1	1	1	1	4194.029	4194.018	0.011
6	0	6	5	0	5	4562.122	4562.117	0.004
3	2	1	2	1	1	4937.572	4937.570	0.002
3	2	2	2	1	2	4963.490	4963.498	-0.008
7	0	7	6	0	6	5321.499	5321.493	0.006
7	2	6	6	2	5	5324.592	5324.586	0.008
7	2	5	6	2	4	5328.228	5328.209	0.018
4	2	2	3	1	2	5685.798	5685.810	-0.012
4	2	3	3	1	3	5737.362	5737.342	0.020
8	0	8	7	0	7	6080.417	6080.424	-0.007
8	1	7	7	1	6	6119.462	6119.493	-0.031

**Note.** *v*<sub>obs</sub> is the observed frequency and *v*<sub>obs</sub> - *v*<sub>calc</sub> is the residual

### D.9 <sup>1</sup>H-NMR spectrum

#### 1-(Phenyl)tricyclo[3.3.1.0<sup>2,8</sup>]nona-3,6-dien-9-one (phe-BA)



### *D.10 References*

Halgren, T.A., *J. Comp. Chem.* 1996, **17**, 490-519.

Becke, A. D., *Phys. Rev.* 1988, **38**, 3098–3100.

Lee, C. T., Yang, W.T., Parr, R.G., *Phys. Rev.* 1988, **37**, 785–789.



Grimme, S., Antony, J., Ehrlich, S., and Krieg, S. *J. Chem. Phys.* 2010, **132**, 154104.

Ditchfield, R., Hehre, W., Pople, J., *J. Chem. Phys.* 1971, **54**, 724-728.

Schwabe, T., Grimme, S., *Phys. Chem. Chem. Phys.* 2007, **9**, 3397-3406.



# Amino acetaldehyde conformers: structure and spectroscopic properties

Pilar Redondo ,  Miguel Sanz-Novo, Antonio Largo and Carmen Barrientos*Computational Chemistry Group, Departamento de Química Física y Química Inorgánica, Facultad de Ciencias, Universidad de Valladolid, E-47011 Valladolid, Spain*

Accepted 2019 December 13. Received 2019 December 11; in original form 2019 November 11

## ABSTRACT

We present a computational study of the different conformers of amino acetaldehyde. This molecule is a precursor of glycine and also an isomer of the detected molecules acetaldehyde and methylformamide. In addition, a previous theoretical result shows that amino acetaldehyde could be formed from the gas phase reaction of formamide with  $\text{CH}_3^+$ . Different computational approaches, going from density functional theory (DFT) to coupled cluster (CC) calculations, are employed for the characterization of the amino acetaldehyde conformers. We locate four low-lying conformation on the singlet potential energy surface (PES), two with a synperiplanar arrangement of the carboxylic oxygen atom and the  $\text{NH}_2$  group, and the other two conformers with an anticlinal disposition. All levels of theory predict the conformer with a synperiplanar arrangement and the H atoms of the  $\text{NH}_2$  group pointing in the direction of the oxygen, denoted as *in-sp*-amino acetaldehyde, as the most stable. The viability of the interconversion processes between the four conformers in space is analysed. Relevant spectroscopic parameters to rotational spectroscopy with ‘spectroscopic’ accuracy at the composite level are reported. Vibrational frequencies and infrared intensities are also computed at the CC with single and double excitations (CCSD) level including anharmonic corrections. This information could help in the experimental characterization of amino acetaldehyde that could be considered as a good candidate molecule to be searched for in space.

**Key words:** astrobiology – astrochemistry – ISM: general – ISM: molecules.

## 1 INTRODUCTION

Complex organic molecules (COMs), organic species containing carbon, nitrogen, and oxygen atoms and having six or more atoms, are considered as possible precursors of prebiotic molecules in the interstellar medium (ISM). These molecules under prebiotic conditions could play a crucial role in the synthesis of biomolecules, the building blocks of life (Quan et al. 2016). The increase on the number and complexity of the detected COMs in the last years is related to the availability of greater and greater accurate radio telescopes. Previous theoretical or experimental spectroscopic characterization of these species is crucial for their eventual identification in space.

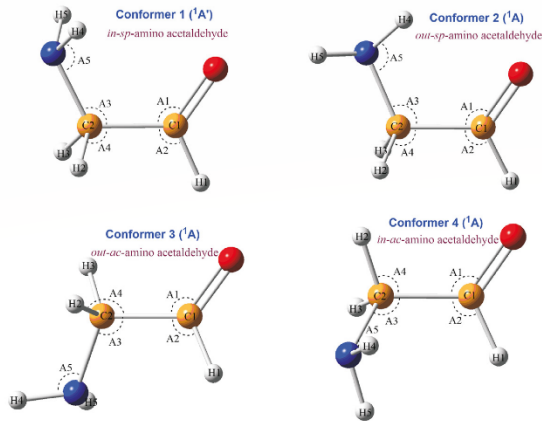
Among the COMs detected, it is found molecular species (isomers) that correspond to the same molecular formula, but contain different functional groups. In this way, examples of isomeric families characterized in space are acetic acid ( $\text{CH}_3\text{COOH}$ ; Mehringer et al. 1997), glycol aldehyde ( $\text{OHCH}_2\text{CHO}$ ; Hollis, Lovas & Jewell 2000), and methyl formate ( $\text{CH}_3\text{OCHO}$ ; Brown et al. 1975; Churchwell & Winniewisser 1975), or acetaldehyde ( $\text{CH}_3\text{CHO}$ ; Gottlieb 1973), ethylene oxide ( $\text{C-C}_2\text{H}_4\text{O}$ ; Dickens et al.

