SCUOLA
NORMALE
SUPERIORE

Chasing Astronomical Complex Organic Molecules

The Role of (Vibro-)Rotational Spectroscopy

Candidate
Alessio Melli

Supervisor
Prof. Dr. Cristina Puzzarini

Internal Advisor
Prof. Vincenzo Barone

Co-Supervisor
Dr. Mattia Melosso

A thesis presented in fulfillment of the requirements for the degree of Doctor of Philosophy in

Astrochemistry

XXXV cycle

"I never met a man who knew
so much about nothing."

Seinfeld, s04e12

Contents

Materials and Methods

Results and Discussion

Appendices

Abstract

Astrochemistry is a multidisciplinary field that embraces complex topics on different space-time scales, delving into the study of molecular species in the domain that is typically associated with astrophysics, *i.e.*, the observable cosmos. As a matter of fact, the determination of the chemical composition of the universe is crucial for understanding its evolution: molecules are excellent diagnostic probes for the derivation of physical characteristics of astronomical objects. The existence of molecules in the interstellar medium has been established more than 80 years ago, and, since then, a huge number of molecular species has been detected. Indeed, the astronomical census now counts about 300 molecules, and *ca.* 90% of them have been identified using radio astronomy techniques, which collect the cosmic emission of radiation in the microwave and radio domains. In particular, the recording of broadband large-scale surveys captures the rotational transitions of all the polar species present in the astronomical object under study. These features are exceptionally molecule-specific; moreover, thanks to the low energies involved, excited rotational states are populated even at extremely low temperatures. For these reasons, radio astronomical observations allow univocal identifications even in cold quiescent interstellar regions. On the other hand, the requirement for accuracy on the frequency of rotational transitions is very demanding, and it is achievable only resorting to experimental studies. Given these premises, this thesis is directed to the collection of necessary experimental spectroscopic data on eight different proved or potential interstellar species (namely allenylacetylene, cyanovinylacetylene, phenylmethanimine, 2- and 3-furonitrile, 3-aminoisoxazole, and iso- and n-propylamine), in order to guide their detection in the interstellar medium in the millimeter- (30–300 GHz) and submillimeter-wave (0.3–3 THz) frequency ranges. To support and complement the experiment, accurate computations exploiting composite schemes rooted in coupled-cluster theory and density functional theory have been carried out for all species, thus providing precise structural information as

well as reliable spectroscopic parameters and, in some cases, the energetic description of the system. For all the investigated species, the outcomes of both theoretical and experimental studies provide the fundamental and robust basis required for guiding the astronomical searches in the interstellar medium.

List of Abbreviations

AO	Atomic orbital
ALMA	Atacama large millimeter/submillimeter array
B3	B3LYP-D3(BJ)/jun-cc-pVDZ level of theory
B3S	B3LYP-D3(BJ)/SNSD level of theory
BO	Born–Oppenheimer
B	Becke exchange functional
BJ	Becke–Johnson damping function
CBS	Complete basis set
CC	Coupled cluster theory
CDMS	Cologne database for molecular spectroscopy
ChS	"Cheap" composite scheme
CMB	Cosmic microwave background radiation
CV	Core–valence correlation
COBRA	Coaxially oriented beam resonator arrangement
COM	Complex organic molecule
D3	Grimme's D3 scheme
DFT	Density functional theory
DSD	Dispersion corrected spin-component scaled double-hybrid functional
FID	Free-induced decay
FM-mmW	Frequency modulated millimeter-wave spectrometer
FTMW	Fourier transform microwave spectrometer
FVP	Flash vacuum pyrolysis
GGA	Generalized gradient approximation
(c)GTO	(contracted) Gaussian-type orbital
HBA	hydrobenzamide

(continues)

HF	Harthee–Fock
HWHM	Half width at half maximum
IRAM	Institut de radioastronomie millimétrique
IRC	Intrinsic reaction coordinate
ISM	Interstellar medium
jB2	B2PLYP-D3(BJ)/jun-cc-pVTZ level of theory
LDA	Local density approximation
LYP	Lee–Yang–Parr correlation functional
LOS	Line of sight
LR	Linear regression approach
LSCD	Lower state combination of differences
LSDA	Local spin-density approximation
LSF	Least-squares fit
LTE	Local thermodynamic equilibrium
(G)MC	(giant) Molecular cloud
GUI	Graphical user interface
mB2	B2PLYP-D3(BJ)/maug-cc-pVTZ level of theory
mB2-dH	B2PLYP-D3(BJ)/maug-cc-pVTZ-dH level of theory
MO	Molecular orbital
MP(*n*)	Møller-Plesset theory (to the *n*-th order)
ngVLA	Next-generation very large array
NMR	Nuclear magnetic resonance
P86	Perdew local correlation functional
PAH	Polycyclic aromatic hydrocarbon
PBE	Perdew–Burke–Ernzerhof functional
PES	Potential energy surface
PLL	Phase-lock loop
QC	Quantum-chemical
revDSD	revDSD-PBEP86-D3(BJ)/jun-cc-pVTZ level of theory
RJ	Rayleight–Jeans
RSPT	Rayleight–Schrödinger perturbation theory
SCF	Self consistent field
SFR	Star-forming region
SD	Slater determinant
SE	Semi-experimental

(continues)

SKAO	Square kilometer array observatory
S/N	Signal-to-noise ratio
TM	Template molecule approach
TMC	Taurus molecular cloud
USCD	Upper state combination of differences
VPT2	Vibrational perturbation theory to the second order
ZPE	Zero-point energy

List of Tables

List of Figures

Note of the Author

I am deeply grateful to my supervisor Cristina Puzzarini for her exceptional mentorship and support throughout all phases of my Ph.D. project, and to Mattia Melosso for his rational and solid guidance. I also sincerely thank Vincenzo Barone, whose expertise helped me advance my research in Pisa. A heartfelt thanks goes also to the past and current members of my research group in Bologna. Lastly, I extend a special thanks to Jens-Uwe Grabow for supervising my research period in Hannover.

Chapter 1

Introduction

In 1987, Astrochemistry was defined as *"the chemistry of astronomical entities in which molecular compounds exist"* by its pioneer, the astronomy professor Alexander Dalgarno.[1] After more than thirty years, this statement still implies the magnitude and the radius of expansion of this field. Indeed, Astrochemistry embraces complex topics on different scales, from the study of planetary atmospheres, to the formation of galaxies, stellar evolution, and even to the origin of life.[1] The reason behind the extension of this field is embodied in the subject of study itself, *i.e.,* molecules. The determination of the chemical composition of the cosmos is crucial for the understanding its evolution: molecules are excellent diagnostic probes for the derivation of physical characteristics of intricate astronomical objects.[2]

The existence of molecules in the interstellar medium (ISM, which refers to the portions of the universe between star systems in a galaxy) has been established more than 80 years ago, with the detection of the electronic transitions of CN and CH using the Mount Wilson 100-inch telescope by McKellar in 1940.[3] One year after, Douglas and Herzberg identified CH^+.[4] However, these discoveries did not obtain the proper attention from astronomers, and only a few other diatomic species were detected in the following 20 years.[2] During this period, in 1955, Townes presented the possibility of detecting atoms and molecules by means of radio observations, but his proposition did not gain the expected follow-up.[2,5] The situation changed only in 1968, when Cheung *et al.* detected the inversion transition of ammonia in emission toward the center of the Milky Way;[2,6] one year later, formaldehyde was identified.[7] Since then, a huge number of molecular species has been detected, and the census of molecules discovered in the ISM and circumstellar shells now counts about 300 species.[8] These molecules include 19 different elements, and range in size from two to

70 atoms; the discovery of such chemical complexity led to the definition of astronomical complex organic molecules (COMs), which are the carbon-bearing species containing more than five atoms.[9,10] A complete list of the interstellar species can be found in the Cologne database for molecular spectroscopy* (CDMS) website.[11,12]

Astrochemistry and radio astronomy

Among the different astronomical techniques, radio astronomy—which works in the radio and microwave range of the electromagnetic spectrum—is the main contributor to molecular detections. In fact, more than the 90% of the interstellar species have been identified thanks to their centimeter-, millimeter-, and submillimeter-wave features.†[8] The reason is twofold:

i. in this range, the primary molecular signals arise from the rotational motion of molecules.[13] Without entering in detail, the molecular rotational energy levels are quantized, and they depend on a set of parameters known as rotational constants.[13] In turn, the latter depend sensibly on the molecular composition and geometry: rotational spectroscopy is able to discern with an outstanding precision isotopologues and conformers of molecules under investigation.[13] This makes rotational features a unique probe for a certain and univocal identification of molecular species;[13]

ii. rotational transitions involve extremely low energies.[13] Therefore, emission can be observed even in very cold regions of the ISM (*e.g.*, Taurus molecular cloud reaches temperatures as low as 10 K).[8] In addition, when the temperature of the object is below the background radiation field temperature, molecules can be observed in absorption against this continuum.[8] Virtually, this makes radio astronomical detections feasible from any source, either in emission or in absorption.[8]

For instance, assuming NH and HC_9N as the extremes of a reasonable range of sizes for typical interstellar molecules, their lowest rotational transitions fall at *ca.* 1 THz and *ca.* 500 MHz, respectively; given this frequency range, the recording of pure rotational spectra of astronomical species is thus nearly exclusively conducted resorting to telescopes operating in the microwave window.

*See also: `cdms.astro.uni-koeln.de/classic/molecules`.

†Centimeter-wave frequency range refers to 3–30 GHz (10–1 cm), millimeter-wave one to 30–300 GHz (10–1 mm), and submillimeter-wave one to 0.3–3 THz (1–0.1 mm). This entire interval is often indicated as microwave region.

However, there are some limitations to this technique. At first, the requirement for accuracy on the rotational transitions is very demanding.[14] Considering a generic COM, state-of-the-art quantum chemical computations are able to provide predictions with an accuracy well within 0.1%. For NH and HC_9N, this would mean an uncertainty of respectively 1 GHz and 500 kHz on their lowest rotational transitions.[14] However, broadband astronomical surveys present thousands of densely stacked features: an accuracy better than 100 kHz is required for an unambiguous identification.[15] This can be obtained only by resorting to experimental studies, whose uncertainties on rotational lines usually are in the 1–30 kHz range.[14] In addition, the experimental characterization needs to be carried out up to the millimeter- and submillimeter-wave ranges, in order to exploit the wide spectral windows of the last generation radio telescopes.[16,17] Indeed, extrapolations from low-frequency measurements (*i.e.*, those carried out in the centimeter-wave range) often provide inaccurate predictions at higher frequencies.[18,19]

A second drawback of rotational spectroscopy is its blindness to completely and highly symmetric molecules.[8,13] As a matter of fact, the intensity of a rotational transition is proportional to the square of the permanent electric dipole moment components of the molecule;[13] highly simmetric molecular species present either very weak or null values for this quantity.[20] Furthermore, with the increasing of the molecular size, the number of accessible rotational energy levels increases significantly; hence, the population is partioned among them, and the number of molecules emitting/absorbing at a given frequency drops.[8,13] This is a common issue for large COMs and for the species whose transitions show splittings (*e.g.*, hyperfine structures). Finally, molecules in the condensed phase cannot rotate, therefore they cannot be detected using radio astronomy.[8,13]

In the field of rotational spectroscopy, the aforementioned quantum-chemical techniques constitute an essential tool, despite only few species have been detected relying solely on them.[21–23] In fact, computational methods play a threefold role: i. interpretation; ii. complementarity; and iii. prediction.[24] A clear example for the first point is offered when dealing with highly-flexible systems: in these cases, conformational analyses can be of great help for understanding what the most contributing conformers are.[24] Moreover, in some instances, the laboratory studies are not able to derive all the physical quantities for the investigated system: high-level theoretical methods can achieve the required accuracy to complement the experiment.[24] Lastly, the use of computational approaches for the prediction and simulation of experimental spectra is essential in order

to districate within hundreds of different rotational features, or when handling spectroscopically "unknown" species.[24]

At the same time, radio astronomy witholds clear advantages with respect to other astronomical techniques, these ranging from the lack of atmospheric blockage to the classical treatment of the signals. As a consequence, and thanks to the recent improvements (*e.g.*, interferometry and aperture synthesis), last generation radio telescopes provide images with the highest resolution possible at any wavelenght.[25] The importance and the progress of this field is evident: in three occasions, the Nobel prize has been awarded to radio astronomers.[26]

However, despite the accelerated pace of new molecular discoveries, the amount of unidentified spectral features is still overwhelming.[8] In particular, in the radio frequency range, surveys continue to reveal hundreds of lines not ascribable to any molecule based on the current spectroscopic databases.[8] Within this context, a significant number of these transitions may be due to vibrationally excited or rare isotopic species which are still lacking a complete characterization;[27] this notwithstanding, much effort still needs to be dedicated to the extension of the astronomical census, which is far from completion.[28]

Astrochemistry and the origin of life

In the debate that aims to solve one of the most difficult questions of humankind, there is one central historic fact broadly accepted in the scientific community: life was initiated by some autocatalytic chemical system.[29] The competing narratives may be classified in two groups:

i. "replication"-first models: they stress the role of oligomeric compounds, which express the autocatalytic capability through their ability to self-replicate. This idea is centered on the work of Troland from 1917.[29,30] The most famous theory, namely the RNA-world hypothesis, is based on the existence of life forms based on ribonucleic acid (RNA), given the evidence of its twofold role as both catalyst (in peptide synthesis) and genome (in RNA viruses);[31,32]

ii. "metabolism"-first models: they emphasize the emergence of cyclic networks, as articulated by Kauffman in the 1980s,[33] inspired from the metabolic cycles found in all extant life.[29] A strong hypothesis is the one of the iron-sulfur world, where mineral catalysts (such as iron sulfide) present near deep sea hydrothermal vents promote a reversed version of the citric acid cycle, which evolved and increased its complexity until resembling cellular shapes.[34]

However, these approaches are not necessarily mutually exclusive, and both might have been crucial elements for life's emergence.[29]

Once accepted that autocatalysis is a central element for the formation of life from non-living matter (*i.e.*, abiogenesis), it follows the question on the origin of the molecular species, an issue particularly crucial in the RNA-world hypothesis.[35] It is broadly accepted that a heavy bombardment of comets and asteroids in the inner solar system *ca.* 3.5 Gyr ago had a crucial impact on abiogenesis.[36] As a consequence, three main theories have developed to address the sources of organic compounds, namely: i. the exogenous delivery by extraterrestrial objects; ii. the synthesis by bombardment shocks; and iii. the synthesis by other energy sources (*i.e.*, endogenous production).[36]

Within this context, Astrochemistry plays a pivotal role. Hypothetically, a proteinogenic "CHON seed" could have been produced somewhere in space, and it could have been delivered to the early Earth, leading to the birth of life forms.[37] Indeed, the identification of prebiotic species (defined as COMs that are connected *via* known chemical routes to simple biomolecules at the basis of life on Earth[38]) or biological building blocks in the ISM could provide insights into the question of the origin of life. This fueled the search of prebiotic COMs over the last years, which led to the detection of several essential species, like glycolaldehyde,[39] its tautomer ethenediol,[40] urea,[41] cyanamide,[42] hydroxylamine,[43] cyanoacetylene,[44] ethanimine,[45] and so on.[11,12]

Scope and structure of this thesis

Within the vast field of Astrochemistry, the focus of this thesis is on the strategy and the work required to collect the necessary spectroscopic data for potential interstellar species, in order to guide their detection in the interstellar medium. To this purpose, an introduction on the physical and chemical conditions that characterize astronomical objects is required. In Section 1.1, the ISM is described in detail from an astrochemical point of view. Then, Section 1.2 covers the reactivity, in order to understand the processes that generate chemical complexity.

As mentioned at the beginning of this Chapter, the identification of astrochemical molecules is carried out by searching for their rotational features. In this respect, this thesis has focused on the recording and analysis of the rotational spectrum of eight different COMs (ten, if including isomers) in the millimeter- and submillimeter-wave ranges. To support and complete the experiment, an accurate computational characterization has

been carried out for all species, providing reliable spectroscopic parameters and, in some cases, the energetic description of the system. Moreover, to ease the experimental study at high frequency, the spectrum in the centimeter-wave range has been recorded and analyzed for a few species. At the end of this Chapter, in Section 1.3, the investigated molecules and their astrochemical relevance are presented.

Subsequently, the Theoretical Background of this thesis is offered: rotational spectroscopy, computational spectroscopy, and radio astronomy need to be presented in detail. Chapter 2 concerns the first topic, and it explains the concepts of "rotational transitions" (including the rules for observing them), "rotational constants", and many others. This provides the reader with the basis for the understanding of the outcomes of rotational experiments. Quantum-chemical methods are presented in Chapter 3, with the computational methodology being explained in Section 3.8. The latter addresses the determination of molecular structures and spectroscopic parameters. In particular, the former is treated in detail in Chapter 4, where a cost-effective approach suitable for large systems is introduced (Section 4.2). Finally, in Chapter 5, the fundamentals of radio astronomy will be presented, together with the methods for deriving chemical information. Then, the Material and Methods part (Chapter 6) reports an account on the computing facilities, the software, and the spectrometers used for the computational and experimental characterization of the molecules under investigation.

The last part of this manuscript covers the Results and Discussion. The methodology defined in Section 4.2 is validated and tested in Chapter 7, while the computational and experimental outcomes for the investigated molecules introduced in Section 1.3 are presented in Chapters 8–13. To conclude, Chapter 14 poses a note on the astrophysical implications of the reported studies and summarizes all findings of this thesis.

In addition, Appendix A collects the supporting information to the presented data for the sake of reproducibility and completeness. Two further works, published during the Ph.D. project and only moderately congruent with the topic of this thesis, are briefly presented in Sections B.1 and B.2. Analogously, Appendix C shows a Python program developed to ease the analysis of dense vibro-rotational spectra.

1.1 The interstellar medium

Recent cosmological investigations determined that only 4.6% of the total mass of the universe is composed by baryons (this category includes every-

thing made up of protons, neutrons, and electrons, therefore also atoms and molecules), and the 98% of it is constituted by hydrogen and helium.[2] Indeed, the former is the most abundant element in the universe; elements from the second row (*e.g.*, carbon, nitrogen, oxygen, *etc.*) have abundances in the 10^{-3}–10^{-4} range with respect to H.[2] The ratio for heavier elements is even smaller, and it reflects the nucleosynthetic processes occurring in stars—for instance, iron is the element with the highest binding energy per nucleon, and it is relatively abundant (3×10^{-5} with respect to H).[2]

Furthermore, the matter distribution is not uniform over the universe. As example, about 90% of the baryon mass of the Milky Way resides in stars, while the remaining is in interstellar matter, namely the ISM, where it tends to concentrate in clumps and filaments of various sizes, called clouds.[2] About the 1% of the mass of the ISM is composed of dust grains, typically 0.1–1 μm in diameter, which are responsible for a large part of the extinction or scattering of the radiation, in addition to their role in reactivity.[37]

Most of the knowledge about the structure of the Milky Way comes from one of the fundamental discoveries in radio astronomy, the 21 cm (*ca.* 14 GHz) line of neutral H in 1951, and, more recently, the R(0) 2.6 mm (115 GHz) line of ^{12}CO.[26,37] The observation of this transition throughout the sky allows the measurement of maps of ^{12}CO density, which reveal structures not visible with the naked eye.[37] This includes regions of nearly no emission, regions with giant molecular clouds (where star formation begins to occur), as well as remnants of supernovae and stellar winds; even the spiral structure of the Milky Way is mapped through CO.[37]

From a chemical point of view, the modeling of the ISM has focused on four different enviroments:[37]

i. diffuse ISM: areas where the number density is in the order of 1–10^2 cm^{-3}, while temperature is hard to define. Indeed, translational temperatures in shock fronts can reach a few hundreds of K; with these densities, temperature is not a sensible concept (for reference, the density of the air at sea level is *ca.* 2.7×10^{19} cm^{-3});[37]

ii. molecular clouds (MCs): their lifetime ranges from 10^6 to 10^8 years, and these are the regions of new star formation. The temperature is on average close to 10 K (but it can reach 100 K in denser areas), and the number density is 10^2–10^7 cm^{-3}. MCs can differ deeply in size, ranging from 0.1 pc* to 10 pc, and in mass, from a few $M_\odot$† to 10^5–10^6 $M_\odot$.

*The parsec (pc) is a unit of lenght, and it corresponds to 3.086×10^{16} m.

†The solar mass ($M_\odot$) is the mass of the Sun, approximately 1.988×10^{30} kg.

Large MCs extending over 10 pc are known as giant MCs (GMCs);[2,37]

iii. circumstellar medium: areas surrounding a star, whose physicochemical conditions strongly depend on the type and the evolution stage of the star;[37]

iv. photon-dominated regions: areas around stars where strong ultraviolet photon fluxes photodissociate and photoionize all molecules to atoms (*e.g.*, HII regions around young stars).[37]

Among these environments, the attention is drawn to molecular clouds, which harbor the richest interstellar chemistry. They originate from the collapse of diffuse clouds due to gravitational instability or cloud-cloud collision, and their higher density with respect to the diffuse ISM shields the H_2 molecules from interstellar UV radiation, preventing their photodissociation at the center of the cloud.[2] As a matter of fact, MCs derive their name from the prominent presence of molecular hydrogen; also, they are often referred to as "dark clouds", as they can be recognized as dark patches blocking the light from background stars.[2]

In these regions, gas-phase chemistry is driven by cosmic-ray ionization and neutral–neutral reactions involving radical species.[2] Furthermore, in addition to UV shielding, grain accretion is another important process taking place at higher densities: ice mantles are absent in the diffuse ISM, while they become prevalent inside molecular clouds.[37] The interaction between the gas phase and the solid state drives the chemical complexity of the MCs. These topics will be treated in detail in Section 1.2.

Upon gravitional collapse, molecular clouds begin to form dense cores (typically 0.1 pc), characterized by higher temperatures (*ca.* 100 K) and densities (10^5–10^8 cm^{-3}), often associated with a nascent star.[8] Typical masses are in the 1–10 $M_\odot$ range.[2] Before harboring protostars, dense cores are referred to as starless cores or prestellar cores; if a protostar is associated with a dense core, then it is denoted as protostellar core.[2] The chemistry of starless cores is an important research area. Saturated COMs can evolve on the iced surfaces of dust grains, but the cold quiescent environment forces these species to remain adsorbed on the surface.[8] As a consequence, the gas-phase inventory tends to be dominated by unsaturated COMs formed by gas-phase reactions (*e.g.*, long-chain carbon molecules like the cyanopolyynes HC_nN, with $n = 3, 5, 7, 9$).[8]

Some well-studied starless cores are located in the Taurus molecular cloud (TMC-1), namely TMC-1 CP and TMC-1 C.[46] TMC-1 is the prototypical dark cloud, located *ca.* 140 pc away from Earth, extended *ca.* 25–35 pc,

Table 1.1: Total number of detections from the first eight most contributing sources. Detections in closely-located regions have been grouped together.