1997), and vinyl alcohol ( $\text{CH}_2\text{CHOH}$ ; Turner & Apponi 2001), with molecular composition [2C, 4H, 2O] and [2C, 4H, O], respectively. The knowledge of abundances and special distribution in the ISM of isomers could be useful to understand the formation routes, gas phase or grain process, of these molecules. In general, it is found that the most stable isomer should be the most abundant one (Lattalais et al. 2009). Nevertheless, since different chemical groups constitute the isomers, their formation routes could be different and, in these cases, the kinetic factors will be key on the abundances. From a kinetic point of view, activation barriers are determinant in the formation rate of the final product.

Acetamide and N-methylformamide, two isomers with a peptide-type bond ( $-\text{NH}-\text{CO}-$ ); have been detected in space (Hollis et al. 2006; Belloche et al. 2017). These molecules are very interesting from a prebiotic point of view, since the knowledge of their formation processes is essential to understand more complex molecules evolution. From the analysis of the relative abundances of both isomers, Belloche et al. (2017) suggest that the formation of these two molecules is controlled by kinetics rather than thermal equilibrium. They also remark that the detection of other isomers of molecular composition [2C, 5H, N, O] could clarify the origin of these systems.

On the other hand, considering the similarity found by Halfen, Ilyushin & Ziurys (2011) between the abundances and distribution

\* E-mail: predondo@qf.uva.es

1828 *P. Redondo et al.*

**Figure 1.** Structure of the four conformers located for amino acetaldehyde on the singlet potential energy surface (PES).

of formamide and acetamide towards SgrB2(N), they propose that acetamide could be synthesized through the reaction between formamide and  $\text{CH}_3^-$ . In a recent study, we performed a theoretical study of this process (Redondo, Barrientos & Largo 2014) and we found that the formation of protonated acetamide and amino acetaldehyde is an exothermic process and takes place through barrier free paths. Therefore, this reaction could be a source of acetamide and amino acetaldehyde, a precursor of glycine by oxidation, in space. Amino acetaldehyde, an isomer of acetamide and N-methylformamide, is located  $\approx 26$  kcal mol $^{-1}$  above acetamide, the lowest energy structure obeying to the chemical formula [2C, 5H, N, O] (Lattelais et al. 2010). In this theoretical study, amino acetaldehyde has been characterized at the density functional theory (DFT), in particular using the B3LYP functional. To the best of our knowledge, there have been no previous experimental studies on this species.

In order to help in the laboratory or astronomical detection of amino acetaldehyde, the main aim of this study is to characterize their conformers at the so-called ‘spectroscopic’ accuracy. To obtain this accuracy we employ a composite approach CCSD(T)/CBS + CV (Heckert, Kállay & Gauss 2005; Heckert et al. 2006) that considers the extrapolation to the complete basis set limit, CBS, and core-valence corrections, CV, at the coupled clusters level including single and double excitations and a perturbative treatment of triple excitations, CCSD(T), level (Raghavachari et al. 1989). This approach obtain bond distances with an accuracy of 0.001–0.002 Å and angles accurate to 0.05°–0.1° (Barone et al. 2013a).

## 2 COMPUTATIONAL METHODS

Different theoretical methodologies have been used within the context of DFT and ab initio approaches. Minima and transition states were characterized at the DFT level using the B3LYP (Becke 1986, 1988a,b; Lee, Yang & Parr 1988) and the B2PLYP functionals (Grimme 2006). The second-order Møller–Plesset (MP2; Møller & Plesset 1934) and coupled cluster with single and double excitations (CCSD; Raghavachari et al. 1989) ab initio theories were also used. In the CCSD calculations, the frozen-core (fc) approximation has been employed in conjunction with the Dunning’s cc-pVTZ (Dunning 1989) basis set. At the B3LYP, B2PLYP, and MP2 levels

of theory, the aug-cc-pVTZ (Dunning 1989; Kendall, Dunning & Harrison 1992) basis set was used.

For each optimized structure, harmonic vibrational frequencies were calculated to characterize the stationary points and estimate the zero-point vibrational (ZPV) energies. In addition, anharmonic corrections were calculated at the CCSD/cc-pVTZ level and B2PLYP/aug-cc-pVTZ levels using a second-order perturbation treatment (VPT2; Barone 2005).

As in our previous work (Redondo, Largo & Barrientos 2018), to obtain more accurate geometrical parameters, equilibrium rotational constants, and electronic energies for the amino acetaldehyde conformers, a composite procedure (Heckert et al. 2005, 2006) was applied. The computed predicted geometrical parameters and equilibrium rotational constants, which we will denote as  $P(\text{comp})$ , are obtained using the following expression (Barone et al. 2013b):

$$P(\text{comp}) = P(\text{CBS}) + \Delta P(\text{CV}) + \Delta P(\text{T}) + \Delta P(\text{aug}), \quad (1)$$

with

$$\Delta P(\text{CV}) = P(\text{CCSD}/\text{cc-pCVTZ, all}) - P(\text{CCSD}/\text{cc-pCVTZ, fc}), \quad (2)$$

$$\Delta P(\text{T}) = P(\text{CCSD}(\text{T})/\text{cc-pVTZ}) - P(\text{CCSD}/\text{cc-pVTZ}), \quad (3)$$

$$\Delta P(\text{aug}) = P(\text{CCSD}/\text{aug-cc-pVTZ}) - P(\text{CCSD}/\text{cc-pVTZ}), \quad (4)$$

where  $P(\text{CBS})$  considers the basis set truncation effects calculated employing the CCSD/cc-pVTZ and CCSD/cc-pVQZ data (Helgaker et al. 1997),  $\Delta P(\text{CV})$  is included to take into account core-valence electron correlation, and the terms  $\Delta P(\text{T})$  and  $\Delta P(\text{aug})$  accounts for the correlation due to the inclusion of triplet excitation at the CCSD level and inclusion of diffuse functions (aug) in the basis set, respectively. Composite energies are estimated using the same approach including the term corresponding to ZPV energy.