Source	Type	# mol.
TMC-1	Dark cloud	86
Sgr B2	SFR	70
IRC+10216	Late-type carbon star	66
LOS cloud[a]		42
Orion	SFR	25
L483	Protostellar core	11
G+0.693–0.027	GMC in Sgr B2 complex	10
W51	SFR	8

[a] Detections made in clouds along the line of sight (LOS) to a background source have been consolidated into LOS cloud.

and cold—the temperature ranges between 10 and 20 K.[8,47] The density at the prestellar cores is estimated to be about 10^5 cm^{-3}.[46] Within TMC-1, various carbon-chains have been detected: HC$_n$N, C$_n$H, C$_n$H$_2$ (cumulene type), C$_n$N, C$_n$O, and C$_n$S; also including anions (such as C$_6$H$^-$ and C$_8$H$^-$) and cations (like C$_n$S$^+$).[2,8]

To date, TMC-1 is the richest astronomical object in terms of molecular detections, counting 86 species; the eight most contributing sources are listed in Table 1.1, together with the source type.[8] Starless cores, and, more in general, dark clouds are the only source type in which all five types of species (*i.e.*, neutral, radical, anions, cations, and cyclic) have been detected, therefore suggesting that these regions may be more chemically complex than they appear at first glance, especially if compared with star-forming regions like Sagittarius B2 and Orion.[8]

The star-forming regions (SFRs) introduced above possess indeed a rich chemical inventory (*cf.* Table 1.1). In these enviroments, the icy surfaces of the grains are heated or are subjected to shocks: the trapped COMs are liberated into the gas phase (*i.e.*, sputtering), and become detectable by radio astronomical techniques.[8] Shocks or protostellar outflows create more complicated physical conditions, injecting additional energy and driving chemical reactivity which was sedated in more quiecent conditions, contributing to the availability in the gas phase of heavy elements, *e.g.*, phosphorous.[8,48] As a matter of fact, the chemical composition of SFRs is different from dark clouds: molecules tend to be more saturated, and the percentage of radicals is significantly lower (7%, with respect to 18% of starless regions). Furthermore, no anions have been detected; indeed,

radicals and anions are highly reactive species, and their depletion may be catalyzed by the liberation of several reaction partners from the surface of ice mantles.[8]

An incredibly chemically-rich SFR is located in the GMC Sagittarius B2 (Sgr B2). According to Table 1.1, it is the second most contributing source to astronomical detections; it is located 120 pc from the center of the Milky Way, and it is one of the largest cores in the galaxy, spanning over a region of *ca.* 40 pc; the mass is estimated to be 10^7 $M_\odot$, while the density ranges from 10^6 cm^{-3} to 10^9 cm^{-3}.[49,50] The cloud is divided into three main cores (north, middle, and main), and it counts more than ten HII regions; the temperature is relatively warm (*ca.* 50–70 K).[50] The chemical composition of Sgr B2 is various and it includes different complex molecules, like acetic acid,[51] glycolaldehyde,[39] aminoacetonitrile,[52] and so on.

While the chemistry of the ISM is dominated by carbon, a number of heavier elements can be found in the circumstellar medium of evolved star sources.[8] An example is the carbon-rich star IRC+10216, which counts almost 60 molecular detections in Table 1.1. Funnily enough, this is the environment where the largest molecule without a carbon atom has been identified, *i.e.*, silane (SiH_4), by Goldhaber and Betz.[53] These astronomical objects inject heavier atoms in the chemical inventory, such as Mg, Fe, and Al. These elements can be detected as components of molecules, *e.g.*, MgC_nN^+, FeCN, AlNC, and various others.[54–56]

Despite most of the molecular discoviers now originate from SFRs, prestellar cores, and carbon stars (*cf.* Table 1.1), the first detections were made in the line of sight (LOS) diffuse clouds that are ubiquitos in the Milky Way.[8] These low-density envirnoments still contribute to the total number of astronomical species, and are collected together as "LOS clouds" in Table 1.1. Indeed, higher temperatures (about 100 K) and a lower shielding from radiation with respect to (G)MCs enhance the chemical reactivity of these objects.[8] Furthermore, LOS clouds are often optically thin, therefore also allowing the use of short wavelenghts, which is inefficient in SFRs and dark clouds. The latter, in fact, would require both an optically thin absorbing medium and a background source of radiation.[8]

1.2 Reactivity in molecular clouds

Having established the chemical complexity of the interstellar medium and circumstellar shells, the goal shifts to the understanding of the chemical processess that generate detected species, focusing again on molecular clouds. As mentioned above, MCs are weakly ionized plasma: ions play

Table 1.2: Common gas-phase reactions in the interstellar medium and their typical rate. Reproduced from Ref. 57.

Name	Reaction	Typical rate
Photodissociation	$AB + h\nu \longrightarrow A + B$	10^{-9} s^{-1}
Neutral–neutral	$A + B \longrightarrow C + D$	4×10^{-11} cm^3 s^{-1}
Ion–molecule	$A^+ + B \longrightarrow C^+ + D$	2×10^{-9} cm^3 s^{-1}
Charge transfer	$A^+ + B \longrightarrow A + B^+$	10^{-9} cm^3 s^{-1}
Radiative association[a]	$A + B \longrightarrow AB + h\nu$	
Dissociative recombination	$A^+ + e^- \longrightarrow C + D$	10^{-7} cm^3 s^{-1}
Associative detachment	$A^- + B \longrightarrow AB + e^-$	10^{-9} cm^3 s^{-1}
Collisional association	$A + B + C \longrightarrow AB + C$	10^{-32} cm^6 s^{-1}

[a] The rate of this reaction is highly reaction-specific.

a central role in their gas-phase chemistry.[57] In fact, molecular synthesis through ion–molecule reactions in the gas phase has been recognized at the beginning of the 1970s, and it successfully explained the basic molecular abundances observed in the ISM.[58–60] Later, the contribution of neutral–neutral reactions was established, even from low-temperature environments.[61] Thanks to the availability of experimental rate coefficients of gas-phase reactions (or, alternatively, their estimates based on theoretical considerations and quantum-chemical calculations), complex reaction networks developed quite readily.[2]

Within this context, the extreme physical conditions of the interstellar medium explained above play a fundamental role. Low temperatures and low pressures strongly affect the reactivity of atomic and molecular species, emphasizing the importance of kinetics over thermodynamics in molecular clouds.[2,57] As an example, one can consider the collision of two hydrogen atoms to form H_2. This tentative molecule is highly unstable, and it cannot discard the excess energy produced by its formation, unless: i. a third body absorbs it; ii. it spontaneously emits a photon.[2] In interstellar conditions, the probability of a collision with a third body is negligible, and the rate costant for radiative stabilization of H_2 is of the same order of magnitude.[2] Therefore, the molecule is eventually dissociated back to two hydrogen atoms. In fact, the most general type of binary reactions which are effective in the gas phase take the form

$$A + B \longrightarrow C + D\,,$$

where the redundant energy produced by the reaction can be taken away as kinetic energy of the products moving in opposite directions.[2] Common

types of gas-phase reactions are listed in Table 1.2, together with an estimate of their typical rate.

An important costraint on reactions in interstellar clouds is also given by the low kinetic temperatures. Endothermic reactions that require an external energy input do not take place; furthermore, also exothermic reactions which have an energy barrier toward the formation of the products do not proceed.[2] In other words, for a reaction to occur, it must be exothermic, and all the energetic barriers must be submerged (*i.e.*, they lie below the sum of the reagents' energies).[2]

As mentioned above, the UV-shielding occurring in MCs placates the photodissociation processes which dominate the diffuse ISM; thus, cosmic rays play a key role as ionizing agent.[62] The H_2 ionization is quickly passed on to other (protonated) molecular species, as shown in Figure 1.1, where the action of collecting all emitted electrons is carried out by polycyclic aromatic hydrocarbons (PAHs).[57]

PAHs represent a family of hydrocarbon molecules with carbon atoms arranged in a honeycomb-like structure through the fusion of six-membered aromatic rings with peripheral hydrogen atoms. The carbon atoms are bonded with each other by three σ bonds, resulting in planar structures.[63] Despite the hydrogen-dominated conditions, PAHs exist in high abundance,[2,64] and their presence has been proven by several well-studied infrared emission features.[63] A few examples of PAH molecular structures are reported in Ref. 63 and references therein.

PAHs play an essential role in the chemical evolution of the ISM by accepting the electrons expelled in the ionization of H_2. Then, PAH anions can be neutralized by photodetachment and by recombination with metal ions (represented by Me^+ in Figure 1.1) and protonated molecular ions (HCO^+ *et similia*).[57] The degree of ionization x of the cloud depends on

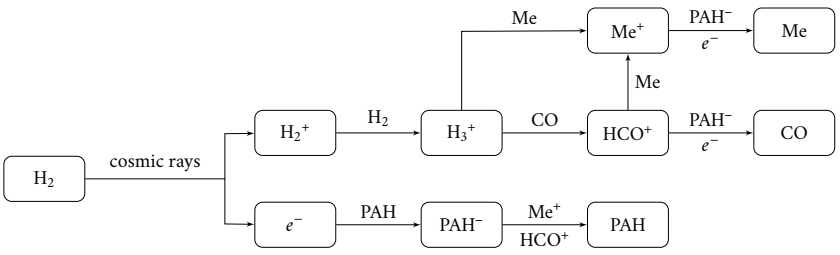


Figure 1.1: Scheme of the ionization mechanism occurring in dense clouds, with HCO^+ and Me^+ representing respectively a general protonated species and a metal ion. Reproduced from Ref. 57.

the total number of molecules n according to

$$x \approx \left(\frac{\zeta^{\mathrm{cr}}}{k_{\mathrm{n}} n}\right)^{1/2}, \tag{1.1}$$

where ζ^{cr} is the cosmic-ray ionization rate and k_{n} is the neutralization coefficient for molecular cations and PAH anions.[57] From Equation (1.1), it derives that the denser the gas, the lower the degree of ionization.[57] In most of the molecular gas in the galaxy, PAHs neutralization is dominated by the absorption of visible photons. However, in dense cores (which have a high visual extinction due to dust particles), PAH anions are the dominant carriers of negative charge.

The presence of PAHs strongly affects the physicochemical characteristics of the cloud: in their absence, ionization is transferred to metals, and recombination of metal ions and electrons balance cosmic-ray ionization.[57] However, this is a slower process if compared with recombination with PAH anions, and the degree of ionization is considerably higher when PAHs are absent.[57] The degree of ionization plays a key role in the determination of molecular abundances; hence, the study of these large aromatic systems is fundamental in the accurate modeling of astronomical objects.[57]

Thus far, gas-phase chemical reactions occurring in interstellar clouds appear as the most basic reactions in chemistry. However, the experimental replication of such elementar reactions is extremely challenging: a typical H_2 density of 10^4 cm^{-3} found in a general MC can be seen as a vacuum of 3.8×10^{-10} Pa.[2] Indeed, various additional factors (*e.g.*, third-body effects, surface catalysis, *etc.*) veer off the laboratory conditions from the actual ones: theoretical and quantum-chemical studies play an integrant role in the modeling of astronomical objects.[2]

In addition to gas-phase reactions, gas–grain processes also play a key role in the chemical complexity of MCs.[57] Dust grains are composed of non-volatile silicates and carbonaceous compounds (including PAHs), and have a typical radius of 0.1 μm. In molecular clouds, this core is surrounded by a mantle of H_2O ice, which contains various molecules.[2] Moreover, dust grains scatter and absorb interstellar UV radiation, contributing to shield the center of molecular clouds from photodissociation.[2]

Indeed, molecular synthesis through grain-surface chemistry has been suggested from the 1970s, right after the discovery of interstellar molecules.[65–70] Reaction networks similar to gas-phase ones have been developed: however, reaction probabilities on grain surfaces are much more uncertain, laboratory experiments are more challenging, and theoretical estimates are less reliable.[2] For these reasons, grain-surface models received less attention than gas-phase models until the beginning of the 1990s. At that point,

although gas-phase chemical systems described well quiecent MCs, they failed to reproduce the abundances of saturated complex organic molecules in SFRs.[2] In addition, absortption of molecules on dust grains (*i.e.*, depletion) has been observationally proved in 1999,[71] implying that chemical reactions occur on the grain surface.[2] The importance of gas–grain interactions is now recognized as critical; from this, large chemical models have been developed to accurately simulate interstellar abundances.[72–74]

Depleted molecules undergo chemical reactions according to two different mechanisms.[2] In the Langmuir–Hinselwood hypothesis, absorbed molecules migrate on the grain surface to encounter a reaction partner. Instead, in the Rideal mechanism, gaseous reagents directly hit atoms or molecules on the grain surface to form products. In the interstellar conditions, the former is supposed to be the most relevant.[2] Therefore, surface mobility becomes an important factor: the absorption site is represented by a local minimum of the interaction potential between the depleted species and the surface.[2] In this context, the motion across the surface is represented *via* the overcoming of an energy barrier between two local minima; for light species such as H atoms, this considers also the possibility of tunnelling between absorption sites.[2] It has been proven that—even if not as fast as a motion relying mainly on tunnelling would do—hydrogen atoms readily move on the grain surface even at temperatures as low as 10 K.[2]

An important feature of reactions occurring on the grain surface is the presence of a third body that can absorb the excess energy, therefore allowing recombination reactions that yield only one product.[2] In fact, grain-surface catalysis paves the way to complex species; however, the presence of an energy barrier still remains an issue, with reactions between closed-shell molecules still being prohibited.[2]

To conclude, the intricate chemical reactivity and the harsh physical conditions of interstellar molecular clouds lead to the formation of exotic molecules that are exceptionally unstable under terrestrial conditions, thus often requiring spectroscopists to exploit complicated generation techniques for their production in laboratory.

1.3 Investigated molecules

The choice of the species to be investigated in this thesis (represented in Figure 1.2) has been driven by various astrochemical issues.

The first molecule listed is allenylacetylene ($H_2CCCHCCH$), which has been selected given its recent detection, which has been carried out in the

centimeter-wave frequency range.[75] Only a few pure hydrocarbons have been identified in the interstellar medium so far (*e.g.*, propene[76]), and this is mainly due to their typically low permanent electric dipole moment. Only the C_nH radicals and anions have moderate-to-large values of this property, and have been detected up to $n = 8$ in the interstellar clouds.[77–89] Instead, pure hydrocarbons like C_2H_2, C_2H_4, and C_2H_6 lack a permanent dipole moment, and therefore are often studied through their derivatives.[75] An example is provided by vinyl acetylene (CH_2CHCCH), which has been found only recently in TMC-1[90] with an abundance twice that of the cyano-derivative, CH_2CHCN, whose detection was confirmed during the early years of radio astronomy.[91] This is mainly due to the huge impact of the cyano group in the value of the electric dipole moment: the transitions of vinyl cyanide are *ca.* 80 times stronger than vinyl acetylene ones. Due to the elusiveness of pure hydrocarbons, their role in the chemistry of dark clouds is still poorly understood.[75] In order to improve the possibility of the identification of allenylacetylene in different astronomical sources, thus exploiting the wide frequency range of last generation radio telescopes,[16,17] the extension of its spectroscopic characterization up to the submillimeter-wave range is deserved.

On a separate note, PAHs are broadly accepted as a common class of interstellar molecules; however, only one small PAH has been identified in the ISM (*i.e.*, indene, see Ref. 92), in addition to their archetype, benzene,[93–95] as well as a few close relatives (*ortho*-benzyne[96] and cyclopentadiene[92]). Many PAHs are poor candidates—as seen for pure hydrocarbons—due to their low/null dipole moment and unfavorably large rotational partition functions. Consequently, the attention moved to their cyano derivatives.[93] In 2018, McGuire *et al.* successfully identified the first cyano-substituted aromatic ring, benzonitrile,[93] followed by the recent detections of 1- and 2-cyanonaphtalene,[97] 1- and 2-cyanocyclopentadiene,[98,99] and

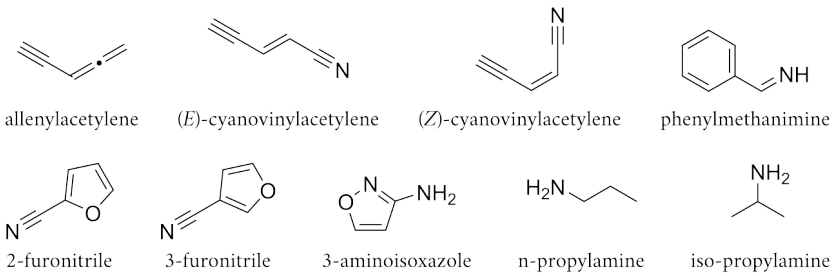


Figure 1.2: Molecules investigated in this thesis.

2-cyanoindene.[100] In addition to the cyano-derivatives, several substituted aromatic species have also beed identified, such as 1- and 2-ethynylcyclopentadiene, ethynyl benzene,[101] and fulvenallene.[102] However, actual chemical models do not have sufficient pathways to reproduce the abundances of these aromatic species.[103] In part, this is due to the lack of observational constraints on potential carbon-chain precursors: abundance measurements of partially saturated species is thus warranted.[103] Among these, the *E* and *Z* isomers of cyanovinylacetylene ($NC(CH)_2CCH$, *cf.* Figure 1.2) are potential candidates, also given the recent detection of the *E* isomer towards TMC-1.[103] In this thesis, the rotational spectrum of both isomers has been recorded and analyzed up to 400 GHz, providing the astronomers with the necessary data to pursue the astronomical search in the millimeter-/submillimeter-wave range.

Another species investigated in this thesis is phenylmethanimine (C_6H_5-CHNH, see Figure 1.2), whose preliminary work carried out during the Master's thesis of the author provided the foundation for further studies, which have been conducted at the beginning of the Ph.D. project. The reason behind the choice of phenylmethanimine lies in its connection to nitrogen-bearing PAHs, and, in particular, to benzonitrile—the latter being, disputably, one of the biggest astrochemical findings of that time,[93,104] considering that cyano-derivatives of aromatic species are excellent observational proxies for their hydrocarbon counterparts.[100] Phenylmethanimine could be generated by hydrogenation on grains (an efficient process[2] which could lead to the formation of partially saturated molecules, see Ref. 105) of benzonitrile. Therefore, the measurements have been extended for both *E* and *Z* isomers, which have been studied at higher frequency, thus providing a reliable basis to guide its astronomical search. In addition, the generation process has been explored in detail.

More recently, motivated by the new detections of aromatic nitriles,[97–99] the attention shifted to cyano-substituted furans, namely 2- and 3-furonitrile (C_4H_4OCN, see Figure 1.2). Their characterization represents the first step towards the investigation of the possible presence of furan in the ISM, a key prebiotic molecule at the basis of simple sugars such as ribose and deoxyribose, the backbone molecules of RNA and DNA.[106] So far, the low values of the electric dipole moment and the large partition function of furan have hampered its detection;[93,107–110] the study of cyano-derivatives such as 2- and 3-furonitrile is therefore warranted. Indeed, these species and their protonated forms have been recently addressed in a quantum-chemical investigation which pointed out the lack of information on their rotational parameters.[111] After conducting the experimental work on these

species, a comprehensive article on 2-furonitrile has been published.[107]

Subsequently, the focus has been moved to prebiotic species, and, in particular, to those related to the RNA-world hypothesis.[31,32] Although being a broadly accepted hypothesis, it is still discussed how RNA units could have formed under the pysicochemical conditions of the early Earth. In fact, the assembling of RNA from a ribose, a phosphate unit, and a nucleobase has been demonstrated to be inefficient;[112,113] hence, alternative pathways have been formulated. In the model suggested by Powner *et al.*, the starting point consists of simple chemical species (*e.g.*, glycolaldehyde, cyanamide, glyceraldehyde, and cyanoacetylene), and 2-aminooxazole constitutes a key intermediate.[114] This mechanism is quite efficient for pyrimidine bases, but it does not account for the formation of purine bases.[115] In 2019, a unified and more comprehensive route has been introduced to explain the simultaneous formation of pyrimidine and purine nucleotides by Becker *et al.*, which still requires small molecules and ribose.[115] To date, it appears to be one of the most valid scenarios investigated. In this pathway, the key reactive species is now represented by 3-aminoisoxazole,[115] represented in Figure 1.2, which has thus been selected because of its high potential prebiotic role.[115–117]

Finally, among the vast range of molecules in some fashion connected with biological building blocks, primary amines play an important role. The NH_2 moiety is found in several compounds, such as neurotrasmitters, hormones, and alkaloids. Moreover, they are closely related to amino acids: since the disproved detection of glycine,[118–120] the search for these species has become a challenge, and effort has also been dedicated to the search for potential precursors. Within this context, the smallest primary amine, *i.e.*, methylamine, has been detected in the ISM in 1974,[121,122] and it has been suggested to play a key role towards the eventual formation of glycine in the ISM;[123] analogously, more complex amines may be seen as precursors of other proteinogenic amino acids. Furthermore, Förstel *et al.* observed the formation of methylamine and ethylamine by exposing ammonia and methane ices to a ionizing radiation, proposing that propylamine could be formed by increasing the methane-to-ammonia ratio.[124] In addition, the detection of methylamine has been followed by the discovery of interstellar vinylamine and ethylamine in 2021.[125] Given these premises, iso-propylamine and n-propylamine (namely $(CH_3)_2CHNH_2$ and $CH_3(CH_2)_2NH_2$, respectively; *cf.* Figure 1.2) are ideal candidates for an astronomical search, and are the last molecules investigated in this thesis.

Theoretical Background

Chapter 2

Rotational Spectroscopy

Due to the intrinsic uniqueness of rotational features and the low energy levels involved, rotational spectroscopy is an essential tool in radio astronomy for the molecular exploration of the interstellar medium. In this Chapter, the reader will be provided with the fundamental theoretical background for rotational spectroscopy, from the definition of the spectroscopic parameters to the rules for observing rotational transitions, in order to provide the necessary knowledge to interpret the experimental outcomes.