Both GAUSSIAN 16 (Frisch et al. 2016) and CFOUR (Stanton et al. 2013) program packages were used for all computations.

## 3 RESULTS AND DISCUSSION

In this section, we will begin by the localization of the amino acetaldehyde conformers and the analysis of their stability. Then we will discuss the interconversion processes between them and finally we will report their spectroscopic properties.

### 3.1 Amino acetaldehyde conformers

In Fig. 1, we show the structure of the four conformers characterized for amino acetaldehyde on the singlet potential energy surface (PES). Conformer 1 has  $C_s$  symmetry with  $^1A'$  electronic ground state, and the other three conformers have no symmetry and their ground state is  $^1A$ . As can be seen in Fig. 1, conformers 1 and 2, which have the oxygen atom and the  $\text{NH}_2$  group on the same side of the plane with a OC1C2N dihedral angle near  $0^\circ$ , reveal a synperiplanar (*sp*) arrangement. In conformer 1 the H atoms of the  $\text{NH}_2$  group are pointing in the direction of the oxygen (*in*), whereas in conformer 2 the hydrogen atoms are addressed outside the skeleton of the molecule (*out*) and we can denote them as *in-sp*- and *out-sp*-amino acetaldehyde, respectively. Conformers 3 and 4, which arise from conformer 1 or 2 by rotation around the C–C bond, with a OC1C2N dihedral angle between  $90^\circ$  and  $150^\circ$  or  $-90^\circ$  and  $-150^\circ$ , are called anticlinal (*ac*). Considering the orientation of hydrogen in the  $\text{NH}_2$  group, conformer 3 is denoted as

**Table 1.** Structural parameters (distances in Angstroms and angles in degrees) for the amino acetaldehyde conformers at different levels of theory.

Level	OC1	C1C2	A1	CIH1	A2	C2N	A3	H2C	A4	H4N	A5	Dih1 <sup>a</sup>	Dih2 <sup>b</sup>	Dih3 <sup>c</sup>	Dih4 <sup>d</sup>
<i>Conformer 1</i>															
B3LYP/aug-cc-pVTZ	1.2040	1.5150	124.7	1.1113	114.6	1.4474	117.4	1.0964	107.0	1.0129	110.4	180.0	0.0	-124.1	-58.5
MP2/aug-cc-pVTZ	1.2147	1.5102	124.0	1.1066	115.2	1.4474	116.6	1.0942	107.2	1.0136	109.4	180.0	0.0	-123.7	-57.3
B2PLYP/aug-cc-pVTZ	1.2086	1.5129	124.4	1.1080	114.9	1.4474	117.0	1.0944	107.1	1.0123	110.0	180.0	0.0	-123.9	-58.0
CCSD/cc-pVTZ	1.2052	1.5137	124.0	1.1059	115.9	1.4498	116.2	1.0940	107.4	1.0123	108.9	180.0	0.0	-123.6	-56.8
Composite	1.2050	1.5104	124.1	1.1061	113.3	1.4430	116.5	1.0940	107.1	1.0111	110.1	180.0	0.0	-123.7	-58.1
<i>Conformer 2</i>															
B3LYP/aug-cc-pVTZ	1.2051	1.5037	124.2	1.1073	114.8	1.4470	112.7	1.1017	105.3	1.0131	110.2	179.4	0.0	-116.9	-20.7
MP2/aug-cc-pVTZ	1.2151	1.5018	123.7	1.1035	115.4	1.4478	111.6	1.0993	105.3	1.0140	109.2	179.0	0.0	-115.2	-24.5
B2PLYP/aug-cc-pVTZ	1.2094	1.5029	124.0	1.1045	115.0	1.4476	112.2	1.0997	105.3	1.0125	109.8	179.2	9.3	-116.0	-22.4
CCSD/cc-pVTZ	1.2053	1.5066	123.9	1.1031	115.2	1.4506	111.5	1.0994	105.5	1.0115	108.6	180.0	10.8	-114.6	-28.1
Composite	1.2057	1.5010	123.7	1.1030	115.4	1.4445	112.0	1.0983	105.7	1.0118	109.9	180.0	8.2	-116.8	-19.0
<i>Conformer 3</i>															
B3LYP/aug-cc-pVTZ	1.2037	1.5159	124.2	1.1086	114.2	1.4583	110.9	1.0951	105.7	1.0100	111.8	179.4	149.8	-93.0	159.3
MP2/aug-cc-pVTZ	1.2145	1.5099	124.1	1.1045	114.3	1.4586	109.7	1.0934	106.1	1.0108	111.0	179.7	145.2	-98.2	162.5
B2PLYP/aug-cc-pVTZ	1.2084	1.5132	124.1	1.1055	114.3	1.4585	110.4	1.0934	105.9	1.0095	111.5	179.5	147.8	-95.3	160.5
CCSD/cc-pVTZ	1.2048	1.5135	124.1	1.1039	114.2	1.4597	109.9	1.0933	106.3	1.0095	110.3	180.0	146.2	-97.0	164.4
Composite	1.2051	1.5101	123.8	1.1042	114.7	1.4547	109.9	1.0932	106.2	1.0082	111.6	180.0	144.6	-98.4	160.3
<i>Conformer 4</i>															
B3LYP/aug-cc-pVTZ	1.2036	1.5175	125.0	1.1134	114.4	1.4589	114.3	1.0922	108.3	1.0121	112.0	178.4	-134.8	122.0	60.8
MP2/aug-cc-pVTZ	1.2142	1.5109	124.8	1.1089	114.6	1.4586	113.8	1.0906	108.3	1.0126	111.1	178.5	-134.1	121.7	59.7
B2PLYP/aug-cc-pVTZ	1.2082	1.5147	124.9	1.1103	114.5	1.4588	114.0	1.0905	108.3	1.0115	111.6	178.4	-134.3	121.9	60.4
CCSD/cc-pVTZ	1.2042	1.5162	124.6	1.1086	114.6	1.4579	114.7	1.0908	108.0	1.0112	110.6	180.0	-142.1	123.3	57.2
Composite	1.2047	1.5122	124.4	1.1088	115.1	1.4532	114.3	1.0905	108.1	1.0100	111.9	180.0	-138.1	123.0	58.3

<sup>a</sup>Dih1 = dih(C2C1OH1).<sup>b</sup>Dih2 = dih(NC2C3O).<sup>c</sup>Dih3 = dih(H2C2C1O).<sup>d</sup>Dih4 = dih(H4NC2C1).**Table 2.** Relative energies, in kcal mol<sup>-1</sup>, for the four conformers of amino acetaldehyde and transition states connecting them computed at different levels of theory (ZPV energies included).

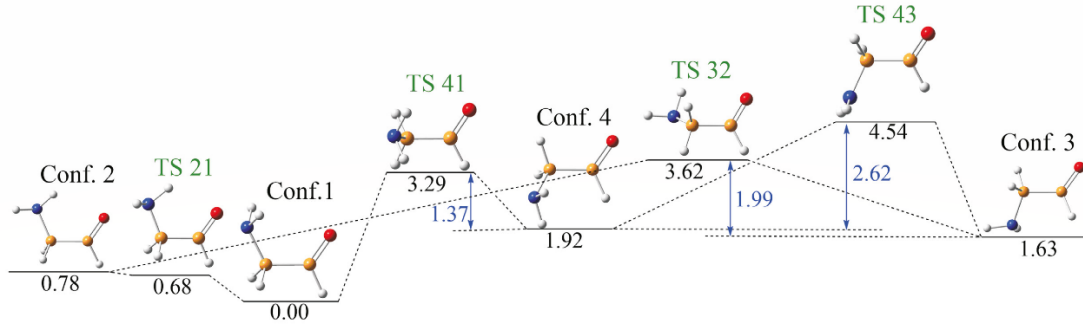
Level	Conf. 1	Conf. 2	Conf. 3	Conf. 4	TS21	TS41	TS32	TS43
B3LYP/aug-cc-pVTZ	0.00	0.56	1.53	1.78	0.48	3.36	4.04	3.57
MP2/aug-cc-pVTZ	0.00	0.68	1.67	2.01	0.62	3.39	3.68	4.43
B2PLYP/aug-cc-pVTZ	0.00	0.63	1.57	1.86	0.56	3.34	3.86	3.93
CCSD/cc-pVTZ	0.00	1.11	1.66	2.00	1.04	3.50	3.76	5.25
CCSD(T)/cc-pVTZ	0.00	1.07	1.78	1.98				
CCSD/cc-pVQZ	0.00	0.91	1.56	1.95				
CCSD/aug-cc-pVTZ	0.00	0.86	1.49	1.90				
CCSD(T)/aug-cc-pVQZ//								
CCSD/cc-pVTZ	0.00	0.78	1.63	1.92	0.68	3.29	3.62	4.54
Composite	0.00	0.44	1.44	1.81				

*out-ac*-amino acetaldehyde, and conformer 4 denoted as *in-ac*-amino acetaldehyde.

The geometrical parameters of the low-lying conformers of amino acetaldehyde computed at different levels of theory are shown in Table 1, whereas in Table 2 their relative energies calculated at the same levels of theory are included. Since the structure of the four conformers only differs in the relative orientation of the functional groups, the bond lengths are very close in the four minima. The most significant change is found for the C–N bond length that is about 0.01 Å shorter in synperiplanar conformers (1 and 2) than in anticlinal ones (3 and 4). If we compare the different levels of calculations employed with the composite values, we can see in Table 1 that the differences found in the bond lengths are in the hundredths of angstroms, and for the angles are less than 1°. The greatest discrepancies correspond to the dihedral angles, specifically in the case of the Dih2 in conformer 2.

Regarding the relative energies of amino acetaldehyde conformers shown in Table 2, all levels of theory used in this work provide the same stability order of conformers. The deviation obtained for the relative energies calculated at different levels of theory is for the four conformers less than 0.7 kcal mol<sup>-1</sup>. The values obtained at the CCSD(T)/aug-cc-pVQZ level on the CCSD/cc-pVTZ geometries (denoted as CCSD(T)/aug-cc-pVQZ//CCSD/cc-pVTZ) are much closer to those estimated with the composite approach than the CCSD/cc-pVTZ ones. It should be pointed out that relative values obtained at the B3LYP and especially at the B2PLYP levels for this system are very close to the composite ones. Therefore, these levels could be employed for predicting relative energies in larger systems.

All levels of theory agree in the prediction of conformer 1, *in-sp*-amino acetaldehyde, as the most stable one. The other synperiplanar conformer, *out-sp*-amino acetaldehyde, the following in energy, is



**Figure 2.** Interconversion processes between the four conformers located for amino acetaldehyde. Relative energies (in black) respect to conformer 1, and interconversion barriers (in blue), calculated at the CCSD(T)/aug-cc-pVQZ//CCSD/cc-pVTZ level. Energies in kcal mol<sup>-1</sup> (ZPV energies are included).

located only 0.44 kcal mol<sup>-1</sup> above conformer 1 at the composite level. Anticlinical conformation is less stable than synperiplanar one, so conformers 3 and 4 are the most energetics. The energy difference between the two anticlinical conformers is only 0.37 kcal mol<sup>-1</sup> at the composite level, similar to that found for synperiplanar ones. The stability of synperiplanar conformers (1 and 2) over anticlinical ones (3 and 4) is related to the interaction between the hydrogen atoms of the group NH<sub>2</sub> and the oxygen one that is greater in conformer 1. However, for anticlinical structures steric factors make conformer 3 slightly more stable than conformer 4, due to the distinct orientation of the hydrogen atoms of the NH<sub>2</sub> group with respect to the heavy atoms skeleton.