Molecules in the gas phase are in constant motion with velocities that depend on their thermal energy. A part of this is collected as rotational energy, which is related to their spinning movement.[126] If a single and isolated molecule is considered, its rotational energy is quantized, leading to individual rotational energy levels. The collection of the transitions between these levels constitutes what is known as the rotational spectrum of the molecule considered.[126] To derive the quantum mechanical description, the starting point is the knowledge of the classic mechanics of rotating bodies.[13] The general procedure for understanding the rotational motion starts from the expression of total energy, E, which is the sum of the total kinetic energy T and the total potential energy V[13]

$$E = T + V. \tag{2.1}$$

If we are interested in "pure" rotation, the total energy contains only the kinetic term T, which includes both the rotation and the translational motion.[13] The best way to get rid of the latter contribution is to resort to a molecule-fixed axis system with the origin at the center of mass. Without going into details, this choice allows for separating the two contributions, with the term due to translation motion no longer considered.[13]

Thus, the classical rotational energy is expressed as[13]

$$E_{\text{rot}} = \frac{1}{2\mathbf{I}}\,\mathbf{P}^2\,, \tag{2.2}$$

where the angular momentum $\mathbf{P}$ is defined as the vector product between the inertia tensor $\mathbf{I}$ and the angular velocity $\boldsymbol{\omega}$ as follows[13]

$$\mathbf{P} = \mathbf{I} \times \boldsymbol{\omega}\,. \tag{2.3}$$

Before moving from the classical to the quantum-mechanical treatment, another simplification is required. This consists in the use of the principal inertia system, a coordinate system in which $\mathbf{I}$ is diagonal, and the three diagonal elements (I_x, I_y, I_z) are called principal moments of inertia.[13] The x, y, z notation is substituted by the a, b, c axes, which are conventionally chosen so that they verify the relationship[13]

$$I_a \leq I_b \leq I_c\,. \tag{2.4}$$

The values of I_a, I_b, I_c are used to classify the type of rotor, as reported in Table 2.1.

Once the classical equation for the rotational energy is derived, the next step consists in the introduction of the quantum mechanical operators. By doing that, from Equation (2.2) it follows[13]

$$\hat{\mathscr{H}}_{\text{rot}} = \frac{\hbar^2}{2\mathbf{I}}\hat{\mathbf{J}}^2\,, \tag{2.5}$$

where $\hat{\mathbf{J}}^2 = \sum_g \hat{J}_g^2$, with g running over the inertial axes. From Equation (2.5), it is clear that the rotational Hamiltonian $\hat{\mathscr{H}}_{\text{rot}}$ commutes with $\hat{\mathbf{J}}^2$, while the following commutation rules for the projections of the latter

Table 2.1: Classification of rotors based on the principal moments of inertia.

Rotor	Moments of inertia
Linear	$I_a = 0, I_b = I_c$
Spherical	$I_a = I_b = I_c$
Symmetric, oblate	$I_a = I_b < I_c$
Symmetric, prolate	$I_a < I_b = I_c$
Asymmetric	$I_a \neq I_b \neq I_c$

operator can be derived[13]

$$\begin{aligned}
&[\hat{\mathbf{J}}^2, \hat{J}_\gamma] = 0\,, \text{ with } \gamma = x, y, z \text{ or } X, Y, Z\,;\\
&[\hat{J}_\gamma, \hat{J}_\eta] = i\epsilon_{\gamma\eta\delta}\hat{J}_\delta\,, \text{ with } \gamma, \eta, \delta = X, Y, Z\,;\\
&[\hat{J}_\gamma, \hat{J}_\eta] = -i\epsilon_{\gamma\eta\delta}\hat{J}_\delta\,, \text{ with } \gamma, \eta, \delta = x, y, z\,,
\end{aligned} \tag{2.6}$$

where $\epsilon_{\gamma\eta\delta}$ is the three-index Levi–Civita symbol. In the notation above, lower-case letters represent the molecule-fixed coordinate system, while the upper-case letters represent the space-fixed coordinate system.* Projections of $\hat{\mathbf{J}}$ on different coordinate systems commute.[13]

Now, considering the Equations (2.5) and (2.6), $\hat{\mathscr{H}}_{\text{rot}}$, in addition to commute with $\hat{\mathbf{J}}^2$, might also commute with only one of its component on each coordinate system, which can be arbitrarly labeled as $\hat{J}_z$ and $\hat{J}_Z$.[13] To give an example, in the case of symmetric-top rotors, where two of the principal inertia moments are equal to each other, Equation (2.5) can be re-arranged as follows (the case of the prolate-top rotor is used, with $I_c = I_b$)[13]

$$\hat{\mathscr{H}}_{\text{rot}} = \frac{\hbar^2}{2I_b}\hat{\mathbf{J}}^2 + \frac{\hbar^2}{2}\left(\frac{1}{I_a} - \frac{1}{I_b}\right)\hat{J}_a^2\,. \tag{2.7}$$

In this case, $\hat{\mathscr{H}}_{\text{rot}}$ commutes with $\hat{\mathbf{J}}^2$ and $\hat{J}_a$, in addition to $\hat{J}_Z$. This means that these operators have a common set of eigenfunctions. This opens the way to a very simply strategy to solve the eigenvalue equation associated to $\hat{\mathscr{H}}_{\text{rot}}$, because the matrix elements of the angular moment operators are known[13]

$$\begin{aligned}
&\langle JKM|\hat{\mathbf{J}}^2|JKM\rangle = \hbar^2 J(J+1)\,;\\
&\langle JKM|\hat{J}_z|JKM\rangle = \hbar K\,;\\
&\langle JKM|\hat{J}_Z|JKM\rangle = \hbar M\,,
\end{aligned} \tag{2.8}$$

with

$$\begin{aligned}
J &= 0, 1, 2, 3, \ldots\,;\\
K &= -J, \ldots, 0, \ldots, +J\,;\\
M &= -J, \ldots, 0, \ldots, +J\,,
\end{aligned} \tag{2.9}$$

where J is the rotational quantum number, K is the quantum number associated to the projection of J along the z axis, M is the quantum number associated to the projection of J along the Z axis, and $|JKM\rangle$ represents a common set of eigenfunctions for the three operators, known as

*The transformation between the molecule-fixed coordinate system (r_g, with $g = x, y, z$) and the space-fixed one (r_F, with $F = X, Y, Z$) is given by $r_F = \sum_g \Phi_{Fg} r_g$, where Φ_{Fg} is the direction cosine matrix which depends on Euler's angles θ, φ, χ defining the orientation of r_g relative to r_F.[13]

symmetric-top eigenfunctions.*[13] In this case, thanks to the prolate-top rotor symmetry, the eigenvalue equation associated to the Hamiltonian of Equation (2.7) can be obtained using the known eigenvalues of $\hat{\mathbf{J}}$ and its projections; this leads to the definition of the rotational energy levels E_{JK}, where the number of M values defines the degeneracy of the levels[13]

$$E_{JK} = hBJ(J+1) + h(A-B)K^2\,, \tag{2.10}$$

where A and B are the rotational constants, defined as (in frequency units)[13]

$$\begin{aligned} A &= \frac{h}{8\pi^2 I_a}\,; \\ B &= \frac{h}{8\pi^2 I_b}\,. \end{aligned} \tag{2.11}$$

In this thesis, the molecules under consideration are asymmetric rotors (see Table 2.1), which, by definition, have no null and no identical principal moments of inertia, thus preventing to re-arrange Equation (2.5) as a function of commuting operators, like Equation (2.7), and deeply increasing the complexity. This arises from the fact that no internal component of the angular momentum is a constant of motion.[13]

2.1 The asymmetric rotor

In the rigid rotor approximation, for an asymmetric-top rotor Equation (2.5) can be written as[13]

$$\hat{\mathscr{H}}_{\text{rot}} = A\hat{J}_a^2 + B\hat{J}_b^2 + C\hat{J}_c^2\,, \tag{2.12}$$

where the rotational constants A and B are defined in Equation (2.11), and the rotational constant C can be analogously defined as[13]

$$C = \frac{h}{8\pi^2 I_c}\,. \tag{2.13}$$

Furthermore, the degree of asymmetry can be described in terms of Ray's parameter, κ, defined as follows[13]

$$\kappa = \frac{2B - A - C}{A - C}\,. \tag{2.14}$$

*$|JKM\rangle = N_{J,K,M}(\sin\frac{\theta}{2})^{|K-M|}(\cos\frac{\theta}{2})^{K+M}F(\sin^2\frac{\theta}{2})e^{iK\varphi}e^{iM\chi}$, where $N_{J,K,M}$ is a normalizing constant satisfying $N^2_{J,K,M}\langle JKM|JKM\rangle = 1$, $F(\sin^2\frac{\theta}{2})$ is a particular series of $\sin^2\frac{\theta}{2}$, and θ, φ, χ are Euler's angles.[13]

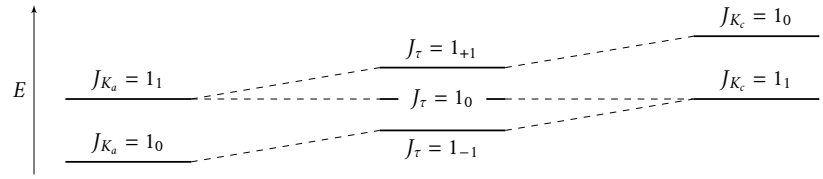


Figure 2.1: Schematic correlation diagram of the asymmetric-top rotor energy levels with prolate and oblate symmetric-top limits.

It is clear from Equation (2.14) that an asymmetric-top rotor with $\kappa \approx -1$ is close to a prolate symmetric-top one, while if $\kappa \approx +1$ it can be approximated to an oblate symmetric-top rotor.[13] The main difference in the energy levels between symmetric-top and asymmetric-top rotors with $\kappa \approx +1, -1$ is in the $\pm K$ sublevels, which are doubly degenerate only in the symmetric-top one. Indeed, given a value of J, the asymmetric-top rotors have $2J + 1$ distinct sublevels, while the symmetric-top ones $J + 1$ (see Figure 2.1 for a schematic representation).[13]

Unlike the other types of rotors, the theoretical treatment of asymmetric rotors is more complex because of the lack of a closed equation for the energy levels.[13] Indeed, in contrast with the symmetric rotors, the rotational Hamiltonian cannot be written in terms of $\hat{\mathbf{J}}$ and one of its projections.[13] Therefore, the corresponding rotational Schrödinger equation has to be solved, and the resolution starts from the "corresponding" symmetric-top rotor, by diagonalizing the Hamiltonian matrix in the $|JKM\rangle$ basis.[13] The Hamiltonian matrix obtained is still diagonal in J and independent of M, but there are non-null off-diagonal terms in K.[13] Hence, the K quantum number is not a "good" quantum number; instead, two *pseudo*-quantum numbers are defined, K_a and K_c, which, together with J, label the energy levels of the asymmetric rotors.[13] K_a and K_c represent the values assumed by $|K|$ in the prolate and oblate limits, respectively. The values of K_a, K_c follow these relations[13]

$$\begin{aligned} K_a + K_c &= J, J + 1\,; \\ \tau = K_a - K_c &= -J, \ldots, 0, \ldots, +J\,, \end{aligned} \tag{2.15}$$

where the *pseudo*-quantum number τ can also be employed in the description of the energy levels of asymmetric-top rotors.[13]

For low J values, Equation (2.12) can be rearranged as follows in order to facilitate the calculation of the energy levels[13]

$$\hat{\mathscr{H}}_{\text{rot}} = \frac{1}{2}(A + C)\hat{\mathbf{J}}^2 + \frac{1}{2}(A - C)\hat{\mathscr{H}}_\kappa\,, \tag{2.16}$$

Table 2.2: Representations: corrispondences between the inertial axes a, b, c and x, y, z.

	I^r	II^r	III^r	I^l	II^l	III^l
x	b	c	a	c	a	b
y	c	a	b	b	c	a
z	a	b	c	a	b	c

where the reduced Hamiltonian $\hat{\mathscr{H}}_\kappa$ is defined as[13]

$$\hat{\mathscr{H}}_\kappa = \hat{J}_a^2 + \kappa \hat{J}_b^2 - \hat{J}_c^2 \,. \tag{2.17}$$

The advantage of this formulation is that the reduced energies (eigenvalues of the reduced Hamiltonian) only depends on the value of κ.[13] However, a closed-form solution for the corresponding eigenvalue equation is not possible; therefore, a general expression for the wave functions is not available. To overcome this problem, the asymmetric-top rotor wave functions are represented by a linear combination of the symmetric-top rotor ones

$$|J\tau M\rangle = \sum_{J,K,M} a_{JKM} \, |JKM\rangle \;, \tag{2.18}$$

where the a_{JKM}'s are numerical coefficients.[13] The resulting eigenvalues, expressed as function of J, are[13]

$$E_{\text{rot}} = \frac{1}{2}(A + C)J(J + 1) + \frac{1}{2}(A - C)E^\kappa_{J,\tau} \,, \tag{2.19}$$

with $E^\kappa_{J,\tau}$ denoting the $2J + 1$ solutions of the secular determinant

$$|\mathbf{E}_\kappa - \mathbf{I}\lambda| = 0 \,, \tag{2.20}$$

where $\mathbf{E}_\kappa$ has nonvanishing elements for $\Delta K = 0, \pm 2$, $\mathbf{I}$ is a unit matrix and λ are the allowed energy levels for an asymmetric rotor.[13] Being $\hat{\mathscr{H}}_\kappa$ not diagonal if $\kappa \neq \pm 1$, an explicit expression for its eigenvalues, and thus for the rotational energy, can be obtained only for low values of J.[13]

The solution of Equation (2.20) can be eased by choosing a suitable representation, *i.e.*, the assignment to the a, b, c axes of the x, y, z ones:[13] the six possible ways are collected in Table 2.2. Clearly, the energy levels are independent of the chosen representation; however, for certain ranges of κ, some simplifications can take place.[13] Knowing that the energy matrix of a prolate rotor ($\kappa = -1$) in the I^r representation is diagonal, for a *quasi*-prolate rotor ($\kappa \approx -1$), the I^r representation is the most suitable, since

the matrix has smaller off-diagonal elements.[13] Analogously, for a *quasi*-oblate rotor, the preferred representation is the III^l one. Indeed, most of the asymmetric-top rotors can be handled by resorting to the I^r and III^l representations.[13]

A further simplification may be obtained by considering the symmetry properties of the Hamiltonian, which can be determined by studying the ellipsoid of inertia, which belongs to the D_2 symmetry point group (see Table 2.3 for the character table). Thus, it is symmetric when the identity operation E and/or the 180° rotation operations C_2^g are applied, with g referring to the inertial axes a, b, c.[13] The four operations transform the angular momentum according to[13]

$$\begin{aligned} E &: P_a \rightarrow P_a, P_b \rightarrow P_b, P_c \rightarrow P_c\,; \\ C_2^a &: P_a \rightarrow P_a, P_b \rightarrow -P_b, P_c \rightarrow -P_c\,; \\ C_2^b &: P_a \rightarrow -P_a, P_b \rightarrow P_b, P_c \rightarrow -P_c\,; \\ C_2^c &: P_a \rightarrow -P_a, P_b \rightarrow -P_b, P_c \rightarrow P_c\,. \end{aligned} \tag{2.21}$$

The Hamiltonian is invariant under these operations and, therefore, it transforms as the A irreducible representation of the D_2 symmetry group. Now, it can be useful to build a set of wave functions which also transform as irreducible representations of this group. By doing that, the general matrix element $\langle k|\hat{\mathcal{H}}_{\text{rot}}|i\rangle$ would be nonzero only between states of the same symmetry. As a consequence, the secular determinant of Equation (2.20) would be factorized into four independent subdeterminants, thus facilitating the diagonalization problem.

Wang's linear combination of the symmetric-top rotor functions $|JKM\rangle$ belong to the D_2 group, and are defined as[13]

$$\begin{aligned} |JKM\gamma\rangle &= \frac{1}{\sqrt{2}}\left[\,|JKM\rangle + (-1)^{\gamma}\,|J-KM\rangle\,\right],\ \text{for } K \neq 0\,; \\ |J0M0\rangle &= |J0M\rangle\,,\ \text{for } K = 0\,, \end{aligned} \tag{2.22}$$

Table 2.3: Character table of the D_2 symmetry point group.

	$\hat{E}$	$\hat{C}_2^a$	$\hat{C}_2^b$	$\hat{C}_2^c$	K_a, K_c
A	1	1	1	1	*ee*
B_a	1	1	−1	−1	*eo*
B_b	1	−1	1	−1	*oo*
B_c	1	−1	−1	1	*oe*

where $\gamma = 0, 1$ and K takes only positive values. By identifying as $\mathbf{X}$ the Wang symmetrizing transformation matrix that performs the change of basis, Equation (2.22) can be re-written as[13]

$$|JKM\gamma\rangle = \tilde{\mathbf{X}}\,|JKM\rangle \ . \tag{2.23}$$

Finally, we can build the asymmetric rotor functions as linear combination of Wang functions[13]

$$|J\tau M\rangle = \sum_K a_K^{J\tau M}\,|JKM\gamma\rangle \ , \tag{2.24}$$

where the sum is carried out only over Wang's functions which transform like the same irreducible representation of the D_2 group as $|J\tau M\rangle$ does. Therefore, there is a significant decrease in the number of terms of Equation (2.24) with respect to the previous Equation (2.18). In the new basis, the new energy matrix $\mathbf{E}'_\kappa$ can be defined as[13]

$$\mathbf{E}'_\kappa = \tilde{\mathbf{X}}\mathbf{E}_\kappa\mathbf{X} = \mathbf{E}_\kappa^+ + \mathbf{O}_\kappa^+ + \mathbf{E}_\kappa^- + \mathbf{O}_\kappa^- \ , \tag{2.25}$$

where the four submatrices group the elements according to the evenness and oddness of K (E and O, respectively) and the $\gamma = 0, 1$ values (+ and −, respectively).

Some symmetry considerations can now be drawn. If the limit case of the symmetric prolate-top rotor is considered, to even values of K corresponds an even wave function (*e*); instead, for odd values of K, it corresponds an odd wave function (*o*), as demostrated by[13]

$$C_2^a : |JKM\rangle \rightarrow (-1)^K\,|JKM\rangle \ . \tag{2.26}$$

Analogously, it can be verified that for the symmetric oblate rotor even and odd values of K lead to an *o* and *e* wave function, respectively. Moving to the asymmetric top, if K_a is even and K_c is odd, the asymmetric wave function is symmetric with respect to a rotation along the a axis and antisymmetric with respect to a rotation about c. Since $C_2^a C_2^c = C_2^b$, the rotation about b must be antisymmetric. By looking at the character table of the D_2 group in Table 2.3, it is evident that the wave function *eo* (K_a, K_c) transforms as the B_a representation. This conclusion can be drawn for all four irreducible representations of D_2, which are collected in the last column of Table 2.3. Furthermore, all symmetrized Wang functions $|JKM\gamma\rangle$ occurring in the same submatrice of Equation (2.25) share the same symmetry;[13] thus, this can be used to classify the submatrices according to their symmetry. Being Wang's functions relative to x, y, z axes, it is

mandatory to know the representation to apply (*c.f.* Table 2.2) to obtain the symmetry information.[13] In summary, the knowledge of the evenness and oddness of K_a and K_c gives access to the symmetry classification of the level, and the factorization of the secular determinant of Equation (2.20) sensibly reduced its size. These developments will be of great help in the derivation of the selection rules.

Selection rules

Spectroscopy is the study of the interaction of the electromagnetic radiation with matter. Hence, it is mandatory to understand which transitions are allowed, which are prohibited and why.[126] From a semi-classical approach of the radiation-matter interaction, transitions are allowed if the corresponding transition moment is not null, where the latter is the integral involving the wave functions of the starting ($|i\rangle$) and arrival ($|k\rangle$) energy levels and the dipole moment operator ($\hat{\boldsymbol{\mu}}$), shortly represented in Dirac notation as $\langle k|\hat{\boldsymbol{\mu}}|i\rangle$.[13] For asymmetric rotors, this can be written as

$$\langle J\tau|\hat{\boldsymbol{\mu}}|J'\tau'\rangle\,, \tag{2.27}$$

where $|J'\tau'\rangle$ is the wave function of the starting energy level, $|J\tau\rangle$ is the one of the arrival energy level, and they depend only on the values of J and τ.[13] In order to treat the interaction with the electric field of the radiation, the projection of the dipole moment operator along a space-fixed axis must be considered.[13] By considering a radiation polarized along the Z axis, the $\hat{\mu}_Z$ operator should be considered, and Equation (2.27) can be written as

$$\langle J\tau|\hat{\mu}_Z|J'\tau'\rangle = \sum_g \mu_g \langle J\tau|\boldsymbol{\Phi}_{Zg}|J'\tau'\rangle\,, \tag{2.28}$$

where μ_g's are the components of the molecular dipole moment along the inertial axes, $\boldsymbol{\Phi}_{Zg}$ is the director cosine matrix, and the sum runs over the g inertial axes,[13] according to

$$\hat{\mu}_F = \sum_g \mu_g \boldsymbol{\Phi}_{Fg}\,, \tag{2.29}$$

where F is any space-fixed coordinate (X, Y, Z). From Equation (2.28), it is clear that at least one of the components of the molecular dipole moment must be not null. This is a general rule for rotational spectroscopy.[13] Furthermore, for each transition, the interation is due to only one of the dipole components, leading to a differentation of the transitions based on the component mediating the interaction: a-type transitions are due to

Table 2.4: Selection rules for asymmetric-top rotors.

Dipole component	ΔK_a	ΔK_c
$\mu_a \neq 0$	$0, \pm 2, \pm 4, \ldots$	$\pm 1, \pm 3, \ldots$
$\mu_b \neq 0$	$\pm 1, \pm 3, \ldots$	$\pm 1, \pm 3, \ldots$
$\mu_c \neq 0$	$\pm 1, \pm 3, \ldots$	$0, \pm 2, \pm 4, \ldots$

μ_a, b-type ones to μ_b, and c-type ones to μ_c.[13] From the director cosine matrix elements, the selection rules for the "good" quantum numbers can be derived. In the asymmetric-top rotors, they refer only to J and M, as stated above, thus leading to[13]

$$\begin{aligned} &\Delta J = 0, \pm 1\,, \\ &\Delta M = 0, \text{ if the radiation is polarized along } Z. \end{aligned} \tag{2.30}$$

To determine the selection rules in K_a, K_c, one can resort to symmetry arguments.[13] In the D_2 simmetry group, the direction cosines have B symmetries; in detail, the a component of $\mathbf{\Phi}_{Zg}$ (from Equation 2.28), ϕ_{Za}, transforms as B_a (eo), ϕ_{Zb} as B_b (oo), and ϕ_{Zc} as B_c (oe). Considering the interaction being mediated by the μ_a component of the molecular dipole moment, the $\mu_a \langle J\tau|\phi_{Za}|J'\tau'\rangle$ term of Equation (2.28) must be non-zero for the transition to occur. This is fulfilled if the integrand has A symmetry, and that is possible only if $|J\tau\rangle$ and $|J'\tau'\rangle$ have A and B_a or B_b and B_c symmetry; indeed,

$$A \otimes B_a \otimes B_a = B_b \otimes B_a \otimes B_c \in A\,. \tag{2.31}$$

In other words, allowed transitions are $ee \leftrightarrow eo$ or $oe \leftrightarrow oo$: for an a-type transition, the parity of K_a must be invariated while K_c's parity must change. Analogously, the K_a, K_c selection rules can be derived for b-type and c-type transitions, and they have been summarized in Table 2.4.