### 3.2 Isomerization processes

Once we have characterized the different conformers of amino acetaldehyde we will now analyse the interconversion processes between them and discuss the results in the context of the ISM. For that, we have located the transition states associated with the interconversion processes among the four conformers. The relative energies of transition states for the different isomerization processes between conformers are included also in Table 2, and in Fig. 2, we represent the different isomerization processes including the isomerization energies. These energies are calculated at the CCSD(T)/aug-cc-pVQZ on the optimized CCSD/cc-pVTZ geometries. This level is shown in previous section that provides relative energies close to that calculated using the composite approach. We denote transition states with TS followed by two numbers that correspond to the conformers they connect. Transition states structures are supported by vibrational analysis, each one presents one imaginary frequency associated with the conversion process. With regard to the geometrical parameters on the optimized transition state structures, as can be observed in Fig. 2, all of them are intermediate between those of the conformers that connect the transition state structures.

We can see in Fig. 2 that the isomerization of the conformer 2, *out-sp*-amino acetaldehyde, to the most stable conformer 1, *in-sp*-amino acetaldehyde, takes place through transition state TS21, where a rotation of the NH<sub>2</sub> group occurs, and does not imply a net activation barrier. Transition state TS21 is located below the conformer 2 (0.10 kcal mol<sup>-1</sup> at the CCSD(T)/aug-cc-pVQZ level). The same result is obtained for all the calculation levels used, as can be seen in Table 2. Therefore, conformer 2 evolves directly to the conformer 1 and its experimental observation would not be possible. The isomerization between the two anticlinical

conformers (3 and 4) takes place through transition state TS43 that is located 2.62 kcal mol<sup>-1</sup> ( $\approx 1320$  K) above the conformer 4, *in-ac*-amino acetaldehyde. Isomer 4 can also isomerize to the more stable isomer 1 through the transition state TS41. In this process a rotation around the C–C bond occurs. The interconversion barrier associated with the conformer 4  $\rightarrow$  conformer 1 process is about 1.37 kcal mol<sup>-1</sup> ( $\approx 690$  K) above conformer 4. Therefore, the most favourable isomerization process for conformer 4, *in-ac*-amino acetaldehyde, from thermodynamic and kinetic points of view, is the interconversion into the most stable conformer 1, *in-sp*-amino acetaldehyde. Finally, conformer 3, *out-ac*-amino acetaldehyde, can isomerize to conformer 2, *out-sp*-amino acetaldehyde. This process presents an activation barrier associated with transition state TS32 about 1.99 kcal mol<sup>-1</sup> ( $\approx 1000$  K). Once conformer 2 is reached, it evolves to the most stable conformer, *in-sp*-amino acetaldehyde. Thus, the isomerization of the conformer 3 to 1 is not a direct process since it takes place through the conformer 2.

The obtained energy barriers for the interconversion processes among the four conformers show that, at the low temperatures of the ISM, conformer 2 will isomerize to conformer 1. The isomerization processes between the conformers 1, 3, and 4 involve net energy barriers. The easiest process corresponds to the interconversion of conformer 4 (the less stable) to conformer 1 (the most stable) with a barrier of 1.37 kcal mol<sup>-1</sup> ( $\approx 690$  K). Therefore, it is worth noticing that the three conformers, 1, 3, and 4, could coexist in the ISM.

### 3.3 Spectroscopic parameters

Besides the conformational landscape of amino acetaldehyde, we report high-level calculated spectroscopic parameters relevant to the investigation of the experimental infrared (IR) and rotational spectrum, which are collected in Tables 3 and 4. The composite approach was used to compute the equilibrium rotational constants ( $A_e$ ,  $B_e$ ,  $C_e$ ), while the corresponding constants for the ground vibrational state ( $A_0$ ,  $B_0$ ,  $C_0$ ) were obtained from the vibration–rotation coupling constants and degeneracy factors for the vibrational modes. The similarities found in structural parameters, given in Table 1, between the two conformations with a synperiplanar disposition (conformers 1 and 2) or with anticlinical orientation (conformers 3 and 4) remain in the calculated rotational constants (Table 3). We have also presented in Table 3 centrifugal distortion parameters in the symmetrically reduced Hamiltonian ( $\Delta_J$ ,  $\Delta_K$ ,  $\Delta_{JK}$ ,  $\delta_J$ , and  $\delta_K$ ). In this case, another interesting spectroscopic tool is taken into account: the presence of a <sup>14</sup>N nucleus. Therefore,

**Table 3.** Spectroscopic parameters for the amino acetaldehyde conformers (in MHz). Equilibrium rotational constants ( $A_e$ ,  $B_e$ ,  $C_e$ ) and rotational constants for the ground vibrational state ( $A_0$ ,  $B_0$ ,  $C_0$ ) are computed with the composite method. Centrifugal distortion parameters in the symmetrically reduced Hamiltonian ( $\Delta_J$ ,  $\Delta_K$ ,  $\Delta_{JK}$ ,  $\delta_J$ , and  $\delta_K$ ), the diagonal elements of the  $^{14}\text{N}$  nuclear quadrupole coupling tensor ( $\chi_{aa}$ ,  $\chi_{bb}$ ,  $\chi_{cc}$ ), and dipole moment components ( $\mu$  in Debye) are computed at the CCSD/cc-pVTZ level.

	Confor. 1	Confor. 2	Confor. 3	Confor. 4
$A_e$	17849.4	17962.9	32345.9	30429.1
$B_e$	6029.6	6290.3	4348.3	4349.4
$C_e$	4747.3	4832.0	4131.7	4158.1
$A_0$	17762.9	17869.9	33197.7	34367.9
$B_0$	5973.8	6199.7	4294.6	4185.2
$C_0$	4704.5	4772.2	4057.6	3929.7
$\Delta_J \times 10^3$	5.64515	5.07310	2.94212	12.5669
$\Delta_{JK} \times 10^3$	-15.8292	-18.0168	-100.718	-476.493
$\Delta_K \times 10^3$	47.9392	51.5130	1993.72	6520.91
$\delta_J \times 10^3$	-1.62698	-1.54878	0.33183	1.85029
$\delta_K \times 10^6$	-28.4672	-107.309	-42.0461	-49.5764
$\mu_a$	0.26	1.23	2.83	1.67
$\mu_b$	1.64	2.18	0.63	1.19
$\mu_c$	0.0	1.24	0.11	0.8
$\chi_{aa}$	-4.3467	-0.5322	1.2532	0.8991
$\chi_{bb}-\chi_{cc}$	0.5809	-4.5809	-2.2590	4.5104

we give the values of the quadrupole coupling constants for the  $^{14}\text{N}$  nucleus ( $\chi_{aa}$ ,  $\chi_{bb}$ , and  $\chi_{cc}$ ). On the one hand, the rotational constants provide information of the mass distribution. In addition, the nuclear quadrupole coupling interactions show a strong dependence on the electronic environment close of the  $^{14}\text{N}$  nucleus. It introduces hyperfine rotational signatures at located sites of amino acetaldehyde and acts as a conclusive probe of the position, orientation, and chemical environment of the  $^{14}\text{N}$  quadrupolar nucleus. Hence, to identify the plausible observed species, the resolution of the nuclear quadrupole hyperfine structure of amino acetaldehyde is needed. Moreover, the

values of the computed dipole moment components (Table 3) are high enough to allow for the observation of these conformers by rotational spectroscopy.

To guide spectral searches, we have simulated the rotational spectrum from the corresponding rotational constants, dipole-moment components, and  $^{14}\text{N}$  nuclear quadrupole coupling tensor constants. In Fig. 3, we present a prediction of the rotational spectra of *in-sp*-amino acetaldehyde (conformer 1) in the 0–500 GHz frequency region using a semirigid-rotor approximation including quartic centrifugal distortion parameters in the symmetrically reduced Hamiltonian. SPCAT program (Pickett 1991) is employed to predict frequencies and intensities from the computed parameters.

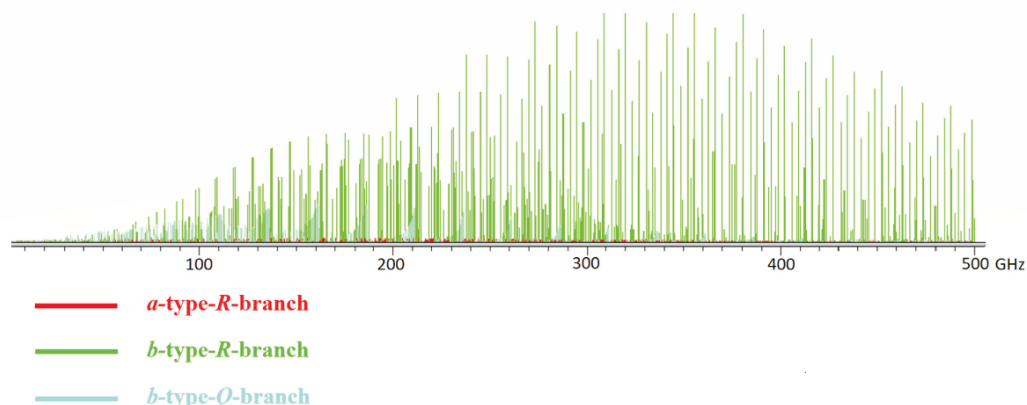
In Table 4, we give CCSD harmonic and anharmonic vibrational frequencies and the corresponding IR intensities for the four amino acetaldehyde conformers. These results could guide their eventual characterization in the gas phase through IR spectroscopy. We were only able to evaluate anharmonic corrections at the CCSD/cc-pVTZ level have only been possible to be evaluated for *in-sp*-amino acetaldehyde (conformer 1) that has  $C_s$  symmetry. For this isomer, we have also computed the anharmonic corrections at the B2PLYP and MP2/aug-cc-pVTZ levels in order to check the accuracy of both approaches. Our results show that anharmonic corrections at the B2PLYP/aug-cc-pVTZ are closer to those calculated at the CCSD level than when the corrections are calculated at the MP2 level. The mean absolute errors with respect to CCSD frequencies are 40 and 14  $\text{cm}^{-1}$  for MP2 and B2PLYP anharmonic corrections, respectively. Therefore, for the rest of conformers, with  $C_1$  symmetry, anharmonic corrections are estimated at the B2PLYP/aug-cc-pVTZ level.

Anharmonic frequencies, in the most of the cases, are lower than the corresponding harmonic ones, according with the usual behaviour. The largest anharmonic corrections are found, in general, in the larger frequencies corresponding to stretching modes. As we can see in Table 4, there is not an evident relation between anharmonic and harmonic vibrational frequencies. Hence, it is not

**Table 4.** Harmonic,  $\omega$ , and anharmonic,  $\nu$ , vibrational frequencies ( $\text{cm}^{-1}$ ), and IR intensities ( $\text{km mol}^{-1}$ ) for the amino acetaldehyde conformers calculated at the CCSD/cc-pVTZ level.