Line intensities

The intensity of rotational lines is proportional to the square of absolute values of the corresponding transition moment element (see Equation 2.27), and, for the $|J\tau\rangle \leftarrow |J'\tau'\rangle$ transition, it can be expressed as[13]

$$|\langle J\tau|\hat{\boldsymbol{\mu}}|J'\tau'\rangle|^2 = \frac{\mu_g^2 S_{ki}}{2J+1}\,, \tag{2.32}$$

where S_{ki} is an useful physical quantity known as line strength, and it is defined as

$$S_{ki} = \sum_{F} \sum_{M,M'} |\langle J\tau M|\Phi_{Fg}|J'\tau'M'\rangle|^2 , \tag{2.33}$$

where the sum extends over the three directions of the space-fixed system and over all values of M and M'. The latter sum considers all possible transitions which are degenerate in the absence of an external field.[13]

Molecular partition function

The molecular partition function is a statistical mechanics concept that describes the distribution of energy among the various quantum states of a system, enabling the calculation of various thermodynamic properties. For a molecule in the gas phase, it is expressed as a function of the electronic (el), vibrational (vib), rotational (rot), and nuclear spin (n) states. By neglecting the coupling between the corresponding degrees of freedom, the molecular partition function Q may be written as product of four terms[13]

$$Q = Q_{\mathrm{el}} Q_{\mathrm{vib}} Q_{\mathrm{rot}} Q_{\mathrm{n}} , \tag{2.34}$$

which represent respectively the electronic, vibrational, rotational, and nuclear spin partition functions. Given the context of this thesis, the expression is given only for the rotational term:[13]

$$\mathrm{Q_{rot}} \approx \frac{1}{\sigma}\left[\left(\frac{\pi}{B^2A}\right)\left(\frac{kT}{h}\right)^3\right]^{1/2} = \left(\frac{5.34\times10^6}{\sigma}\right)\left(\frac{T^3}{B^2A}\right)^{1/2} , \tag{2.35}$$

for a symmetic-top rotor, and

$$\mathrm{Q_{rot}} = \left(\frac{5.34\times10^6}{\sigma}\right)\left(\frac{T^3}{ABC}\right)^{1/2} \tag{2.36}$$

for an asymmetric-top rotor. In both Equations (2.35) and (2.36), σ is a measure of the degree of symmetry.[13] The reader is referred to Ref. 13 for details.

2.2 Vibrational effects

To describe the effect of vibration on the rotational motion, the starting point is the definition of a vibro-rotational Hamiltonian, in Watson's formulation[127]

$$\hat{\mathscr{H}}_{\mathrm{vibrot}} = \frac{1}{2}\sum_{g,h}(\hat{J}_g - \hat{\pi}_g)\mu_{gh}(\hat{J}_h - \hat{\pi}_h) + \frac{1}{2}\sum_{r}\omega_r\hat{p}_r^2 + V(\mathbf{q}) - \frac{1}{8}\sum_{g}\mu_{gg} , \tag{2.37}$$

where g, h refer to the principal inertia system axes a, b, c, q_r's are the dimensionless normal coordinates ($q_r = \omega_r^{1/2} Q_r$, where Q_r is the normal coordinate and ω_r is the harmonic wavenumber associated to the r-th vibrational mode), and $\hat{p}_r$'s are their conjugate momenta.[127] In Watson's vibro-rotational Hamiltonian the following terms also appear: i. the components of the internal vibrational angular momentum $\hat{\pi}_g$, $\hat{\pi}_h$; ii. the rotational angular momentum operator along the g-axis, $\hat{J}_g$; and iii. the inverse moment of inertia tensor μ_{gh}, which is a function of q.[127] The vibrational angular momentum $\hat{\pi}_g$ is defined as

$$\hat{\pi}_g = \sum_{r,s} \zeta_{rs}^g \left(\frac{\omega_s}{\omega_r}\right)^{1/2} q_r \hat{p}_s \,, \tag{2.38}$$

where ζ_{rs}^g's are the elements of the antisymmetric Coriolis zeta matrix ($\zeta_{rs}^g = -\zeta_{sr}^g$, and $\zeta_{rs}^g = 0$ if $r = s$), which couples q_r and q_s through the rotation around the g-axis.[13,127] The third term in the right-hand side of Equation (2.37) is the potential function in terms of the dimensionless normal coordinates, and the final term is the so-called Watson term (U), which is due to the use of a normal-coordinate representation and represents a small nearly-constant mass-dependent correction.[127]

The eigenvalues of the vibro-rotational Hamiltonian may be approached in a perturbation expansion, with the inverse inertia tensor and the potential being expanded as Taylor series in the normal coordinates q_r[127]

$$\mu_{gh} = \mu_{gh}^e + \sum_r \mu_{gh}^k q_r + \frac{1}{2} \sum_{r,s} \mu_{gh}^{kl} q_r q_s + \cdots \,, \tag{2.39}$$

and

$$V(\mathbf{q}) = \frac{1}{2} \sum_r \omega_r q_r^2 + \frac{1}{6} \sum_{r,s,t} \phi_{rst} q_r q_s q_t + \frac{1}{24} \sum_{r,s,t,u} \phi_{rstu} q_r q_s q_t q_u + \cdots \,, \tag{2.40}$$

where

$$\mu_{gh}^e = \frac{\delta_{gh}}{I_g^e} \,; \tag{2.41}$$

$$\mu_{gh}^r = \left(\frac{\partial \mu_{gh}}{\partial q_r}\right)_e = -\frac{a_r^{gh}}{I_g^e I_h^e} \,; \tag{2.42}$$

$$\mu_{gh}^{rs} = \left(\frac{\partial^2 \mu_{gh}}{\partial q_r \partial q_s}\right)_e = \frac{3}{4} \sum_\xi \frac{a_r^{g\xi} a_s^{\xi h} + a_s^{g\xi} a_r^{\xi h}}{I_g^e I_h^e I_\xi^e} \,, \tag{2.43}$$

with

$$a_r^{gh} = \left(\frac{\partial I_{gh}}{\partial q_r}\right)_e . \tag{2.44}$$

The leading terms of the expansions are the inverse principal moments of inertia at the equilibrium structure (Equation 2.41) and the harmonic potential (first term on the right-hand side of Equation 2.40).[127] The corrections to the harmonic potential are given by cubic, quartic, *etc.* terms, as shown in Equation (2.40). It can also be shown, from Equations (2.42) and (2.43), that first and second derivatives of the effective inverse inertia tensor elements can be written as a function of the first derivative of the inertia tensor with respect to the dimensionless normal coordinates (Equation 2.44).[127]

By incorporating Equations (2.39) and (2.40) in Equation (2.37), the terms of the vibro-rotational Hamiltonian can be grouped by power of vibrational (q, $\hat{p}$) and rotational ($\hat{J}$) operators[128]

$$\begin{aligned}\hat{\mathscr{H}}_{\text{vibrot}} = {} & \hat{\mathscr{H}}_{20} + \hat{\mathscr{H}}_{30} + \hat{\mathscr{H}}_{40} + \cdots \text{ (vibrational terms)} \\ & + \hat{\mathscr{H}}_{21} + \hat{\mathscr{H}}_{31} + \hat{\mathscr{H}}_{41} + \cdots \text{ (Coriolis terms)} \\ & + \hat{\mathscr{H}}_{02} + \hat{\mathscr{H}}_{12} + \hat{\mathscr{H}}_{22} + \cdots \text{ (rotational terms)} ,\end{aligned} \tag{2.45}$$

where, in $\hat{\mathscr{H}}_{mn}$, m represents the total power of q and $\hat{p}$ operators and n represents the total power of $\hat{J}$. Using Rayleigh–Schrödinger perturbation theory (RSPT), one can re-write Equation (2.45) as[127]

$$\hat{\mathscr{H}}_{\text{vibrot}} = \hat{\mathscr{H}}^0 + \hat{\mathscr{H}}' , \tag{2.46}$$

where

$$\begin{aligned}\hat{\mathscr{H}}^0 &= \hat{\mathscr{H}}_{20} + \hat{\mathscr{H}}_{02} \\ &= \sum_g B^g \hat{J}_g^2 + \frac{1}{2}\sum_r \omega_r (\hat{p}_r^2 + q_r^2) .\end{aligned} \tag{2.47}$$

This is the foundation of the vibrational perturbation theory to the second order (VPT2).[129] According to Equation (2.47), the unperturbed Hamiltonian $\hat{\mathscr{H}}^0$ consists of the rigid-rotor Hamiltonian and the harmonic oscillator Hamiltonian, while all the remaining terms of Equation (2.45) lead to the perturbation Hamiltonian $\hat{\mathscr{H}}'$.[127] Focusing on rotational constants, perturbation theory to the second order* leads to three different contributions to rotational constants.[127] The first derives from the quadratic term

*The first-order correction is defined by $E^{\text{I}} = \langle i|\hat{\mathscr{H}}'|i\rangle$, while the second-order one is $E^{\text{II}} = \sum_{k\neq i} \frac{\langle i|\hat{\mathscr{H}}'|k\rangle\langle k|\hat{\mathscr{H}}'|i\rangle}{E_i^0 - E_k^0}$, where $E_i^0 = \langle i|\hat{\mathscr{H}}^0|i\rangle$, $E_k^0 = \langle k|\hat{\mathscr{H}}^0|k\rangle$.

in the expansion of the effective inverse inertia tensor

$$B^g_{\text{eff}} \Leftarrow \frac{1}{2}\left\langle i \left| \frac{1}{2}\sum_{r,s} \mu^{rs}_{gg} q_r q_s \right| i \right\rangle , \tag{2.48}$$

while the second one is due to the Coriolis operator

$$B^g_{\text{eff}} \Leftarrow \frac{1}{I_g^2}\sum_{k\neq i} \frac{\langle i|\hat{\pi}_g|k\rangle \langle k|\hat{\pi}_g|i\rangle}{E^0_i - E^0_k} , \tag{2.49}$$

and the third one comes from the inverse mass tensor together with cubic anharmonicity

$$B^g_{\text{eff}} \Leftarrow \frac{1}{2}\sum_{k\neq i} \frac{\langle i|\sum_r \mu^r_{gg} q_r|k\rangle \langle k|\frac{1}{6}\sum_{r,s,t}\phi_{rst} q_r q_s q_t|i\rangle}{E^0_i - E^0_k} + \\ + \text{hermitian conjugate.} \tag{2.50}$$

If the vibrational state $|i\rangle$ is the vibrational ground state, Equations (2.48), (2.49), and (2.50) lead to the expression for the effective rotational constants in the vibrational ground state, *i.e.*, the so-called B^g_0:*

$$B^g_0 = B^g_e - \frac{d_r}{2}\sum_r \alpha^g_r ; \tag{2.51}$$

where B^g_e is the equilibrium rotational constant and the sum runs on all d_r-degenerate vibrational modes r.[127] The α^g_r's are the vibration-rotation interaction constants in the VPT2 framework, and they take the form[127]

$$\alpha^g_r = -2(B^g_e)^2\left[\sum_h \frac{3(a^{gh}_r)^2}{4I^h_e} + \sum_s \frac{(\zeta^g_{rs})^2(3\omega_r^2+\omega_s^2)}{\omega_r(\omega_r^2-\omega_s^2)} + \frac{1}{2}\sum_s \frac{\phi_{rrs}a^{gg}_s}{\omega_s^{3/2}}\right]. \tag{2.52}$$

2.3 Centrifugal distortion

Because of the non-rigidity of chemical bonds, the effect of centrifugal distortion needs to be taken into account in the treatment of rotational

*By generalizing Equation (2.51) for a vibrational state v, it is obtained $B^g_v = B^g_e + - \sum_r \alpha^g_r\left(v_r + \frac{d_r}{2}\right)$, where v refers to the vibrational state, and v_r (which takes zero and positive integral values) denotes the quantum number associated to the r-th vibrational mode.

motion. To do so, the rotational terms of the vibro-rotational Hamiltonian of Equation (2.45) should be considered, but a further simplification is required. The application of a contact transformation perturbation procedure, which will not be explained here (thus generally represented as a linear transformation),[127]

$$\tilde{\mathcal{H}}_{\text{vibrot}} = \mathbf{T}^{-1}\hat{\mathcal{H}}_{\text{vibrot}}\mathbf{T}\,, \tag{2.53}$$

leads to a block-diagonal effective Hamiltonian $\tilde{\mathcal{H}}_{\text{vibrot}} \equiv \sum_{m,n}\tilde{\mathcal{H}}_{mn}$, with the $m = 0$ terms defining the pure rotational and centrifugal distortion contributions. Thus, the rotational terms can be written as[127]

$$\begin{aligned}\tilde{\mathcal{H}}_{\text{rot}} &= \tilde{\mathcal{H}}_{02} + \tilde{\mathcal{H}}_{04} + \tilde{\mathcal{H}}_{06} + \cdots \\ &= \frac{1}{2}\sum_{g,h}\mu^{e}_{gh}\hat{J}_g\hat{J}_h + \frac{1}{4}\sum_{g,h,l,m}\tau_{ghlm}\hat{J}_g\hat{J}_h\hat{J}_l\hat{J}_m + \\ &+ \sum_{g,h,l,m,n,o}\tau_{ghlmno}\hat{J}_g\hat{J}_h\hat{J}_l\hat{J}_m\hat{J}_n\hat{J}_o + \cdots\,,\end{aligned} \tag{2.54}$$

where μ^{e}_{gh} and $\hat{J}_g$ have already been defined. All the sums run over the inertial axes. The first term in Equation (2.54) is the rigid-rotor Hamiltonian ($\hat{\mathcal{H}}_{02} = \tilde{\mathcal{H}}_{02}$) and the following terms represent the quartic and sextic centrifugal distortion Hamiltonians, τ_{ghlm} and τ_{ghlmno} being the quartic and sextic centrifugal distortion constants, respectively.[128]

However, to relate the experimentally determinable parameters with Equation (2.54), one must remove the indeterminacies caused by the use of the contact transformation.[128] From now on, only the quartic centrifugal distortion constants will be considered, but analogous results can be derived for the sextic ones. The former are defined as[127]

$$\tau_{ghlm} = -\frac{1}{2}\sum_{r}\frac{\partial\mu_{gh}}{\partial Q_r}(\omega_r)^{-1}\frac{\partial\mu_{lm}}{\partial Q_r}\,, \tag{2.55}$$

where, as before, ω_r is the r-th harmonic frequency, and Q_r is the corresponding normal coordinate. The first transformation leads to the so-called reduced Hamiltonian[128]

$$\tilde{\mathcal{H}}'_{04} = \frac{1}{4}\sum_{gh}\tau'_{gghh}\hat{J}^2_g\hat{J}^2_h\,, \tag{2.56}$$

resulting also in a small correction to the effective rotational constants. Another contact transformation is required, leading to $\tilde{\mathcal{H}}^{\text{red}}_{04}$. Different results are obtained depending on the reduction chosen. In the following,

the results for two widely used reductions are reported, also including the sextic terms[13,128]

$$\tilde{\mathscr{H}}^{A}_{\mathrm{rot}} = \sum_g B^A_g \hat{J}^2_g - \Delta_J (\hat{J}^2)^2 - \Delta_{JK} \hat{J}^2 \hat{J}^2_z - \Delta_K \hat{J}^4_z + \Phi_J (\hat{J}^2)^3 + \\ + \Phi_{JK} (\hat{J}^2)^2 \hat{J}^2_z + \Phi_{KJ} \hat{J}^2 \hat{J}^6_z - \frac{1}{2}[(\delta_J \hat{J}^2 + \delta_K \hat{J}^2_z), (\hat{J}^2_+ + \hat{J}^2_-)]_+ + \\ + \frac{1}{2}[(\phi_J (\hat{J}^2)^2 + \phi_{JK} \hat{J}^2 \hat{J}^2_z + \phi_K \hat{J}^4_z), (\hat{J}^2_+ + \hat{J}^2_-)] + \cdots , \quad (2.57)$$

$$\tilde{\mathscr{H}}^{S}_{\mathrm{rot}} = \sum_g B^S_g \hat{J}^2_g - D_J (\hat{J}^2)^2 - D_{JK} \hat{J}^2 \hat{J}^2_z - D_K \hat{J}^4_z + d_1 \hat{J}^2 (\hat{J}^2_+ + \hat{J}^2_-) + \\ + d_2 (\hat{J}^4_+ + \hat{J}^4_-) + H_J (\hat{J}^2)^3 + H_{JK} (\hat{J}^2)^2 \hat{J}^2_z + H_{KJ} \hat{J}^2 \hat{J}^6_z + h_1 (\hat{J}^2)^2 (\hat{J}^2_+ + \hat{J}^2_-) + \\ + h_2 \hat{J}^2 (\hat{J}^4_+ + \hat{J}^4_-) + h_3 (\hat{J}^6_+ + \hat{J}^6_-) + \cdots . \quad (2.58)$$

In both Equations (2.57) and (2.58), $\hat{J}_\pm$ represent the ladder operators* and $[\hat{o}, \hat{p}]_+$ is the anticommutator. Equation (2.57) defines the Watson *A*-reduction Hamiltonian, while the Equation (2.58) the Watson *S*-reduction Hamiltonian.[13] All the parameters appearing in Equations (2.57) and (2.58) are quantities experimentally determinable from the analysis of the rotational spectra. The reader is referred to Ref. 13 for higher-order terms.

2.4 Hyperfine structure

A further issue in rotational spectroscopy is the effect of non-zero nuclear spins because, when present in the molecule under consideration, they give rise to non-negligible interactions with molecular electric and/or magnetic fields.[13] The overall result of these interactions is the splitting of the rotational energy levels, and, consequently, the splitting of the rotational transitions, thus leading to the so-called hyperfine structure.[13] The main interaction, which is the only one addressed in this thesis, occurs between a nuclear quadrupole moment and the electric-field gradient at the corresponding nucleus. This results from a non-spherical distribution of the nuclear charge (which gives rise to the nuclear quadrupole) and a non-spherical distribution of the electronic charge about the nucleus (which gives rise to an electric field-gradient at the nucleus). In order to

*$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y$.

have quadrupolar nucleus, the nuclear spin quantum number I of the A nucleus must be greater or equal to 1 ($I_A \geq 1$). The nuclear quadrupole interaction exerts a twisting torque on the nucleus depending on the field gradient, which, in gases, depends on the rotational state of the molecule.[13]

The classical interaction energy of the nuclear charge with a potential V can be defined in terms of the nuclear charge density ρ_A in a space-fixed coordinate system

$$E = \int \rho_A V \mathrm{d}\tau_A \,, \tag{2.59}$$

with $\rho_A = \rho(X, Y, Z)$, $\mathrm{d}\tau_A = \mathrm{d}X\mathrm{d}Y\mathrm{d}Z$, and the integration is over the nuclear volume. After expanding the potential as Taylor's series and considering ρ_A constant throughout the volume, the quadrupole interaction energy is derived

$$\begin{aligned} E_Q = {} & \frac{1}{2}\left(\frac{\partial^2 V}{\partial X^2}\right)_0 \int \rho_A X_A^2 \mathrm{d}\tau_A + \frac{1}{2}\left(\frac{\partial^2 V}{\partial Y^2}\right)_0 \int \rho_A Y_A^2 \mathrm{d}\tau_A + \\ & + \frac{1}{2}\left(\frac{\partial^2 V}{\partial Z^2}\right)_0 \int \rho_A Z_A^2 \mathrm{d}\tau_A + \frac{1}{2}\left(\frac{\partial^2 V}{\partial X \partial Y}\right)_0 \int \rho_A X_A Y_A \mathrm{d}\tau_A + \\ & + \frac{1}{2}\left(\frac{\partial^2 V}{\partial X \partial Z}\right)_0 \int \rho_A X_A Z_A \mathrm{d}\tau_A + \frac{1}{2}\left(\frac{\partial^2 V}{\partial Y \partial Z}\right)_0 \int \rho_A Y_A Z_A \mathrm{d}\tau_A \,. \end{aligned} \tag{2.60}$$

It has to be pointed out that the first two terms arising from the Taylor's expansion, *viz.* the monopole and the (null) dipole interactions, are ignored in Equation (2.60). Further semplifications can be applied by chosing the x, y, z body-fixed system, with z along the spin axis **I**: this translates in x, y, z being the principal quadrupole axes, and, therefore, the vanishing of the cross terms.

Given the premises on the quadrupole interaction, for a rotating molecule in absence of an external field, the nuclear spin **I** couples with the molecular rotational angular momentum **J**, to form a resultant **F** according to the coupling scheme[13]

$$\mathbf{F} = \mathbf{J} + \mathbf{I} \,. \tag{2.61}$$

J and **I** can be considered as precessing about **F**, as shown in Figure 2.2. The "good" quantum numbers now are F, M_F, J, and I, with

$$\begin{aligned} F &= J + 1, J + I -, \ldots, |J - I| \,; \\ M_F &= F, F - 1, \ldots, F \,, \end{aligned} \tag{2.62}$$

and

$$\begin{aligned} \langle F M_F | \mathbf{F}^2 | F' M_F' \rangle &= \hbar^2 F(F+1) \,; \\ \langle F M_F | F_Z | F' M_F' \rangle &= \hbar M_F \,. \end{aligned} \tag{2.63}$$

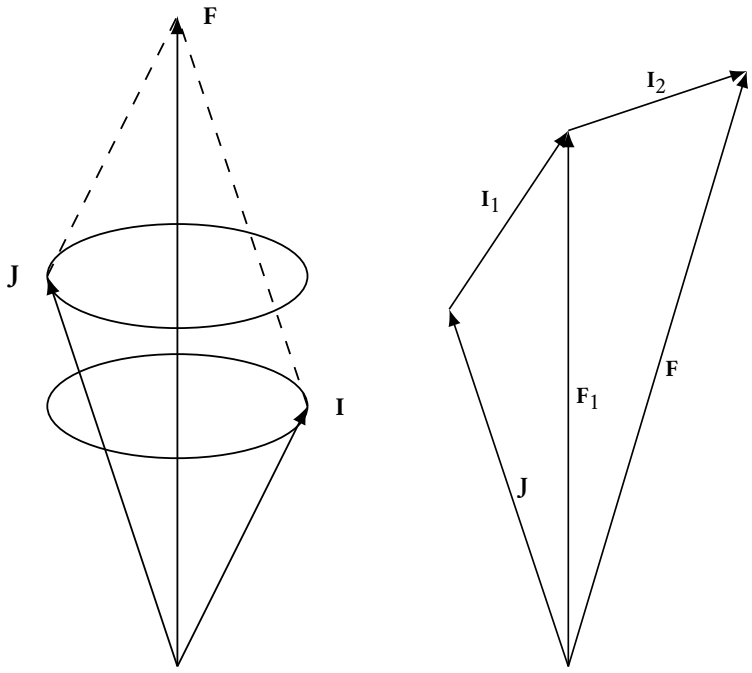


Figure 2.2: Coupling of the nuclear spin **I** with the molecular rotational angular momentum **J** (left panel); coupling of two nuclear spins $\mathbf{I}_1 > \mathbf{I}_2$ with the molecular rotational angular momentum **J** (right panel).