Mode	Conformer 1				Conformer 2 <sup>a</sup>				Conformer 3 <sup>a</sup>				Conformer 4 <sup>a</sup>			
	$\omega$	$I_{\text{har}}$	$\nu$	$I_{\text{anhar}}$	$\omega$	$I_{\text{har}}$	$\nu$	$I_{\text{anhar}}$	$\omega$	$I_{\text{har}}$	$\nu$	$I_{\text{anhar}}$	$\omega$	$I_{\text{har}}$	$\nu$	$I_{\text{anhar}}$
a', a	125	47.2	144	79.9	143	39.9	29	13.7	84	10.2	85	9.2	58	9.6	28	6.6
a', a	248	13.6	256	2.8	219	13.1	206	20.3	252	44.6	227	37.5	278	51.3	234	35.2
a', a	268	21.7	269	19.5	295	7.1	461	9287.2	352	14.1	348	10.3	346	9.3	341	21.8
a', a	708	0.5	714	0.4	701	24.3	677	23.1	537	5.3	536	4.5	516	9.3	527	8.3
a', a	740	2.7	725	2.2	763	5.2	754	59.8	762	8.9	730	14.6	750	0.5	652	1.2
a', a	868	64.5	845	86.3	857	105.0	792	59.7	873	136.6	807	91.6	864	192.6	769	157.0
a', a	970	126.3	928	50.5	928	47.5	889	48.1	1009	4.8	995	4.1	996	0.8	985	1.7
a', a	1019	0.3	1026	0.0	1103	4.9	1069	2.7	1092	18.5	1061	22.0	1085	2.7	1045	10.5
a', a	1174	13.5	1143	8.8	1148	3.0	1126	4.8	1131	4.8	1095	6.5	1137	5.1	1121	12.9
a', a	1222	0.1	1202	0.5	1245	9.0	1205	7.1	1210	0.9	1186	0.3	1215	3.5	1189	0.0
a', a	1393	4.5	1356	4.3	1275	0.9	1251	1.9	1327	13.1	1289	14.1	1357	8.5	1321	13.1
a', a	1407	0.6	1370	0.9	1434	10.2	1397	14.0	1398	0.7	1366	1.4	1396	0.7	1365	0.1
a', a	1429	11.3	1405	4.2	1449	4.8	1410	3.3	1438	8.3	1403	3.7	1436	3.3	1397	3.2
a', a	1476	10.6	1438	7.4	1510	10.8	1474	0.0	1500	5.5	1456	4.6	1485	4.3	1433	2.4
a', a	1695	19.1	1605	2.2	1666	48.2	1649	17.5	1689	22.4	1656	18.3	1684	21.9	1649	17.4
a', a	1841	121.6	1869	–	1840	122.9	1816	80.8	1840	134.6	1806	117.2	1839	132.4	1808	100.6
a', a	2947	107.8	2822	60.9	2974	60.2	2845	22.0	2976	72.3	2826	72.3	2919	104.9	2710	51.6
a', a	3052	31.0	2931	24.8	2999	68.8	2844	35.6	3048	27.0	2943	21.6	3076	15.6	2939	10.6
a', a	3086	13.0	2935	14.0	3059	29.7	2885	32.5	3093	18.7	2950	15.8	3125	10.8	2974	15.8
a', a	3532	1.5	3383	0.5	3549	4.9	3388	5.6	3547	0.7	3400	0.3	3535	0.8	3392	0.0
a', a	3603	3.0	3431	1.4	3642	9.9	3467	4.7	3630	3.8	3464	0.5	3620	2.1	3452	0.0

<sup>a</sup>Anharmonic contributions are computed at the B2PLYP/aug-cc-pVTZ level.



**Figure 3.** Calculated rotational spectrum at 50 K of *in-sp*-amino acetaldehyde (conformer 1) in the 0–500 GHz frequency region using a semirigid-rotor approximation including quartic centrifugal distortion parameters in the symmetrically reduced Hamiltonian.

possible to find a general scale factor to bring the harmonic frequencies to the anharmonic ones. The same situation is found when we compare anharmonic intensities with the harmonic ones. Thus, we can infer that in order to predict reliable IR spectra for these conformers it is mandatory to include anharmonic contributions.

The band associated with the C–O stretching mode is one of the most intense in the IR spectrum of all conformers. This band is located between  $1806\text{ cm}^{-1}$  in conformer 3 and  $1869\text{ cm}^{-1}$  in conformer 1. The two anticlinal conformers, 3 and 4, present an IR spectrum similar. It is related with the very close geometries, which only differ in the orientation of the hydrogen atoms of the  $\text{NH}_2$  group.

#### 4 CONCLUSIONS

In this work, we have performed a study of amino acetaldehyde conformers using different quantum chemical approaches. We have characterized four low-lying conformers on the singlet PES. Conformers 1 and 2 have a synperiplanar (*sp*) arrangement with the carboxylic oxygen atom and the  $\text{NH}_2$  group on the same side of the plane. In conformer 1, denoted as *in-sp*-amino acetaldehyde, the H atoms of the  $\text{NH}_2$  group point in the direction of the oxygen, whereas in conformer 2 the hydrogen atoms are addressed outside the skeleton of the molecule and is called *out-sp*-amino acetaldehyde. Conformers 3 and 4 present an anticlinal (*ac*) disposition of the carboxylic oxygen atom and the  $\text{NH}_2$  group. Considering the orientation of hydrogen in the  $\text{NH}_2$  group, conformer 3 is denoted as *out-ac*-amino acetaldehyde, and conformer 4 as *in-ac*-amino acetaldehyde.

The same stability order of conformers is obtained at all levels of theory used in this work. The variations of the relative energies computed at the different levels are less than  $0.7\text{ kcal mol}^{-1}$ . Conformer 1, *in-sp*-amino acetaldehyde, is the most stable followed by the other synperiplanar conformer, *out-sp*-amino acetaldehyde, which is located only  $0.44\text{ kcal mol}^{-1}$  above conformer 1 at the composite level. Synperiplanar conformers (1 and 2) are more stable than anticlinal ones (3 and 4) due to the interaction between the hydrogen atom of the group  $\text{NH}_2$  and the oxygen one. Conformers 3 and 4 are located  $1.44$  and  $1.81\text{ kcal mol}^{-1}$  above the conformer 1 at the composite level, respectively.

Isomerization processes between the four conformers are also studied. The obtained energy barriers for the interconversion processes show that, at the low temperatures of the ISM, the isomerization of conformer 2 to conformer 1 (the most stable) is a barrier-free process, therefore, conformer 2 will isomerize to 1 and its experimental observation would not be possible. The isomerization processes between the conformers 1, 3, and 4 involve a net energy barrier. The easiest process corresponds to the interconversion of conformer 4 (the less stable) to conformer 1 (the most stable) with a barrier of  $1.37\text{ kcal mol}^{-1}$  ( $\approx 690\text{ K}$ ) at the CCSD(T)/aug-cc-pVQZ//CCSD/cc-pVTZ level. However, under the conditions of the ISM the three conformers, 1, 3, and 4, could coexist.

For the four conformers we have obtained the relevant spectroscopic parameters to rotational spectroscopy with ‘spectroscopic’ accuracy at the composite level. We provided accurate theoretical values of the  $^{14}\text{N}$  nuclear quadrupole coupling constants together with the rotational constants in order to reproduce the spectrum to enable its distinct possible identification. For observations in low-frequency regions like the ones accessible for the Green Bank Telescope (GBT), the hyperfine components of the transitions, which normally are propagated several MHz, could transpire to be a handicap for interpreting the radioastronomical observations. Vibrational frequencies and IR intensities are also reported at the CCSD level including anharmonic corrections.

These results could help in the experimental characterization of amino acetaldehyde in the laboratory or in space by radio or IR spectroscopy.

#### ACKNOWLEDGEMENTS

Financial support from the Spanish Ministerio de Economía Industria y Competitividad (Grant AYA2017-87515-P) and the Junta de Castilla y León (Grant VA010G18) is gratefully acknowledged. MS-N acknowledges funding from a Predoctoral FPU Grant (FPU17/02987).

#### REFERENCES

- Barone V., 2005, *J. Chem. Phys.*, **122**, 014108  
 Barone V., Biczysko M., Bloino J., Puzzarini C., 2013a, *Phys. Chem. Chem. Phys.*, **15**, 1358



- Barone V., Biczysko M., Bloino J., Puzzarini C., 2013b, *J. Chem. Theory Comput.*, 9, 1533
- Becke A. D., 1986, *J. Chem. Phys.*, 84, 4524
- Becke A. D., 1988a, *J. Chem. Phys.*, 88, 1053
- Becke A. D., 1988b, *J. Chem. Phys.*, 88, 2547
- Belloche A. et al., 2017, *A&A*, 601, A49
- Brown R. D., Crofts J. G., Gardner F. F., Godfrey P. D., Robinson B. J., Whiteoak J. B., 1975, *ApJ*, 197, L29
- Churchwell E., Winnewisser G., 1975, *A&A*, 45, 229
- Dickens J. E., Irvine W. M., Ohishi M., Ikeda M., Ishikawa S., Nummelin A., Hjalmarson Å., 1997, *ApJ*, 489, 753
- Dunning T. H., 1989, *J. Chem. Phys.*, 90, 1007
- Frisch M. J. et al., 2016, *Gaussian 16*. Gaussian, Inc., Wallingford, CT
- Gottlieb C. A., 1973, in Gordon M. A., Snyder L. E., eds, *Molecules in the Galactic Environment*. Wiley, New York, p. 181
- Grimme S., 2006, *J. Chem. Phys.*, 124, 034108
- Hallén D. T., Ilyushin V. V., Ziurys L. M., 2011, *ApJ*, 743, 60
- Heckert M., Kállay M., Gauss J., 2005, *Mol. Phys.*, 103, 2109
- Heckert M., Kállay M., Tew D. P., Klopper W., Gauss J., 2006, *J. Chem. Phys.*, 125, 044108
- Helgaker T., Klopper W., Koch H., Noga J., 1997, *J. Chem. Phys.*, 106, 9639
- Hollis J. M., Lovas F. J., Jewell P. R., 2000, *ApJ*, 540, L107
- Hollis J. M., Lovas F. J., Remijan A. J., Jewell P. R., Ilyushin V. V., Kleiner I., 2006, *ApJ*, 643, L25
- Kendall R. A., Dunning T. H., Harrison R. J., 1992, *J. Chem. Phys.*, 96, 6796
- Lattalais M., Pauzat F., Ellinger Y., Ceccarelli C., 2009, *ApJ*, 696, L133
- Lattalais M., Pauzat F., Ellinger Y., Ceccarelli C., 2010, *A&A*, 519, A30
- Lee C., Yang W., Parr R. G., 1988, *Phys. Rev. B (Condensed Matter)*, 37, 785
- Mehring D. M., Snyder L. E., Miao Y., Lovas F., 1997, *ApJ*, 480, L71
- Møller C., Plesset M., 1934, *Phys. Rev.*, 46, 618
- Pickett H. M., 1991, *J. Mol. Spectrosc.*, 148, 371
- Quan D., Herbst E., Corby J. F., Durr A., Hassel G., 2016, *ApJ*, 824, 129
- Raghavachari K., Trucks G. W., Pople J. A., Head-Gordon M., 1989, *Chem. Phys. Lett.*, 157, 479
- Redondo P., Barrientos C., Largo A., 2014, *ApJ*, 793, 32
- Redondo P., Largo A., Barrientos C., 2018, *MNRAS*, 478, 3042
- Stanton J. F., Gauss J., Harding M. E., Szalay P. G., 2013, CFOUR, a quantum chemical program package
- Turner B. E., Apponi A. J., 2001, *ApJ*, 561, L207

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.