Moving to transitions, new selection rules involving F need to be defined, namely[13]

$$\Delta F = 0, \pm 1\,. \tag{2.64}$$

The derivation of the quantum mechanical operator for the nuclear quadrupole interaction energy is rather complicate; the reader is referred to Ref. 13 for details. Briefly, from Equation (2.60), the general expression of the Hamiltonian describing this type of interactions can be written as[13]

$$\hat{\mathscr{H}}_Q = \frac{eQq_J}{I(2I-1)2J(2J-1)} \times \left[3(\hat{\mathbf{I}}\cdot\hat{\mathbf{J}})^2 + \frac{3}{2}(\hat{\mathbf{I}}\cdot\hat{\mathbf{J}}) - \hat{\mathbf{I}}^2\hat{\mathbf{J}}^2\right], \tag{2.65}$$

where eQ is the nuclear quadrupole moment and q_J is the expectation value of the space-fixed component of the electric-field gradient tensor at the same nucleus.[13] The value of q_J depends on the type of rotor. By resorting to the coupling scheme of Equation (2.61), the eigenvalues of the general Hamiltonian of Equation (2.65) can be obtained

$$E_Q = \frac{eQq_J}{I(2I-1)2J(2J-1)}\left[\tfrac{3}{4}E(E+1) - J(J+1)I(I+1)\right], \tag{2.66}$$

with

$$E = F(F+1) - J(J+1) - I(I+1)\,. \tag{2.67}$$

The value of q_J for an asymmetric-top rotor is given by:

$$\begin{aligned} q_J &= \left\langle \left(\frac{\partial^2 V}{\partial Z^2}\right)_0 \right\rangle_{M_J=J} \\ &= \sum_g q_{gg} \left\langle J\tau M_J = J|\Phi_{Zg}^2|J\tau M_J = J\right\rangle\,. \end{aligned} \tag{2.68}$$

Except for low values of J, the matrix elements of the squared director cosines matrix can not be expressed in a closed form. Thus, they are expressed in terms of the reduced energies and of Ray's asymmetry parameter $\varkappa$. The reader is referred to Ref. 13 for details.

To conclude, it is important to introduce the physical quantities that are experimentally determinable. These are the nuclear quadrupole coupling constants, defined as

$$\chi_{gg} = eQq_{gg}\,, \tag{2.69}$$

where q_{gg} is the gg-element of the electric-field gradient tensor at the quadrupolar nucleus. Furthermore, the nuclear quadrupole coupling tensor is traceless.

Presence of two quadrupolar nuclei

The complication of the rotational hyperfine structure of a molecule increases if more than one nucleus presents a quadrupolar coupling. A perturbative treatment can be employed if one nuclear quadrupole coupling constant is significantly larger than the second ($\chi_1 \gtrapprox 10\chi_2$); instead, the composite Hamiltonian for nuclei with comparable quadrupole interactions is defined as sum of separate Hamiltonians of the form of Equation (2.65).[13] For the matrix element with two quadrupolar nuclei, the reader is referred to Ref. 13.

2.5 Inversion motions

Non-rigid molecules constitute an interesting class of molecules whose rotational and vibro-rotational spectra have been studied extensively by spectroscopists. They are characterized by a potential energy surface that includes energy barriers separating two or more identical equilibrium configurations. Within this class, the most investigated phenomenon is

the internal rotation, which involves the relative rotation of two groups connected by a single bond; *e.g.* the rotation of a methyl moiety. In the latter case, the potential energy takes the shape of a periodic sinusoidal function. However, none of the species investigated has an internal rotor; therefore, interested reader is referred to Refs. 13 and 130 for a detailed treatment. Within this thesis, the focus is on another important phenomenon: the inversion motion.

The simplest example consists in the inversion of a pyramidal symmetric-top XA_3 molecule (as ammonia), *i.e.*, when the X atom moves through the A_3 plane in order to reach an identical but inverted configuration. This motion is theoretically possible in any non-planar molecule, but it is significant for only a few of them. In these cases, the potential energy curve along the inversion coordinate takes the shape of a double-well (see Figure 2.3), with two identical minima separated by a barrier, denoted as inversion barrier.[5,13] The height of the barrier depends on the way the molecule has to contort to reach the identical conformation thorough the inversion motion.[130] From a classical point of view, the molecule could not undergo inversion if the vibrational state lies below the potential maximum. In quantum mechanics, this is made possible by the tunnel effect. Inversion motions are not an exclusive prerogative of symmetric XA_3 molecules; this phenomenon is noted also in more complex species which contain flexible –XA or $–XA_2$ moieties, as hydroxyl or amino groups.[131]

As a consequence, the vibrational energy levels of the electronic ground state are splitted: the $v_{\text{inv}} = 0^{\pm}, 1^{\pm}, \ldots$ nomenclature is used, where the "±" label denotes the parity of the overall wave function with respect to inversion, and v_{inv} is the vibrational quantum number of the inversion motion. To derive the selection rules for the allowed transitions between vibrational, rotational and inversional states, one can resort to symmetry considerations.[5,130] Being $\langle k|\hat{\boldsymbol{\mu}}|i\rangle$ a general dipole-induced transition moment (see Section 2.1), the selection rules differ depending on the behavior of the μ_g component ($g = a, b, c$) of the molecular electric dipole moment upon inversion. If the transition is mediated by a component that changes sign if the inversion takes place, the parity selection rules is[5,130]

$$+ \leftrightarrow - , \tag{2.70}$$

meaning that $\langle k|\hat{\boldsymbol{\mu}}|i\rangle$ is non-null only if $|k\rangle$ and $|i\rangle$ have different parity with respect to inversion. Instead, if the inversion does not affect the sign of the dipole component, the transitions take place within states of the

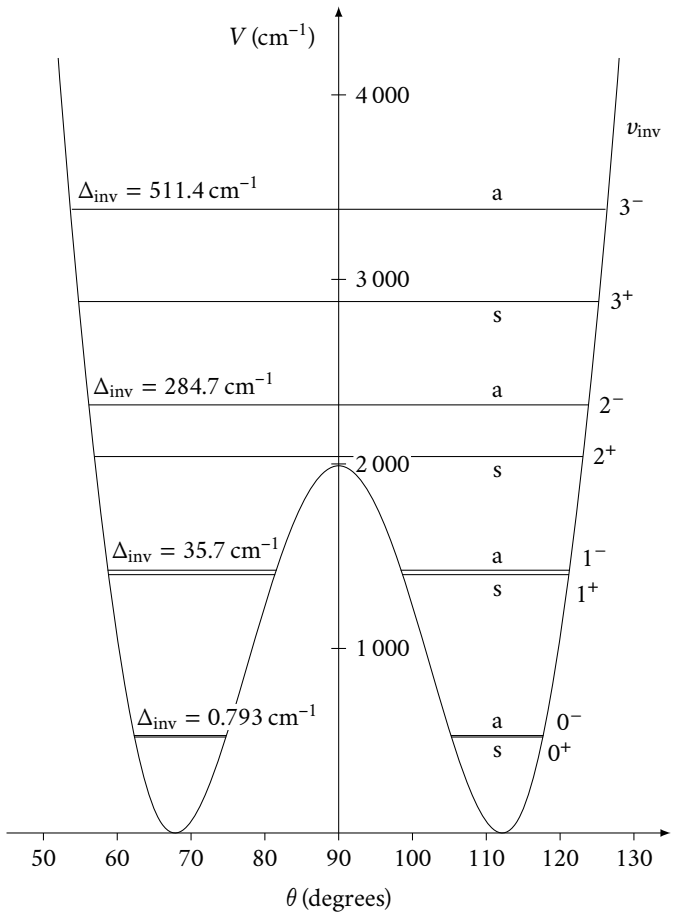


Figure 2.3: Double-well potential function of ammonia with inversion splittings (Δ_{inv}) and values for the vibrational quantum number of the inversion motion (v_{inv}). The angle θ represents the bending coordinate (namely the angle between the N–H bond and the C_3 axis), while V is the potential energy; "a" and "s" refer to the antisymmetric and symmetric wave functions, respectively. Reproduced from Ref. 130.

same symmetry:[5,130]

$$\begin{aligned} + &\leftrightarrow + ; \\ - &\leftrightarrow - . \end{aligned} \tag{2.71}$$

Chapter 3

Computational Spectroscopy

Theoretical studies in the field of molecular spectroscopy play a threefold role: interpretation, complementarity, and prediction. In order fully exploit them, the parameters introduced in Chapter 2 need to be computed. Before moving to the computational approach employed (Section 3.8), an overview of quantum chemistry is warranted: the reader will make acquaintance with the main *ab initio* and density functional theory methods, and with the concept of basis set.

The starting point is represented by the time-independent non-relativistic Schrödinger equation[132]

$$\hat{\mathscr{H}}_{\text{mol}}(\mathbf{r},\mathbf{R})\,|\Psi(\mathbf{r},\mathbf{R})\rangle = E_{\text{mol}}\,|\Psi(\mathbf{r},\mathbf{R})\rangle\ , \tag{3.1}$$

where $\hat{\mathscr{H}}_{\text{mol}}(\mathbf{r},\mathbf{R})$ is the molecular Hamiltonian, $\mathbf{r}$ and $\mathbf{R}$ are the position vectors of electrons and nuclei, respectively, $|\Psi(\mathbf{r},\mathbf{R})\rangle$ is the molecular wave function, and E_{mol} is the total energy of the system.[132] Equation (3.1) can be solved exactly only for hydrogen and hydrogenoid (*i.e.*, monoelectronic) atoms, thus requiring the introduction of approximations. To this purpose, the Born–Oppenheimer (BO) approximation is commonly used to separate the electronic and nuclear degrees of freedom.[132] It relies on the large mass difference between nuclei and electrons ($\frac{m_p}{m_e} \approx 5.4 \times 10^{-4}$, where m_p is the mass of the proton, and m_e is the electron mass), which causes a slower motion of the former with respect to the latter.[132] Thus, while the electrons are moving, the nuclei can be considered fixed. This approximation leads to the formulation of an electronic and a nuclear Schrödinger equations[132]

$$\hat{\mathscr{H}}_{\text{el}}(\mathbf{r},\mathbf{R})\,|\Psi_{\text{el}}(\mathbf{r},\mathbf{R})\rangle = E_{\text{el}}(\mathbf{R})\,|\Psi_{\text{el}}(\mathbf{r},\mathbf{R})\rangle\ ; \tag{3.2}$$

$$\left(\hat{T}_{\text{N}}(\mathbf{R}) + V(\mathbf{R})\right)|\Psi_{\text{N}}(\mathbf{R})\rangle = E_{\text{mol}}\,|\Psi_{\text{N}}(\mathbf{R})\rangle\ , \tag{3.3}$$

where

$$V(\mathbf{R}) = f(E_{\text{el}}(\mathbf{R}))\,. \tag{3.4}$$

In Equation (3.2), $\hat{\mathscr{H}}_{\text{el}}(\mathbf{r},\mathbf{R})$ is the electronic Hamiltonian and $\Psi_{\text{el}}(\mathbf{r},\mathbf{R})$ is the electronic wave function, and they both depend only parametrically on $\mathbf{R}$; in Equation (3.3), $\hat{T}_{\text{N}}$ is the nuclear kinetic energy operator and $\Psi_{\text{N}}(\mathbf{R})$ is the nuclear wave function.[132] The electronic Hamiltonian of Equation (3.2) is defined as follows

$$\hat{\mathscr{H}}_{\text{el}}(\mathbf{r},\mathbf{R}) = \hat{T}_{\text{el}}(\mathbf{r}) + \hat{V}_{\text{N,el}}(\mathbf{r},\mathbf{R}) + \hat{V}_{\text{el}}(\mathbf{r}) + \hat{V}_{\text{N}}(\mathbf{R})\,, \tag{3.5}$$

where: i. $\hat{T}_{\text{el}}(\mathbf{r})$ represents the electronic kinetic energy; ii. $\hat{V}_{\text{N,el}}(\mathbf{r},\mathbf{R})$ is the attractive electron-nuclei potential energy; iii. $\hat{V}_{\text{el}}(\mathbf{r})$ represents the repulsive potential energy between electrons; iv. $\hat{V}_{\text{N}}(\mathbf{R})$ is the electron-electron repulsive potential energy. Equation (3.4) defines the potential energy surface (PES), which is a hypersurface function of the $3N-6$ internal coordinates (for a non-linear molecule with N atoms), whose description is obtained by the interpolation of electronic energies, $E_{\text{el}}(\mathbf{R})$—which are the solutions of Equation (3.2) at different nuclear coordinates—calculated for a significant set of nuclear geometries.[132]

At this point, the aim is to solve the nuclear Schrödinger equation (Equation 3.3) where, given the context of this thesis, the Hamiltonian takes Watson's formulation from Equation (2.37). However, the derivation of the PES defined in Equation (3.4) must be addressed beforehand, thus requiring the solution of the electronic Schrödinger equation (Equation 3.2). To this purpose, different methods have been developed over the years:[132] their applications are referred to as electronic structure or quantum-chemical calculations, and they are presented in the following Sections. Before addressing them, the next Section introduces a suitable mathematical expression for the potential of the Hamiltonian of Equation (3.3).

3.1 Force field representation

If the considered system is small- to medium-sized, the PES may be constructed by interpolating a grid of points by means of variational procedures. This is not the case for the molecules investigated in this thesis, which range from medium- to large-sized, thus limiting the evaluation only to limited portions of the PES. The choice of the portions depends on the aim of the study; rotational spectroscopy requires the determination of the minimum energy structures together with the description of the surrounding PES (if more than one minimum is present, first-order

saddle points can also be of interest for an energetic characterization of interconversion barriers). To this purpose, a convenient model is provided by force field representation, which can be obtained by a Taylor expansion of the potential around a stationary point:[127]

$$V(\mathbf{q}) = V(q_e) + \sum_r \left(\frac{\partial E}{\partial q_r}\right)_e q_r + \frac{1}{2}\sum_{r,s}\left(\frac{\partial^2 E}{\partial q_r \partial q_s}\right)_e q_r q_s + \\ + \frac{1}{3!}\sum_{r,s,t}\left(\frac{\partial^3 E}{\partial q_r \partial q_s \partial q_t}\right)_e q_r q_s q_t + \\ + \frac{1}{4!}\sum_{r,s,t,u}\left(\frac{\partial^4 E}{\partial q_r \partial q_s \partial q_t \partial q_u}\right)_e q_r q_s q_t q_u + \cdots, \quad (3.6)$$

with q denoting the internal coordinates. The first derivatives of the energy (evaluated at the equilibrium) clearly vanish, while $V(q_e)$ can be arbitrarily set to zero. Therefore, Equation (3.6) can be re-written as[127]

$$V(\mathbf{q}) = \frac{1}{2}\sum_{r,s} f_{rs} q_r q_s + \frac{1}{3!}\sum_{r,s,t} f_{rst} q_r q_s q_t + \frac{1}{4!}\sum_{r,s,t,u} f_{rstu} q_r q_s q_t q_u + \cdots, \quad (3.7)$$

where f_{rs}, f_{rst} and f_{rstu} are the second, third and quartic derivatives of the energy with respect to the nuclear coordinates evaluated at the equilibrium structure; they are denoted as quadratic, cubic and quartic force constants, respectively.[127] The collection of the quadratic force constants defines the harmonic force field, that of the cubic force constants the cubic force field, and so on. A further simplification of Equation (3.7) can be obtained by resorting to normal coordinates $\mathbf{Q}$[127]

$$V(\mathbf{Q}) = \frac{1}{2}\sum_i \phi_i Q_i^2 + \frac{1}{3!}\sum_{i,j,k} \phi_{ijk} Q_i Q_j Q_k + \cdots, \quad (3.8)$$

where only the diagonal terms $\phi_{ii} \equiv \phi_i$ survive in the set of quadratic force constants, which are the square of the harmonic vibrational frequencies. This formulation of the potential enters the Hamiltonian of Equation (3.3) as shown in Equation 2.40.

From a computational point of view, ϕ_i's are obtained as analytical second derivatives of the electronic energies with respect to nuclear coordinates, with the normal coordinate representation being mostly used.[128] Instead, the computation of the higher-order force constants requires numerical-differentation techniques. In the framework of this thesis, only

the cubic and semi-diagonal quartic force constants are required[127]

$$\phi_{ijk} = \frac{\phi^{+}_{jk} - \phi^{-}_{jk}}{\Delta_i}\,; \tag{3.9}$$

$$\phi_{iijk} = \frac{\phi^{+}_{jk} - \phi^{-}_{jk} - 2\phi_{jk}}{\Delta_i^2}\,, \tag{3.10}$$

where $\phi^{\pm}_{jk}$ refers to the jk-th quadratic force constant evaluated at the $\mathbf{Q} \pm \Delta_i$ geometry, with Δ_i being the displacement along the i-th coordinate.[127]

3.2 Independent-particle approximation

As mentioned above, several quantum-chemical methods have been proposed to solve the electronic Schrödinger equation (Equation 3.2). The cornerstone is represented by the Harthee–Fock (HF) theory, which introduces a significant simplification, *i.e.*, the independent-particle approximation. This model considers each electron independent one from the other by approximating the electron-electron interaction in Equation (3.5) as a mean field: this allows to write the electronic Hamiltonian as a sum of mono-electronic Hamiltonians.[132] Within the HF theory, each electron is described by an orbital, and the total wave function is the antisimmetrized product of these orbitals, which takes the form of a single Slater determinant (SD)[133]

$$\begin{aligned}\Phi &= \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(1) & \psi_2(1) & \cdots & \psi_N(1) \\ \psi_1(2) & \psi_2(2) & \cdots & \psi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(N) & \psi_2(N) & \cdots & \psi_N(N) \end{vmatrix} \\ &= A\psi_1\psi_2\ldots\psi_N\,,\end{aligned} \tag{3.11}$$

where $\psi_i(\mu)$ is a spinorbital, *i.e.*, the combination, for the μ-th electron, of a function of the spatial orbital and the spin function, and A is the antisymmetrizer. In the HF (or self-consistent field, SCF) model, the spinorbitals are varied to minimize the energy expetation value of the single-determinant wave function, according to the variational principle.[133] The minimization is achieved by solving a set of coupled one-electron eigenvalue equations

$$\hat{F}_i\psi_i = \epsilon_i\psi_i\,, \tag{3.12}$$

where the mono-electronic Fock operator $\hat{F}_i$ is defined as

$$\hat{F}_i = \hat{h}_i + \sum_j^N (\hat{J}_{ij} - \hat{K}_{ij}) \,. \tag{3.13}$$

In Equation (3.13), $\hat{h}_i$ is the mono-electronic operator describing the kinetic energy of one electron and its attractive interaction with all the nuclei, while $\hat{J}_j, \hat{K}_j$ account for the mean-field description of the two-electron interactions.[127] In more detail, the Coulomb operator $\hat{J}_j$ is defined as

$$\hat{J}_j \left|\psi_i(2)\right\rangle = \left\langle \psi_i(1) \middle| \hat{r}_{12} \middle| \psi_j(1) \right\rangle \left|\psi_j(2)\right\rangle \,, \tag{3.14}$$

where $\hat{r}_{ij}$ is a two-electron operator giving the electron–electron repulsion. Equation (3.14) represents the repulsion experienced by the electron 2 in orbital ψ_i from electron 1 in orbital ψ_j.[127,132] Instead, the exchange operator $\hat{K}_j$ accounts for the modification of the electron–electron repulsion due to spin correlation, and it takes the form[127,132]

$$\hat{K}_j \left|\psi_i(2)\right\rangle = \left\langle \psi_j(1) \middle| \hat{r}_{12} \middle| \psi_j(1) \right\rangle \left|\psi_i(2)\right\rangle \,. \tag{3.15}$$

Note that the Fock operator from Equation (3.13) depends on all the spinorbitals. Therefore, iterative procedures are required to acquire consistency between the spinorbitals in the definition of $\hat{F}_i$ and the spinorbitals obtained as its eigenfunctions.[133]

Level of theory

Nowadays, the HF method is mostly used to derive the reference wave function to be used in *post*-HF methods. Indeed, the HF function represents an adequate zeroth-order wave function in the absence of near degenerancies between the contribution of several SDs to the wave function. These cases, also named multi-determinantal (or multi-configurational) cases, will not be addressed in this thesis.

In *post*-HF methods, the solution of Equation (3.2) is improved by incorporating in the treatment the so-called electron correlation, E_c, defined as[132]

$$E_c = E - E^{HF} \,, \tag{3.16}$$

where E is the total energy of the system and E^{HF} is the HF energy extrapolated to the complete basis set (CBS) limit. Usually, HF accounts for 90–99% of the total energy, but the remaining 10–1% is fundamental for chemical purposes.[132] Since the description of electron correlation can be

done only in an approximated way, several computational methods have been developed to solve the electronic Schrödinger equation, with different accuracy. Another important point is the description of the atomic orbitals (AOs) to be used to build the molecular ones. For this purpose, sets of basis functions are used.[132] The combination of the method chosen for describing the electronic wave function and the basis set employed defines the level of theory.[132]

In this thesis, different levels of theory will be employed based on accuracy and computational cost. In the following Sections, an overview of the basis sets and the theoretical methods used is presented. To provide the reader with an unit of measurement of the computational cost, the formal scaling of the main methods used in this thesis is shown in Table 3.1, where M refers to the number of basis functions.[132]

Table 3.1: Formal scaling of the rate limiting step of main *ab initio* quantum chemical methods of use in this thesis.

Scaling	Methods
M^4	HF
M^5	MP2
M^6	CCSD
M^7	CCSD(T)
M^8	CCSDT
M^{10}	CCSDTQ

3.3 Basis sets

One of the inherent approximations of computational methods is the use of basis sets—an array of functions used to expand an unknown function (as the molecular orbital, MO) in order to describe it. If the basis set is complete, there is no approximation; however, this means that an infinite number of functions is being used, which is impossible in actual calculations.[132] On the other hand, the smaller basis set, the poorer the representation; it is therefore necessary to find a suitable compromise between accuracy and computational requirements.

The basis sets used in this thesis to build the AOs are Gaussian-type orbitals (GTOs), which, in terms of polar (r, θ, φ) or Cartesian (x, y, z)

coordinates, are defined as

$$\begin{aligned} \chi_{\alpha nlm}(r,\theta,\varphi) &= N Y_{l,m}(\theta,\varphi) r^{2n-2-l} \mathrm{e}^{-\alpha r^2}\,; \\ \chi_{\alpha nlm}(x,y,z) &= N x^{l_x} y^{l_y} z^{l_z} \mathrm{e}^{-\alpha r^2}\,, \end{aligned} \tag{3.17}$$

where: i. α is an exponent controlling the width of the GTO; ii. n is the principal quantum number; iii. l, m are the angular momentum and its component along z; iv. N is a normalization constant; and v. $Y_{l,m}(\theta,\varphi)$ are the spherical harmonic functions.[132,134] The sum of l_x, l_y, and l_z determines the type of orbital (*e.g.*, $l_x + l_y + l_z = 1$ describes a p orbital). In order to reduce the computational effort, modern programs geared on Cartesian coordinates exploit spherical functions when dealing with high angular momentum orbitals: in fact, considering for instance a generic d orbital, the Cartesian functions have six different components ($xy, xz, yz, x^2, y^2, z^2$), while the spherical ones only five ($Y_{2,-2}, Y_{2,-2}, Y_{2,0}, Y_{2,1}, Y_{2,2}$).[132]

However, GTOs are not used directly; to better mimic the behavior of hydrogenic AOs, linear combination of GTOs are employed:

$$\xi_{\alpha nlm} = \sum_{a}^{N} c_a \chi_{\alpha nlm}\,, \tag{3.18}$$

where c_a is the contraction coefficient.[132,134] The term "contraction" is referred to the use of linear combinations; thus, the new functions take the name of contracted GTOs (cGTOs), while the χ's are now referred to as primitives. There are two different ways of contracting a set of primitives, *viz.* general and segmented contraction forms. The latter partitions the set of primitives into smaller sets, each of them generating different cGTOs; the former, instead, uses all primitives on a given atom to generate different cGTOs, but with differents contraction coefficients. This is the method used to build the following basis sets.[132]

The basis sets employed in this thesis belong to the series developed by Dunning and co-workers[135–137] and their derivatives.[138–140] They are generally denoted by the following acronym

`(aug-)cc-p(C)VnZ,`

where the parentheses mean that the included term might not be present. The `VnZ` part of the acronym stands for "valence n-uple ζ", where n denotes the number of cGTOs used to describe one valence orbital, while core orbitals are described by only one cGTO.[132] This is however not the case if `C` is present, this meaning that more than one function are also used

for core orbitals. The cc part stands for "correlation-consistent", which means that the basis set has been optimized for being used with correlated methods.[132] Polarization functions, which are higher angular momentum functions with respect to the valence space (*e.g.*, p functions on the H atoms) are crucial and are represented by the p letter.[132] Instead, the aug- part of the acronym refers to the use of additional diffuse functions, characterized by a small α exponent (*cf.* Equation 3.17).[132]

The derivatives of Dunning's original set employed in this thesis are the minimally augmented and calendar basis sets. The former (maug-cc-pVTZ) is generated from the aug-cc-pVTZ one by truncating the diffuse space to the minimal augmentation level, while preserving the same exponential parameters of the augmented case. For example, in this basis set, the carbon atom possesses only s and p functions, but retains the aug- coefficients on them.[138,141] Starting from the minimally augmented basis set, the maug-cc-pVTZ-dH one is obtained by removal of the d functions on the H atoms (-dH).

The calendar basis set family is defined as

mon-cc-pVTZ,

where mon- refers to the "month", *i.e.*, jul-, jun-, may-, and apr-. They constitute a convergent sequence of partially augmented basis sets obtained by successively deleting diffuse subshells on heavier atoms from the aug- set.[140] As an example, the carbon atom in the the aug- and jul- triple-ζ basis sets has diffuse s, p, d, and f; by removal of the f subshell, the jun- set is obtained; then, by removal of the d subshell, the may- basis is obtained, and so on. For a triple-ζ basis set, maug- and may- are equivalent.[139]

Finally, a small basis set employed in this thesis is the so-called SNSD,* a double-ζ basis set obtained from the N07D basis set[142–144] by inclusion of diffuse s functions on all atoms and one set of diffuse polarized functions (d functions on heavy atoms and p on H atoms).

3.4 Second order Møller-Plesset theory

Rayleigh–Schrödinger perturbation theory (RSPT) applied to the solution of the electronic Schrödinger equation leads to Møller-Plesset (MP) theory to the n-th order (MPn).[132] The most known and used method of this series is that with $n = 2$, defined as Møller-Plesset theory to the second order (MP2). In this formulation, the perturbation theory is used to partition the

*Available for download from: skies-village.it/smartdownload/download.php.

exact Hamiltonian in an unperturbed Hamiltonian, defined as the sum of all mono-electronic Fock operators[132]

$$\hat{\mathscr{H}}^0 = \sum_{i=1}^{N} \hat{F}_i \,, \tag{3.19}$$

and a perturbation Hamiltonian, defined as[132]

$$\hat{\mathscr{H}}' = \hat{\mathscr{H}} - \hat{\mathscr{H}}^0 \,. \tag{3.20}$$

By applying the RSPT to the second order, it is obtained[132]

$$E^0 = \langle \Psi^0 | \hat{\mathscr{H}}^0 | \Psi^0 \rangle = \sum_{i=1}^{N} \epsilon_i \,; \tag{3.21}$$

$$E^{\mathrm{I}} = \langle \Psi^0 | \hat{\mathscr{H}}' | \Psi^0 \rangle = E^{\mathrm{SCF}} - E^0 \,; \tag{3.22}$$

$$E^{\mathrm{II}} = \langle \Psi^0 | \hat{\mathscr{H}}' | \Psi^1 \rangle = -\sum_{i,j}^{\mathrm{occ}} \sum_{a,b}^{\mathrm{vir}} \frac{\langle \Psi_{ij}^{ab} | \hat{\mathscr{H}}' | \Psi^0 \rangle \langle \Psi^0 | \hat{\mathscr{H}}' | \Psi_{ij}^{ab} \rangle}{\langle \Psi_{ij}^{ab} | \hat{\mathscr{H}}' | \Psi_{ij}^{ab} \rangle - E^0} \,, \tag{3.23}$$

where $\Psi^0 \equiv \Psi^{\mathrm{HF}}$, which is the HF wave function, and ϵ_i are the orbitalic energies. The HF-SCF energy is obtained at the first order.[132] Being $\hat{\mathscr{H}}'$ a two-electron operator, only doubly excited determinants contribute to Equation (3.23). The notation $(ai|bj)$, $(aj|bi)$ can be used for defining Coulomb and exchange integrals, respectively (*cf.* Equations 3.14 and 3.15), and the MP2 correction to HF energy becomes[132]

$$E^{\mathrm{MP2}} = -\frac{1}{4} \sum_{i,j}^{\mathrm{occ}} \sum_{a,b}^{\mathrm{vir}} \frac{[(ai|bj) - (aj|bi)]^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \,, \tag{3.24}$$

where the denominator represents the difference in the energies of the involved MOs, as follows from Koopmans' theorem,* and the factor ¼ is due to the equivalence $\Psi_{ij}^{ab} = -\Psi_{ji}^{ab} = -\Psi_{ij}^{ba} = \Psi_{ji}^{ba}$.[132]

3.5 Coupled cluster theory

In the coupled cluster (CC) formulation,[133] the electronic wave function is given by

$$\Psi = \exp(\hat{T})\Psi^0 \,, \tag{3.25}$$

*The ionization energy within the frozen MO approximation is given simply as the orbital energy.[132]

where $\Psi^0 \equiv \Psi^{\mathrm{HF}}$ and $\hat{T}$ is the cluster operator in the general formulation[133]

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots . \tag{3.26}$$

The generic $\hat{T}_m$ operator describes the m-th excitation and can be written by using second quantization. In the following expression, $i, j, k, \ldots$ are employed for labeling the occupied orbitals and $a, b, c, \ldots$ for the virtual ones as[133]

$$\hat{T}_m = \frac{1}{(m!)^2} \sum_{i,j,k,\ldots}^{\mathrm{occ}} \sum_{a,b,c,\ldots}^{\mathrm{vir}} t_{ijk\ldots}^{abc\ldots} \hat{a}^\dagger \hat{i} \hat{b}^\dagger \hat{j} \ldots , \tag{3.27}$$

where $\hat{o}^\dagger$ and $\hat{p}$ represent creation and annihilation operators, respectively, producing an m-fold excitation by creating an electron in the virtual o orbital and destroying an electron in the occupied orbital p.[133] The $t_{ijk\ldots}^{abc\ldots}$ coefficients are the cluster amplitudes. The $(m!)^{-2}$ factor accounts for the redundancy created by the unrestricted summations. For example, for the double excitation operator[133]

$$t_{ij}^{ab} = -t_{ji}^{ab} = -t_{ij}^{ba} = t_{ji}^{ba} . \tag{3.28}$$

Once the cluster operator and the CC wave function are defined, the Schrödinger equation can be written as[133]

$$\hat{\mathscr{H}} \exp(\hat{T}) \Psi^0 = E \exp(\hat{T}) \Psi^0 , \tag{3.29}$$

from which the energy and amplitudes can be derived by solving the following equations, respectively[133]

$$\langle \Psi^0 | \exp(-\hat{T}) \hat{\mathscr{H}} \exp(\hat{T}) | \Psi^0 \rangle = E ; \tag{3.30}$$

$$\langle \Psi^m | \exp(-\hat{T}) \hat{\mathscr{H}} \exp(\hat{T}) | \Psi^0 \rangle = 0 , \tag{3.31}$$

where Ψ^m represents a generic m-fold excitated configuration. The solution of Equations (3.30) and (3.31) requires the truncation of the cluster operator of Equation (3.26) to a given excitation order, thus leading to different methods.[133] The CCD method includes only double excitations, while CCSD considers singles and doubles excitations. However, the main aim of the CC theory is providing accurate results, thus requiring the inclusion of triple excitations at an affordable computational cost.[127] This requirement led to the formulation of the CCSD(T) ansatz, which includes triple excitations by means of a perturbative treatment.[145,146] This method is defined as the "gold standard" for accurate computations, thanks to an overestimation of the triples contribution which often leads to results that are qualitatively and quantitatively in between CCSDT (full treatment of triples) and CCSDTQ (full quadruples).[127]

3.6 Density functional theory

Density functional* theory (DFT) originated from the need of finding a different way to describe the electronic structure of a system, in order to overcome the shortcomings of *ab initio* methods.[132] Its basis is provided by the Hohenberg–Kohn theorem, which states that the electronic ground state energy depends on the electronic density ρ.[132] Furthermore, the total number N of electrons is related to ρ and a set of spatial coordinates $\mathbf{r}$ by

$$N = \int \rho(\mathbf{r}) \mathrm{d}\mathbf{r}\,, \tag{3.32}$$

and nuclei correspond to electron density maxima.[132]

The electronic energy can be written by using the Thomas–Fermi formulation

$$E[\rho(\mathbf{r})] = T_{\mathrm{ni}}[\rho(\mathbf{r})] + V_{\mathrm{N,el}}[\rho(\mathbf{r})] + V_{\mathrm{el}}[\rho(\mathbf{r})] + E_{\mathrm{xc}}[\rho(\mathbf{r})]\,, \tag{3.33}$$

which consists of the following terms: i. the kinetic energy of non-interacting electrons, $T_{\mathrm{ni}}[\rho(\mathbf{r})]$; ii. the attractive interaction energy between nuclei and electrons, $V_{\mathrm{N,el}}[\rho(\mathbf{r})]$; iii. the classical repulsion energy between electrons, $V_{\mathrm{el}}[\rho(\mathbf{r})]$; and iv. the exchange-correlation energy, $E_{\mathrm{xc}}[\rho(\mathbf{r})]$, which includes correlation terms for the electron-electron interaction and the non-classical repulsion energy between electrons.[132] From this formulation, the quantum Kohn–Sham (KS) mono-electronic operator can be derived

$$\hat{h}_i^{\mathrm{KS}} = -\frac{1}{2}\nabla_i^2 - \sum_K \frac{Z_K}{|r_i - r_K|} + \int \frac{\rho(\mathbf{r}')}{|r_i - r_i'|}\mathrm{d}\mathbf{r}' + \frac{\partial E_{\mathrm{xc}}[\rho(\mathbf{r})]}{\partial \rho(\mathbf{r})}\,, \tag{3.34}$$

where

$$E_{\mathrm{xc}}[\rho(\mathbf{r})] = \int \rho(\mathbf{r}) \epsilon_{\mathrm{xc}}[\rho(\mathbf{r})] \mathrm{d}\mathbf{r}\,, \tag{3.35}$$

and it is thus written as interaction between electron density and the exchange-correlation energy density functional $\epsilon_{\mathrm{xc}}[\rho(\mathbf{r})]$.[132] As stated in the Hohenberg–Kohn theorem, there is a one-to-one correspondence between the electron density and the nuclear potential, and thereby also between the Hamiltonian operator and the energy; in other words, the energy is a unique functional of the electron density. According to that, an exact solution of the problem should exist; however, there is no information about the mathematical expression.[132]

*A functional, $F[f(x)]$, is the prescription for producing a numerical value from a function, $f(x)$, depending on variables (x).[132]

Therefore, approximations are needed to solve the eigenvalue equation corresponding to the KS operator. A widespread one is the local density approximation* (LDA), which affirms that ϵ_{xc} only depends on the local value of $\rho(\mathbf{r})$.[132] LDA can be improved by including the dependence on the density gradient, thus leading to the generalized gradient approximation (GGA)[132]

$$\epsilon_{xc}^{GGA} = \epsilon_{xc}^{LDA} + \Delta\epsilon_{xc}[\nabla\rho(\mathbf{r}), \rho(\mathbf{r})] \,. \tag{3.36}$$

DFT functionals are often denoted by the acronym deriving by the combination of the exchange and correlation functionals used in the formulation. A largely used exchange functional is Becke's (B),[147] in conjunction with the Lee–Yang–Parr (LYP) correlation functional,[148,149] whose combination leads to the BLYP functional. Another example of widespread exchange-correlation functional is the Perdew–Burke–Ernzerhof one (PBE),[150,151] often used together with the local correlation functional P86 by Perdew.[152]

Hybrid DFT methods

To further improve the DFT functionals, a fraction of the HF-SCF exchange energy can be included, thus leading to hybrid functionals. One of the mostly used is the B3LYP functional

$$E_{xc}^{B3LYP} = (1-a)E_x^{LDA} + aE_x^{HF} + b\Delta E_x^{B} + (1-c)E_c^{LDA} + cE_c^{LYP} \,, \tag{3.37}$$

where the "3" in the acronym refers to the number of empirical parameters a, b, c.[153] A further improvement can be obtained by the inclusion of a fraction of correlation energy at the MP2 level of theory. This leads to the definition the B2PLYP functional, a double-hybrid functional whose exchange-correlation energy is expressed as[154]

$$E_{xc}^{B2PLYP} = (1-a)E_x^{GGA} + aE_x^{HF} + bE_c^{GGA} + cE_c^{MP2} \,. \tag{3.38}$$

The introduction of spin-component scaling in the MP2-like term (which uses different empirical coefficients for the same- and opposite-spin perturbational terms, c_s and c_o, respectively), toghether with London-type empirical corrections (which consists of a molecular-mechanical term for long-range interactions of the van der Waals type and hydrogen-bond type) in double-hybrid functionals led to the definition of DSD methods,[155,156] whose exchange-correlation energy takes the form:

$$E_{xc}^{DSD} = (1-a)E_x^{GGA} + aE_x^{HF} + bE_c^{GGA} + c_o E_o^{MP2} + c_s E_s^{MP2} + s_6 E_D \,, \tag{3.39}$$

*For closed shell systems, it is equivalent to the local spin-density approximation (LSDA), which accounts for spin density differences in the LDA.[132]

where s_6 is also an empirical coefficient.[155,156] In this thesis, DSD methods are used together with the Perdew–Burke–Ernzerhof and P86 functionals (DSD-PBEP86)[155] exploiting Yang's revision[157] (revDSD-PBEP86).[158]

Finally, DFT functionals are usually employed in conjunction with empirical corrections for describing dispersion forces. For this purpose, the D3 scheme, proposed by Grimme,[159] in combination with the Becke–Johnson (BJ) damping function[160] has been chosen, thus leading to the B3LYP-D3(BJ), B2PLYP-D3(BJ), and revDSD-PBEP86-D3(BJ) methods. The scaling of the hybrid and double-hybrid functionals can be approximated to the scaling of the *ab initio* part in their definition (see Table 3.1).

3.7 Composite schemes

The approximations required to solve the electronic Schrödinger equation, such as the truncation of the wave function and of the basis set, lead to errors in quantum chemical computations. To recover them (and thus to improve the accuracy) without steeply raising the computational cost, composite schemes have been defined.[128] In these approaches, the additivity approximation is exploited, which states that all contributions that are considered important can be evaluated at the best possible level each according to the molecular size and combined together.

Concerning the energetic characterization, the first composite scheme here considered is the CCSD(T)/CBS+CV+fT+fQ model, introduced by Gauss and co-workers,[161,162] which defines the energy as

$$E_{\text{full}} = E_{\text{HF}}^{\text{CBS}} + \Delta E_{\text{c}}^{\text{CBS}} + \Delta E_{\text{CV}} + \Delta E_{\text{fT}} + \Delta E_{\text{fQ}}\,, \tag{3.40}$$

where, on the right-hand side:

i. the first term accounts for the extrapolation to the CBS limit of the HF/cc-pVnZ electronic energies, using the three-point formula by Feller[163]

$$E_{\text{HF}}^{\text{CBS}} = \frac{E(n) \times E(n-2) - [E(n-1)]^2}{E(n) - 2E(n-1) + E(n-2)}\,, \tag{3.41}$$

with n = T, Q, 5;

ii. the second term evaluates the CCSD(T)/pVnZ correlation energy, within the frozen-core (fc) approximation, extrapolated to the CBS limit, using the two-point n^{-3} formula by Helgaker and co-workers[164]

$$\Delta E_{\text{c}}^{\text{CBS}} = \frac{n^3 E(n) - (n-1)^3 E(n-1)}{n^3 - (n-1)^3} - E(n-1)\,, \tag{3.42}$$

with n = T, Q;

iii. the third term is the core–valence (CV) correlation correction, and is calculated as energy difference

$$\Delta E_{\mathrm{CV}} = E^{\mathrm{CVTZ}}_{\text{all-CCSD(T)}} - E^{\mathrm{CVTZ}}_{\text{fc-CCSD(T)}}\,, \tag{3.43}$$

where "all-" stands for the correlation of all electrons;*

iv. the fourth term is the contribution of the full treatment of the triple excitations, which is determined as energy difference

$$\Delta E_{\mathrm{fT}} = E^{\mathrm{VTZ}}_{\mathrm{CCSDT}} - E^{\mathrm{VTZ}}_{\mathrm{CCSD(T)}}\,; \tag{3.44}$$

v. the last term is the contribution of the inclusion of quadruples

$$\Delta E_{\mathrm{fQ}} = E^{\mathrm{VDZ}}_{\mathrm{CCSDTQ}} - E^{\mathrm{VDZ}}_{\mathrm{CCSDT}}\,. \tag{3.45}$$

From this scheme, the CCSD(T)/CBS+CV approach can be obtained by neglecting the last two terms terms on the right-hand side of Equation (3.40), thus sensibly reducing the computational demand.

In order to accurately describe larger systems, the so-called "cheap" composite scheme (ChS) has been developed,[165] whose denomination is due to its reduced computational cost with respect to a full CC approach. The ChS is based on the fc-CCSD(T)/cc-pVTZ level as starting point and incorporates the corrections for the extrapolation to the CBS limit and for the core-correlation effects at the MP2 level.[165] The expression for the ChS energy is thus given by[165]

$$E_{\mathrm{ChS}} = E^{\mathrm{VTZ}}_{\mathrm{CCSD(T)}} + \Delta E^{\mathrm{CBS}}_{\mathrm{MP2}} + \Delta E^{\mathrm{CV}}_{\mathrm{MP2}}\,, \tag{3.46}$$

where:

i. the first term is the single-point energy evaluated at the fc-CCSD(T)/cc-pVTZ level of theory;

ii. the second term considers the MP2/cc-pVnZ energies extrapolated to the CBS limit by using Equation (3.42) with n = T, Q (without discriminating the HF-SCF and correlation contributions);

*In the frozen-core approximantion, the number of determinants is limited to only those that can be generated by exciting the valence electrons.[132]

iii. the last term is the core–valence correlation correction, and is calculated as in Equation (3.43), but using the MP2/cc-pCVTZ level of theory.

In some cases, Equation (3.46) incorporates additional term, ΔE_{aug}, which accounts for the inclusion of diffuse functions in the basis set. It is calcolated as

$$\Delta E_{\text{aug}} = E_{\text{MP2}}^{\text{augVTZ}} - E_{\text{MP2}}^{\text{VTZ}}\,, \tag{3.47}$$

but it lacks a solid theoretical fundament. Indeed, its inclusion is controversial: it leads to significant improvements in geometry optimizations (see later on), where it tends to compensate the over-estimation of the MP2 extrapolation to the CBS limit; instead, in energetic determinations, it can provide erratic results.[166] To overcome this shortcoming, a variant of the ChS has been proposed by Alessandrini *et al.*,[166] namely the jun-"cheap" composite scheme (jun-ChS). Its formulation is analogous to Equation (3.46), where the cc-pVnZ basis sets are replaced by the partially augmented jun-cc-pVnZ basis sets.[166]

Moving to structural determinations, Equation (3.40) can be exploited in energy gradients to be used in geometry optimizations

$$\frac{\mathrm{d}E_{\text{full}}}{\mathrm{d}x} = \frac{\mathrm{d}E_{\text{HF}}^{\text{CBS}}}{\mathrm{d}x} + \frac{\mathrm{d}E_{\text{c}}^{\text{CBS}}}{\mathrm{d}x} + \frac{\mathrm{d}E_{\text{CV}}}{\mathrm{d}x} + \frac{\mathrm{d}E_{\text{fT}}}{\mathrm{d}x} + \frac{\mathrm{d}E_{\text{fQ}}}{\mathrm{d}x}\,, \tag{3.48}$$

while neglecting the last two terms the CCSD(T)/CBS+CV scheme is obtained. Analogously, the ChS (and jun-ChS) can also be employed for structural determinations. In this case, the additivity approximation is directly applied to the geometrical parameters, thus assuming that their behavior is similar to that of energy. Hence, the formulation for a generic structural parameter r is given by[165]

$$r_{\text{ChS}} = r_{\text{CCSD(T)}}^{\text{VTZ}} + \Delta r_{\text{MP2}}^{\text{CBS}} + \Delta r_{\text{MP2}}^{\text{CV}}\,, \tag{3.49}$$

where:

i. the first term is the structural parameter obtained from the geometry optimization at the fc-CCSD(T)/(jun-)cc-pVTZ level of theory;

ii. the second term accounts for the extrapolation to the CBS limit at the MP2/(jun-)cc-pVnZ level of theory by using the n^{-3} formula;[164]

iii. the last term is the core–valence correlation correction, analogous to Equation (3.43), again obtained using the MP2/cc-pCVTZ level of theory.

3.8 Computational methodology

To guide the experimental investigations, the starting point is the spectroscopic characterization of the minima. Indeed, an accurate estimate of the corresponding effective rotational constants is mandatory, which often provides sufficient information to identify and assign the rotational transitions of the species considered. In the VPT2 framework, vibrational ground-state rotational constants B_0^g can be obtained from Equation (2.51) by defining the vibrational correction $\Delta B_{0,\mathrm{vib}}^g$ as

$$\Delta B_{0,\mathrm{vib}}^g = -\frac{d_r}{2} \sum_r \alpha_r^g \,; \tag{3.50}$$

thus, it derives

$$B_0^g = B_e^g + \Delta B_{0,\mathrm{vib}}^g \,, \tag{3.51}$$

where the equilibrium rotational constants (namely the first term on the right-hand side of Equation 3.51) are straightforwardly derived from the optimized structure, while the vibrational corrections (*i.e.*, the second term) are obtained from the anharmonic force field calculation (see the treatment in Chapter 2).[167]

The two contributes affect the effective rotational constants with a different weight, where (for semi-rigid molecules) the B_e^g's usually account for more than 99%.[168] Thus, the major computational effort is devoted to the accurate determination of the equilibrium structure carried out as explained above and where the choice of the level of theory depends on molecular size. Since anharmonic force field calculations are sensibly more demanding than geometry optimizations and since vibrational corrections provide a small contribution, the $\Delta B_{0,\mathrm{vib}}^g$ corrections are usually obtained at a lower level of theory without affecting the overall accuracy.[169–172]

As a by-product of anharmonic force field calculations, the sextic centrifugal distortion constants are obtained. Quartic centrifugal distortion constants are instead derived as a by-product of harmonic force field computations, which are usually performed at a higher level of theory. Further spectroscopically relevant properties of the system are obtained as first derivatives of the energy (namely, first-order properties), *viz.* the nuclear quadrupole coupling constants and the electric dipole moments. Concerning the energetic description of the system, the zero-point energy (ZPE) correction can be obtained within the harmonic approximation, thus requiring harmonic force field calculations, but an improved determination can be derived from anharmonic force field calculations.

Chapter 4

Determination of Molecular Structures

Section 3.8 presented the computational methodology, pointing out the importance of accurate structural determinations. In light of that, this Chapter aims to provide the reader with the necessary tools for obtaining precise geometrical parameters for the species under investigation. At first, an estimate of the accuracy required to guide rotational spectroscopy experiments is provided; then, Sections 4.1 and 4.2 present joint theoretical-experimental approaches for accurate structure determinations.

The definition of "accuracy" varies alongside with the field of study. Within the rotational spectroscopy framework, the experimentally achievable precision on rotational parameters is astounding: usually, 1 part in 10^7.[13] However, extabilishing the required accuracy for the computational groundwork in order to successfully guide an experiment is a rather complicated task—it has to be noted that this Chapter is not devoted to define the level of accuracy to conduct an astronomical search relying solely on quantum-chemical (QC) methods.

As a rule of thumb, a 0.1% error on rotational constants can suffice—which implies a 10 MHz confidence interval on a 10 GHz B.*[173] In addition to this estimate, two main factors need to be accounted for: i. the value of J in the transition with the lowest measurable frequency; ii. the presence in the gas phase of molecules different from the investigated species. The impact of the former is easily deducible from Equations (2.10) and (2.19): a given uncertainty on B amplyfies with the increasing of J. A common ex-

*The superscript denoting the inertial axis is dropped for the sake of conciseness.

perimental work, in fact, starts by measuring the rotational spectrum at the lowest possible frequency available. In the framework of this thesis, given the average order of magnitude of B for medium-sized asymmetric-top rotors, ideally, the first measurements for a new species are carried out in the centimeter-wave frequency range (3–30 GHz) and then extended to the millimeter- and submillimeter-wave regions (30–300 GHz and 0.3–3 THz, respectively).

The second factor, instead, is merely justified by the increased difficulty in spectral assignments because of the presence of by-products, which lead to overcrowded spectra. This is common (and especially challenging) when the generation of the sample in the gas phase is carried out by electric-discharge or pyrolysis techniques—which are prone to the production of tens of species.

Now, in order to quantify the effect of small discrepancies in the geometry on the value of B, the vibrational contribution to the effective rotational constants (see Equation 3.51) can be neglected.[168] Therefore, in terms of the fractional change k in the equilibrium rotational constant ($\delta_k B_e / B_e$), the following relationship is obtained[173]

$$\sum_k \frac{\delta_k B_e / B_e}{\delta r_k / r_k} = -2\,, \tag{4.1}$$

where $\delta r_k / r_k$ is the fractional change on the corresponding bond distance. From Equation (4.1), it can be derived that an accuracy of *ca.* 0.5 mÅ on bond lenghts is required to obtain an error on B_e close to 0.1%. Analogously, for bond angles, uncertainties close to 0.1° yield a 0.1% error on B_e.[173] Therefore, the accuracy in the equilibrium geometries that determines a sufficiently-low error on B is achievable only by resorting to high-level quantum-chemical calculations—*i.e.*, by exploiting computationally-demanding composite schemes (see Section 3.7 and Table 3.1). Their estimated precision on bond lenghts and angles is summarized in Table 4.1. To overcome the limitations related to the computational cost, new procedures are being developed in order to obtain accurate structures for larger systems, like the TM-SE(-LR) approach defined in Section 4.2.

4.1 Interplay of theory and experiment

Rotational spectroscopy is the most precise technique for the determination of molecular structures of isolated molecules in the gas phase. In fact, the geometrical information is stored in the formulation of the rotational constants B_v (*cf.* Equations 2.11 and 2.13), which can be determined with an

Table 4.1: Accuracy of structural parameters at different levels of theory, defined as absolute error on bond lengths and angles.

Level of theory	Accuracy	Refs.
CCSD(T)/CBS+CV+fT+fQ	1 mÅ, 0.05°	128, 161, 162
CCSD(T)/CBS+CV	1–2 mÅ, 0.05–0.1°	128, 174–176
"cheap"	1–2 mÅ, 0.1–0.2°	168, 176, 177
jun-"cheap"	1–2 mÅ, 0.1–0.2°	178, 179

impressive precision from the analysis of the experimental spectrum in the v-th vibrational state.[13] This notwithstanding, the evaluation of structural parameters is rarely carried out at a comparable accuracy, and the reason lies in the effect of vibration on molecular dimensions, whose influence is significant even in the ground vibrational state (see Equation 2.51 and the related note).[13]

Therefore, several methods have been introduced over the years to account at various degrees for vibrational effects, leading to the definition of different types of structures. In this respect, it is essential to distinguish the equilibrium structure (namely, the geometry at the bottom of the potential well and denoted as r_e) from the effective one for the ground vibrational state (*i.e.*, the r_0 geometry, which is calculated from the B_0 rotational constants). Within the BO approximation, the r_e structures are invariant upon isotopic substitution: this allows to exploit isotopically-substituted species to increase the number of determinable rotational constants for the same reference geometry. Lastly, r_e solely depends on the electronic structure of the molecular system, and is therefore directly comparable with QC outcomes.

Given these premises, in 1953 Kraitchman proposed a convenient method for the calculation of the position of an atom in a molecule by exploiting the changes in the moments of inertia resulting from a single isotopic substitution. The structure determined by resorting to this approach takes the name of substitution geometry, and it is denoted as r_s.[180] In this model, zero-point vibrational effects tend to cancel-out when calculating the interatomic distances: if various combinations of isotopic data are used, r_s bond lenghts are usually shorter than the r_0 values and close to the r_e ones. Subsequently, in 1973, Watson introduced the mass-dependent r_m structure, an extension of the substitution model that however requires an extensive set of accurate isotopic data.[13,181] It is based on the use of the effective moments of inertia (I_0) of a number of isotopic species to deter-

mine, to first-order, the mass dependence of the vibrational contributions, and thus to obtain an estimate of the equilibrium structure.[13,181] However, given its high demand in terms of laboratory work and considering its shortcomings when dealing with significant changes of mass due to specific isotopic substitutions (such as the replacement of H by D), its application is strongly limited.[13]

On a separate note, one should recall that a full-experimental determination of the equilibrium geometry is feasible only in rare cases. For instance, considering a linear XYZ triatomic molecule, extensive laboratory work is required for the determation of B_0 and B_v in three singly excited vibrational states, which gives access to the experimental α_r's, and, *via* Equations (3.51), (2.11), and (2.13), to the r_e.[130]

Semi-experimental equilibrium structures

Introduced by Pulay *et al.* in 1978,[182] the semi-experimental (SE) equilibrium structures, r_e^{SE}, are considered the best approach for structural equilibrium determination of isolated molecules.[183] They are obtained from the least-squares fit (LSF) of the experimental rotational constants of different isotopologues corrected for computed vibrational contributions.[184] In fact, the SE equilibrium rotational constants are defined as[169,170]

$$B_e^{\mathrm{SE}} = B_0^{\mathrm{exp}} - \Delta B_{0,\mathrm{vib}}^{\mathrm{theo}}\,, \tag{4.2}$$

where, B_0^{exp} are the experimental ground-state rotational constants, and $\Delta B_{0,\mathrm{vib}}^{\mathrm{theo}}$'s are determined from Equations (2.51) and (2.52) within the VPT2 framework.[167] In the context of this thesis, the electronic contributions to the B_0^{exp} are instead neglected.[169,170]

From a computational point of view, the bottleneck of the SE protocol is the calculation of the cubic force field in order to determine the $\Delta B_{0,\mathrm{vib}}^{\mathrm{theo}}$ corrections (see Chapter 3). The CCSD(T) is considered as the "gold standard" for this task, but its computational scaling makes it suitable only for small systems with less than 10 atoms.[169,170] To overcome this limitation, recent works (see Refs. 169–172) have used (double-)hybrid DFT methods to determine the vibrational contribution and, therefore, to obtain the r_e^{SE} structures at a reduced computational cost. The key of this DFT-based strategy is in the determination of the sum of the α_r's from Equation (2.51), rather than the individual constants, which provides sufficiently accurate values thanks to error-cancellation. The thorough work carried out in Refs. 169–172 identified B3LYP-D3(BJ)/SNSD and B2PLYP-D3(BJ)/cc-pVTZ levels of theory as suitable for the calculation of cubic force fields for medium- to big-sized systems.

A second drawback of the SE approach is in the number of experimental data required for a complete structural characterization. Indeed, in the determination of r_e^{SE}, the spectroscopic information on the isotopologues may be limited, and choosing which parameters can be optimized might not be a trivial task. A first solution to this under-determination problem is to select a subset of parameters to keep fixed at the best possible computed geometry. This simple strategy provides satisfactory results, especially when supported by accurate QC calculations.[185] Furthermore, the subset of fixed parameters can be corrected according to the template molecule approach and/or to the linear regression method (*e.g.*, the TM-SE procedure defined in Section 4.2),[171,172] which are explained in the next paragraphs. An alternative approach to the under-determination issue has been proposed in Refs. 185 and 186, and it is based on predicate observations. Within this method, the computed structural parameters are used as additional input in the LSF procedure, with an appropiate weighting scheme, usually in relation to their accuracy. The reader is referred to Ref. 186 for details.

Template molecule approach

In order to deal with medium- to large-sized systems, new approaches based on cost-effective computational strategies have been developed in the last years. The template molecule (TM) approach, defined in Refs. 169 and 170, relies on the availability of a highly-accurate structure for a smaller (but similar) species, in order to derive corrected geometrical parameters for the larger system of interest. Given r_e^{TM} as a generic equilibrium parameter of the system of interest, this can be defined as

$$r_e^{\mathrm{TM}} = r_e^{\mathrm{lc}} + \Delta r_e(\mathrm{TM})\,, \tag{4.3}$$

where r_e^{lc} is the equilibrium parameter of the system of interest evaluated using a low-cost (lc) computational method and $\Delta r_e(\mathrm{TM})$ is the correction due to the TM approach. The latter is evaluated as

$$\Delta r_e(\mathrm{TM}) = r_e^{\mathrm{ha}}(\mathrm{TM}) - r_e^{\mathrm{lc}}(\mathrm{TM})\,, \tag{4.4}$$

where $r_e^{\mathrm{ha}}(\mathrm{TM})$ and $r_e^{\mathrm{lc}}(\mathrm{TM})$ are the highly-accurate (ha) and the low-cost equilibrium parameters for the smaller system, which acts as template for the larger one. Naturally, the low-cost method of Equation (4.4) is the same of Equation (4.3), and it is usually based on DFT. Instead, for the ha level, composite schemes rooted in CC techinques are expected to provide the required accuracy (see Table 4.1). An alternative, as explained later, is to resort to SE equilibrium structures (*i.e.*, the TM-SE method).

Linear regression approach

Starting from Equation (4.3), an analogous expression can be obtained for the linear regression (LR) method:[187]

$$r_e^{\mathrm{LR}} = r_e^{\mathrm{lc}} + \Delta r_e(\mathrm{LR})\,, \tag{4.5}$$

where $\Delta r_e(\mathrm{LR})$ is the correction due to the LR method, and it is expressed as

$$\Delta r_e(\mathrm{LR}) = a r_e^{\mathrm{lc}} + b\,, \tag{4.6}$$

where a, b are empirical parameters.[187] By substituing Equation (4.6) into Equation (4.5):

$$r_e^{\mathrm{LR}} = (1 + a) r_e^{\mathrm{lc}} + b\,. \tag{4.7}$$

Equation (4.7) shows the slope $(1 + a)$ and the intercept (b) of the linear regression, which depend on the atoms involved (namely the type of bond) and on the level of theory (which is DFT-based).[187] The LR coefficients are obtained by resorting to large databases of accurate SE structures; the outcomes are tabulated in Refs. 187 and 188. Analogously to the TM method, the LR approach improves the accuracy of the equilibrium geometrical parameters without increasing the computational effort.

4.2 Definition of the TM-SE(-LR) approach

The combination of the TM (and LR) method with the database of SE equilibrium geometries is the birthplace of the TM-SE(-LR) approach by Melli, Tonolo, Barone and Puzzarini.[188] The conceptualisation, validation and application of this procedure is a fundamental part of this thesis; the former is treated in this Section, while the reader is referred to Chapter 7 for the latter.

The idea behind the TM-SE approach is to identify, in the molecule under study, a fragment whose SE equilibrium structure is available with high-accuracy and use it to correct the geometrical parameters as in Equations (4.3) and (4.4). In this case, Equation (4.4) can be re-written as

$$\Delta r_e(\mathrm{TM}) = r_e^{\mathrm{SE}}(\mathrm{TM}) - r_e^{\mathrm{lc}}(\mathrm{TM})\,, \tag{4.8}$$

where $r_e^{\mathrm{SE}}(\mathrm{TM})$ is the semi-experimental equilibrium structure of the fragment.

The LR approach can be used for inter-fragments parameters (TM-SE-LR method). A schematic representation of this method is sketched in Figure 4.1, which shows the application of the TM-SE-LR method to

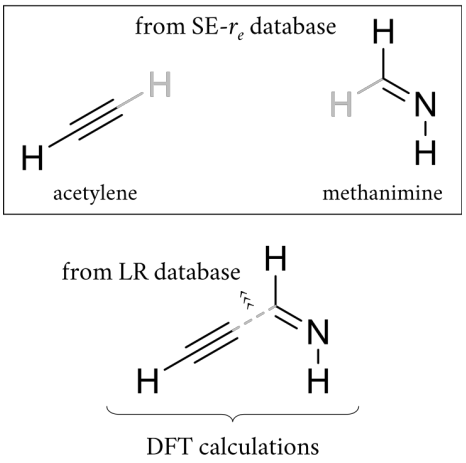


Figure 4.1: Scheme of the TM-SE(-LR) approach for (*Z*)-propargylimine.

(*Z*)-propargylimine (*cf.* Figure A.1). The procedure is composed of two steps; in particular:

i. the selection of the fragments (*i.e.*, the HCC– and –CHNH moieties) and their structural improvement using the TM-SE approach. The DFT structure is corrected by resorting to the available SE equilibrium structures of acetylene and methanimine, according to Equations (4.3) and (4.8);

ii. the improvement of the linkage of the fragments by exploiting the LR method. The DFT C–C bond distance is corrected using Equation (4.7).

If the second step is skipped and the bond distance is kept at the DFT value, the geometry is denoted as $r_e^{\text{TM-SE}}$; otherwise, $r_e^{\text{TM-SE-LR}}$.

In other words, the TM-SE fragments can be seen as a collection of bulding blocks which can be connected between each other to form new complex molecules with high accuracy, and with the computational rate determining step being the low-cost geometry optimization for the larger system.

Chapter 5

Fundamentals of Radio Astronomy

As stated in Chapter 1, radio astronomy techniques are of fundamental importance in the extension of the chemical astronomical census. In order to gain insight on the challenges and the accuracy demands of these techniques, this Chapter contains a preamble on the basic principles of radio astronomy, followed by the methods for deriving chemical informations from astronomical data.

Radio astronomy, namely the study of astronomical sources by means of radio frequencies, is a pivotal branch of astronomy. Its birthplace can be associated with the discovery of the first radio emission from the Milky Way by Jansky, in 1933; however, its bloom started in the 1950s, after the discovery of the 21 cm hydrogen line, as mentioned in Chapter 1.[25] Nowadays, it is a broad discipline with a wide range of uses, from galaxy kinetics to the study of the cosmic magnetic field, passing through the investigation of the chemical composition at different stages of cosmic evolution.

The underlying reasons for its large number of applications are manyfold. At first, radio waves are able to penetrate dust and gas, contrary to most other wavebands; this allows to observe into galaxies and molecular clouds (MCs).[25] Moreover, radio quanta store the lowest energies of all the electromagnetic spectrum (*e.g.*, 10^{-7} eV at 30 MHz, or 10^{-3} eV at 300 GHz); as a consequence, radio signals can be treated as classical waves up to the THz regime. This allows their conversion into complex voltages in the receiver, and, thus, their amplification and manipulation.[25] In addition, thanks to the development of interferometry and aperture-synthesis imaging, last generation radio telescopes provide the highest resolution images possible at any wavelenght. This made it accessible to study even

the furthest regions, near to the event horizon around the black hole in the centre of the Milky Way.[25]

Another advantage of radio astronomy lies in the absence of atmospheric blockage. Indeed, the Earth atmosphere is a barrier to all but two types of wavelenght: radio and optical (including the near-infrared). Actually, the radio window is not totally available: H_2O absorbs at 22.2 GHz and 183 GHz, while a group of lines belonging to O_2 fall close to 60 GHz, with an additional transition near 119 GHz. Other abundant molecules such as CO_2 and N_2 absorb at higher frequencies (above 300 GHz).[189] However, the water vapor issue can be partially overcome by choosing dry and high-altitude areas (at 5 000 m from sea level, the optical depth from H_2O is reduced to *ca.* 8%), while little can be done from Earth surface to prevent oxygen and other species absorption. To avoid such absorption and mainly for other regions of the electromagnetic spectrum, measurements can be carried out by resorting to airborne facilities or spacecraft observations, which however put strong constraints on the size of the telescopes.

On the other hand, the present challenge of radio astronomy comes from the reliance of the modern communication systems on radio transmissions. Interference-free regions suitable for radio astronomical observation are lacking; indeed, the current generation of telescopes is built in unconvenient remote environments, away from human activity.[25]

Within the vast range of applications of radio astronomy, in this thesis the focus is drawn to its role in Astrochemistry. On this basis, the next Section addresses the basics of radio astronomy, while Section 5.2 is dedicated to the extraction of chemical information stored in astronomical spectra.

Selected facilities

To provide a complete description of radio astronomy techniques, a brief presentation of some of the most advanced and high-sensitivity radio telescopes is deserved. According to their contribution to molecular detection (see Table 5.1), a few facilities can be considered:

i. Institut de radioastronomie millimétrique (IRAM) 30 m radio telescope: it is a single-dish telescope located on the Pico Veleta (2 850 m from sea level) in the Spanish Sierra Nevada, and it covers the 80–370 GHz frequency range;[190]

ii. Yebes 40 m radio telescope: it is a single-dish telescope located in Yebes (Spain), at an altitude of 931 m, and it operates in the 4.9–90 GHz range;[17]

iii. Atacama large millimeter/submillimeter array (ALMA): it is an array of 66 radio antennas (54 antennas with a 12 m dish and 12 with a 7 m dish) situated in Chajnantor Plateau (5 058 m from sea level), in the heart of the Atacama Desert (Chile). The antennas can be configured in different ways, and they are spaced at distances which range from 150 m to 16 km, covering a wide portion of the microwave range, from 35 GHz to 950 GHz with outstanding sensitivity. Despite the lower number of detection (*cf.* Table 5.1), it needs to be pointed out that the identified species are extremely complex (*e.g.*, urea,[41] methoxymethanol,[191] propionamide,[192] *etc.*), therefore requiring an exceptional sensitivity and high spectral and spatial resolution. As far as the millimeter/submillimeter-wave frequency domain is concerned, ALMA represents the best facility for the observation of molecular signatures in a variety of astrophysical objects, including remote galaxies and planets;[16,193]

In addition, a glance at next-generation facilities is deserved. The square kilometer array observatory (SKAO) will count thousands of tree-like antennas in Australia and 197 dish antennas in South Africa, and will be operating in the 50–350 MHz and 0.350–15.4 GHz ranges, respectively. Second, the next-generation very large array (ngVLA) will be composed of 263 dish antennas placed throughout the USA, and will operate from 1.2 GHz to 116 GHz. These facilities will have a profound impact on radio astronomy and Astrochemistry, given the fact that the most complex species detected to date are already pushing the limits of ALMA.[8]

Table 5.1: Total number of detections from the first eight most contributing facilities.

Facility	Dish diameter	# mol.
IRAM	30 m	77
Yebes	40 m	60
NRAO	36 ft	33
GBT	100 m	32
NRAO/ARO	12 m	28
Nobeyama	45 m	15
NRAO	140 ft	13
ALMA	7–12 m	10

5.1 Basic definitions

Radio electromagnetic radiation is a wave phenomenon. However, when the scale of the system involved is much larger than the wavelenght, the radiation is considered to travel in straight lines, *i.e.*, rays.[189] The brightness (or specific intensity) I_ν for radiation in free space is independent of the distance along a ray, and, neglecting diffraction and extinction effects, it varies only if an interceding material (*e.g.*, a cloud) is present, according to the equation of transfer:

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}s} = -\kappa_\nu I_\nu + \epsilon_\nu \,, \tag{5.1}$$

where ds represents the infinitesimal thickness of the cloud, while κ_ν and ϵ_ν are respectively its absorption coefficient and the emissivity.[189]

Now, one can consider the radiation emitted from a non-zero temperature objects, namely thermal emission. In general, it depends on the composition of the object as well as its temperature. In an extreme case, where all radiation is absorbed at any wavelenght without reflection, the object takes the name of a blackbody:[189] in this specific case, there is a complete equilibrium of the radiation with its surroundings, and this condition is called local thermodynamic equilibrium (LTE).[189] By solving the differential equation of radiative transfer in LTE conditions, it derives that the brightness distribution is described by the Plank function, $B_\nu(T)$:

$$I_\nu = B_\nu(T) = \frac{2h\nu^3}{c^2}\left(\exp\left(\frac{h\nu}{kT}\right) - 1\right)^{-1} , \tag{5.2}$$

which solely depends on the thermodynamic temperature T and it is instead independent on the material.[189] As an example, the isotropic cosmic microwave background (CMB) radiation—namely, the remnant of the early phases of the universe, according to the Big Bang model—can be well approximated by a blackbody with a temperature of $T = T_{\mathrm{CMB}} = 2.725\,\mathrm{K}$.[189]

By integrating Equation (5.2) over ν, the total brightness of a blackbody is obtained. Its maximum value occours at

$$\nu_{\max} = 58.789T \,, \tag{5.3}$$

where $\nu_{\max}$ and T are expressed in GHz and K, respectively.[189] Equation (5.3) is known as Wien's displacement law. As a matter of fact, the maximum of the CMB radiation falls in the microwave region (about 160 GHz). Instead, if $h\nu \ll kT$, the Rayleigh–Jeans (RJ) condition holds, and Equation (5.2) for

a blackbody takes the form

$$B_\nu^{\mathrm{RJ}}(T) = \frac{2\nu^2}{c^2} kT\,. \tag{5.4}$$

This approximation can be used for all thermal radio sources which satisfy the condition

$$\nu \ll 20T\,, \tag{5.5}$$

where ν and T are again expressed in GHz and K, respectively.[25] An important feature of the RJ approximation lies in the strict proportionality between the brightness and the temperature of the blackbody.[189] Indeed, it become of standard use in astronomy to measure the brightness in terms of brightness temperature T_b, defined as the RJ temperature of an equivalent black body which will give the same infinitesimal power of the celestial source.[189] The brightness temperature can be derived from measurement of the flux and the source size, and, if the apparent source size is used, the main beam brightness temperature T_{mb} is obtained.[189]

In general, the absorption coefficient and the emissivity (*cf.* Equation 5.1) are functions of s.[2] Defining the optical depth as

$$\mathrm{d}\tau_\nu = \kappa_\nu \mathrm{d}s\,, \tag{5.6}$$

the equation of transfer (Equation 5.1) can be re-written as

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = -I_\nu + S_\nu\,, \tag{5.7}$$

where S_ν is the source function

$$S_\nu = \frac{\epsilon_\nu}{\kappa_\nu}\,. \tag{5.8}$$

Once the value of the source function is known, Equation (5.7) can be solved and the intensity at any position can be obtained. As boundary condition, the intensity at $s = 0$ is set at $B(T')$, *i.e.*, the blackbody radiation at temperature T'. If the cloud exists only for $s \geq 0$, at $s = 0$ the CMB radiation ($T' = T_{\mathrm{CMB}}$) is considered.[2] By integration over s, at $s = L$ it is obtained

$$I_\nu = S_\nu + \exp(-\tau_\nu)\left[B_\nu(T_{\mathrm{CMB}}) - S_\nu\right], \tag{5.9}$$

with

$$\tau_\nu = \kappa_\nu L\,, \tag{5.10}$$

where L is the physical thickness of the cloud. If the absorption coefficient is small, optically thin cases (namely, low values of τ_ν) can occour even for large physical thicknesses.[2]

In order to evaluate the source function, one can consider the cloud as enclosed by a blackbody at temperature T in thermal equilibrium with the cloud.[2] The intensity observed from the cloud is $B_\nu(T)$; indeed, the cloud remains a blackbody, and $T_{\mathrm{CMB}} = T$. Therefore, Equation (5.9) can be re-written as

$$B_\nu(T) = S_\nu + \exp(-\tau_\nu)\left[B_\nu(T) - S_\nu\right]. \tag{5.11}$$

Being the blackbody radiation independent of the composition, the relation expressed in Equation (5.11) holds for every value of τ_ν; this occours if

$$S_\nu = B_\nu(T), \tag{5.12}$$

or, in other words, if Kirchhoff's law is valid. Indeed, the absorption coefficient and the emissivity are not independent, but their ratio (namely, the source function) is constrained by the temperature through the Planck function.[2]

By substituing Equation (5.12) into Equation (5.9), it is obtained

$$I_\nu = B_\nu(T) + \exp(-\tau_\nu)\left[B_\nu(T_{\mathrm{CMB}}) - B_\nu(T)\right]. \tag{5.13}$$

However, in actual observations, to avoid the influence of the atmospheric conditions and of the telescope system, the target position and a nearby position without any cloud emission or absorption, with $T = T_{\mathrm{CMB}}$, are alternatively observed.[2] For the off-source position,

$$I'_\nu = B_\nu(T_{\mathrm{CMB}}), \tag{5.14}$$

and, therefore, their difference is expressed as

$$\Delta I_\nu = I_\nu - I'_\nu = \left[B_\nu(T) - B_\nu(T_{\mathrm{CMB}})\right]\left[1 - \exp(\tau_\nu)\right], \tag{5.15}$$

which is the actual intensity observed by radio telescopes.[2] Finally, as pointed out above, intensities in astronomy are usually represented by a temperature scale; in RJ conditions ($h\nu \ll kT$, see Equation 5.4), Equation (5.15) can be re-written as

$$\Delta T = (T - T_{\mathrm{CMB}})\left[1 - \exp(-\tau_\nu)\right]. \tag{5.16}$$

5.2 Derivation of molecular abundances

An expression for the optical depth can be derived by considering the radiation-matter interaction in a two-level (lower, l, and upper, u) system. While the mathematical steps are left to the reader,[2] the following

expression is obtained

$$\tau_{ul} = \frac{8\pi^3 S_{ul}\mu^2}{3h\Delta v U(T)}\left[\exp\left(\frac{h\nu_{ul}}{kT}\right) - 1\right]\exp\left(-\frac{E_u}{kT}\right)N\,, \tag{5.17}$$

where: i. S_{ul} is the line strenght (see Equation 2.33); ii. μ is the dipole moment; iii. Δv is the line width* expressed as velocity width; iv. $U(T)$ is the partition function at temperature T (*cf.* Equations 2.35 and 2.36); v. ν_{ul} is the energy difference between the two energy levels expressed in frequency; vi. E_u is the energy of the upper level; and vii. N is the column density. The velocity width is derived from[2]

$$\Delta v = \frac{\Delta\nu}{\nu_{ul}}c\,, \tag{5.18}$$

where $\Delta\nu$ is the line width in frequency units. The velocity width is the usual expression for line widths in astronomy, while the line width largely depends on Doppler effect.[2] The column density N is related to the depth of the cloud by

$$N = nL\,, \tag{5.19}$$

where n is the total number of molecules.[2] N is thus a measure of the amount of intervening matter between the observer and the source. From Equation (5.17), it is clear that τ depends on the column density, on the line strength, and on the square of the dipole moment. Unless L is known, only the column density can be derived from observations.

In the derivation of Equation (5.17), the level population is assumed to be represented by the Boltzmann distribution. If this condition is satisfied, the system is in the LTE conditions.[2] Furthermore, the excitation temperature T_{ex} for the $u \leftarrow l$ transition can be defined as

$$\frac{n_u}{n_l} = \frac{g_u}{g_l}\exp\left(-\frac{h\nu}{kT_{ex}}\right), \tag{5.20}$$

where n and g are respectively the number of particles and the degenerancy of the involved states. In LTE conditions, T_{ex} coincides with the rotational temperature.

When $\tau \ll 1$, the line is referred to as optically thin, usually implying a low line strenght or a low column density. In this case, assuming that the RJ and the LTE conditions apply and $T \gg T_{\mathrm{CMB}}$, it is obtained[2]

$$\Delta T\Delta v = \frac{8\pi^3 \nu_{ul} S_{ul}\mu^2}{3kU(T)}\exp\left(-\frac{E_u}{kT}\right)N\,. \tag{5.21}$$

*The line width depends mainly on the Doppler effect due to the motion of the cloud. In the Equation (5.17), it has been considered as a portion of a square wave.[2]

By rearranging Equation (5.21), the following expression is derived[2]

$$\ln\left(\frac{3kW}{8\pi^3\nu_{ul}S_{ul}\mu^2}\right) = \ln\left(\frac{N}{U(T)}\right) - \frac{E_u}{kT}\,, \tag{5.22}$$

where W is the integrated intensity ($W = \Delta T\Delta v$ in the approximation of a square wave). If the LTE condition is satisfied, the left-hand side of Equation (5.22) can be plotted against E_u for different transitions of the same species and a linear relation is obtained.[2] In such a case, T and N can be derived from slope and intercept, respectively.[2] This method, known as the rotation diagram method,[194] is usually applied to warm and dense clouds such as star-forming cores, where all the conditions are satisfied. This method is not appropriate for cold clouds, as LTE does not always hold.[2]

Instead, if $\tau \gg 1$, we have the so-called optically thick case, which is usually due to high abundances, and the intensity is simply[2]

$$\Delta T = T\,, \tag{5.23}$$

if the RJ condition is satisfied and $T \gg T_{\mathrm{CMB}}$. In this case, all the photons emitted are reabsorbed in the cloud and the intensity is given by the surface temperature, without allowing the derivation of the column density. In this case, rare isotopologues may be used for determining the column density by making use of the isotopic ratio.[2] However, one has to keep in mind that the latter depends on the distance from the center of the galaxy and on isotope-selective photodissociation reactions. Eventually, the hyperfine structure of the transitions may help in determining column densities for optically thick species.[2]

Equation (5.17) contanis further assumptions beyond those delineated above. In a detailed description, the physical non-homogeneities along the line of sight of the clouds have to be taken into account, as well as the limited spatial resolution of the telescopes.[2] The latter can cause the so-called beam dilution effect, when the line intensity is strongly decreased due to the source size being smaller than the telescope resolution.[2] Moreover, the column densities are function of $N(\mathrm{H_2})$, which is itself a function of the column density of rare isotopic species of CO. Thus, effects such as CO depletion and isotope-selective photodissociation of CO have to be considered.[2]

Materials and Methods

Chapter 6

Computational and Experimental Facilities

Having addressed the Theoretical Background in Chapters 2–5, and before moving to the presentation of the results and their discussion (Chapters 7–13), a description of the computational and experimental facilities is mandatory. This is the focus of the present Chapter. In particular, Section 6.1 addresses the quantum-chemical (and non) software, together with the computing facilities. Subsequently, the employed spectrometers are described in Sections 6.2 and 6.3. Finally, the laboratory conditions of choice are reported for each species in Section 6.4.

Unlike the spectroscopic and astrochemical underlying reasons described in Section 1.3—which are various and contingent—there is a relatively uniform experimental approach that is applicable to the molecules investigated in this thesis. At first, the rotational spectrum is predicted using the computed spectroscopic parameters and, when possible, data from previous characterizations. Then, the laboratory study is carried out in the millimeter-wave range. In some cases, a prior characterization using a high-resolution centimeter-wave spectrometer is also conducted. Along the entire thesis, the notation used to refer to rotational transitions is $J_{K_a,K_c} \leftarrow J'_{K'_a,K'_c}$, where the prime denotes the (*pseudo-*)quantum numbers of the energy level from which the transition starts (as also indicated by the arrow direction). In these experimental steps, the criticalities due to the physical properties and availability of the investigated molecule arise and need to be addressed. Finally, the rotational spectrum is analysed, and a set of accurate spectroscopic parameters is determined. This operation is carried out by fitting the rotational transitions to the global molecular Hamiltonian in a weighted non-linear fit procedure, the weight being

proportional to the inverse square of the frequency uncertainty.

6.1 Software and high performance computing

In order to perform the calculations described in Chapter 3, according to the computational methodology explained in Section 3.8, the access to high performance computing facilities is mandatory. In this thesis, the workstations of the Rotational and Computational Spectroscopy Group* (Department of Chemistry "Giacomo Ciamician", Alma Mater Studiorum University of Bologna, Italy) led by Prof. Dr. Cristina Puzzarini and the Avogadro computing cluster of Scuola Normale Superiore (Pisa, Italy), thanks to Prof. Vincenzo Barone, have been employed. In detail, all the DFT and MP2 calculations have been performed using Gaussian16, Revision C.01,[195] while the CC calculations have been computed using the Cfour package[196] and its interface with Mrcc for CCSDTQ calculations.[197,198] Basis sets are available from the libraries provided by the QC programs. The LSF procedure described in Chapter 4 has been carried out resorting to the Xrefit module of the Cfour package.[196]

Rotational spectra visualization mainly relied on the Pgopher program,[199–201] while the analysis (*i.e.*, the determination of the spectroscopic parameters of Chapter 2) has been carried out using the Spfit/Spcat subroutines of the Calpgm suite of programs by Pickett,[202] available from the Prospe website†.

Finally, the Astromol package provides a database of molecules detected in space, and has been used to generate Tables 1.1 and 5.1.[203]

6.2 The FM-mmW spectrometer

The frequency modulated millimeter-wave (FM-mmW) spectrometer, operating in the range from 80 GHz to 1.2 THz, is located at the Department of Chemistry "Giacomo Ciamician" at the Alma Mater Studiorum University of Bologna, in the laboratory of the Rotational and Computational Spectroscopy Group led by Prof. Dr. Cristina Puzzarini. Depending on the signal-to-noise ratio, if the Lamb-dip technique‡ is not exploited, the average experimental accuracy achievable on the transition frequency lies in

*See also: `site.unibo.it/rotational-computational-spectroscopy/en`.

†See also: `info.ifpan.edu.pl/~kisiel/prospe`.

‡In the Lamb-dip technique Doppler broadening is eliminated and, therefore, the resolution of the transition is sensibly improved.[204] The reader is referred to 204 for details.

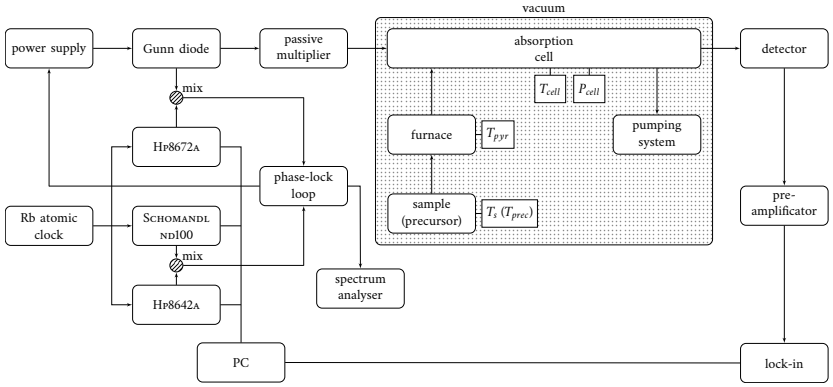


Figure 6.1: Block diagram of the FM-mmW spectrometer. See text for the description of the operating system. T_{cell} and P_{cell} are the temperature and pressure of the absorption cell, respectively. T_{pyr} refers to the temperature of the pyrolysis oven, while T_s (T_{prec}) is the temperature of the sample (precursor).

the 10–60 kHz range depending on the frequecy region, the signal-to-noise ratio (S/N), and the blending by other features.

Operating system

Following the block diagram in Figure 6.1, the microwave radiation source of the FM-mmW spectrometer is a Gunn diode working in the 80–115 GHz range (J.E. CARLSTROM CO.), which can be coupled with a series of passive frequency multipliers in order to reach higher frequencies up to 1.2 THz. In particular, a frequency tripler (VDI) and a quadrupler (RPG) have been employed in this thesis, while the sextupler and nonupler have not been used. To cover the entire working range, the fundamental band can be extended to the 75–134 GHz range by resorting to different Gunn diodes. For its stabilization, the Gunn diode frequency is mixed with the radiation coming from a 2–18 GHz synthetizer (HP8672A), and their difference enters a phase-lock loop (PLL). In the PLL, the difference signal is locked to a 75 MHz modulated signal obtained by mixing the radiation coming from a centimeter-wave synthetizer (HP8642A) and from a SCHOMANDL ND100 frequency synthetizer. This combination is used to exploit the faster scan operation of the latter and the frequency modulation of the former. Finally, the PLL acts on the power supply of the Gunn diode to stabilize and modulate its emission. The electronics are referenced to a Rb atomic clock with an accuracy of 10^{-3} Hz.

The radiation is focused into the absorption cell using a parabolic

mirror. The cell is a Pyrex® glass tube (3.25 m long, 10 cm in diameter) with high-density polyethilene windows, which is kept at room temperature (T_{cell}). A pumping system—consisting in a diffusion and a rotary pump—ensures high vacuum conditions, with the working P_{cell} being of the order of a few mTorr (*ca.* 1 kPa). The sample reservoir (possibly coated with heating tape or immersed in an ice bath, thus acting on T_s or T_{prec}) can be directly connected to the cell through a flow-regulating valve. If the flash vacuum pyrolysis (FVP) is exploited, a quartz tube heated by a 30 cm-long furnace (with a maximum T_{pyr} of 1 200 °C) is positioned between the absorption cell and the precursor reservoir. This configuration is shown in Figure 6.1. Different setups using a double-pass configuration, an electric glow-discharge system or a heated absorption cell are available (*e.g.*, see Refs. 205, 206 and 193, respectively).

Up to the fourth-harmonic region (*ca.* 450 GHz), the detection system consists of Schottky barrier diodes (Vdi) connected to a lock-in amplifier, the latter demodulating the signal at twice the modulation frequency: this allows the recording of a $2f$ spectrum profile. A InSb bolometer is used at higher frequencies. The settings of the electronics and the spectrum recording, are guided using a home-made program operating on a PC.

6.3 The COBRA-FTMW spectrometer

The coaxially oriented beam resonator arrangement (COBRA) Fourier-transform microwave (FTMW) high-resolution spectrometer, COBRA-FTMW, working in the 3–26.5 GHz region, is located at the Institut für Physikalische Chemie und Elektrochemie of the Leibniz Universität in Hannover, Germany, at the Molecular Beam Spectroscopy laboratory* led by Prof. Dr. Jens-Uwe Grabow. In normal conditions, the experimental accuracy on the transition frequency recorded with this spectrometer is 2–3 kHz.

Operating system

Its working principle is based on the observation of the free induction decay (FID) of a molecular beam from a pulsed excitation. The FID originates from the relaxation of the system from a stationary equilibrium due to the interaction with an electromagnetic field to the original equilibrium conditions.[207] This spectrometer design is an improved version of the previous molecular beam Fabry–Pérot resonator spectrometers, and it is

*See also: `pci.uni-hannover.de/en/research/research-groups/grabow-group`.

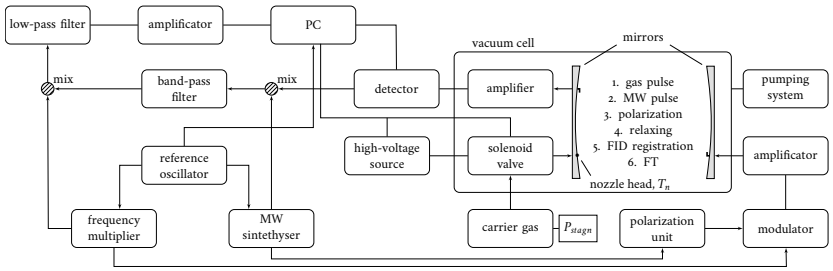


Figure 6.2: Block diagram of the COBRA-FTMW spectrometer. See text for the description of the operating system. T_n is the temperature of the nozzle head, while P_{stagn} is the stagnation pressure of the carrier gas.

able to achieve a linewidth (half width at half maximum, HWHM) relative to the molecular emission frequency better than 10^{-7}.[207]

Following the steps shown in the enumeration inside Figure 6.2, at first, the carrier gas (He, Ne or Ar, at the stagnation pressure P_{stagn}) is introduced in the chamber by a solenoid valve, tipically operated with a pulse length of 250 µs.[207] To register the spectra, the species of interest should be released in the gas phase, and different options are available for this purpose. If the sample has a significant vapor pressure, a mixture of carrier gas and sample can be prepared and collected in a storage cylinder and flew into the nozzle. If the vapor pressure is low, the liquid or solid sample can be filled in the nozzle head, eventually combined with a heater coil (the temperature of the nozzle head being T_n), and carried to the gas phase by the flux of the carrier gas.[207]

Then, a short coherent monochromatic radiation pulse excites the gas sample in the resonator *via* an L-shaped antenna, with the stability of the frequency being obtained by synchronization to a radio station. The polarization unit assures the required sensitivity even in the investigation of species with a small transition moment.[207] At the same time, the detection system is protected from the pulse. At this stage, the sample is macroscopically polarized (third step in Figure 6.2), and when the pulse is removed, it starts to relax by emission.

Subsequently, the detection phase begins: a switch opens the signal path from the receiver antenna and closes the path from the excitation section, preventing additional noise to be recorded.[207] Indeed, all sources of noise are significant, since very weak molecular signals lay under the thermal background noise. Thanks to the coaxial arrangement of the microwave resonator and of the gas expansion source, both sensitivity and resolution can be greatly improved with respect to a Fabry–Pérot

resonator.[207] Indeed, the molecular beam propagates along the symmetry axis of the two mirrors, maximizing the transit time and the volume of the gas interacting with the active region.[207] The coaxial arrangement is also responsible for the Doppler splitting of all transitions, the emission occuring in the same or opposite direction of the molecular beam.

Finally, the molecular signal in the microwave is downconverted to radio frequency before being digitalized, due to the huge amount of data to be sampled.[207] A phase quadrature detection is used, followed by performing a complex Fourier transformation of the signal. A series of noise suppressors and the repetition of the 1–6 sequence shown in Figure 6.2 thousands of times guarantee the desired S/N.[207]

Resonator specifications

The resonator consists in a pair of confocal mirrors, sketched in Figure 6.2, the fixed one incorporating an electromechanical valve and two pairs of antennas for covering the whole 3–26.5 GHz range, while the mobile one is used to adjust the resonator lenght thanks to a motorized micrometer, following the relationship

$$\frac{\Delta d}{\Delta \omega} = -\frac{d}{\omega}\,, \tag{6.1}$$

where d is the resonator lenght, ω is the resonance frequency, and Δd and $\Delta\omega$ are small increases in their values.[207] Thanks to that and to the total control of the instrument by the Ftmw++ program on a PC, the frequency scanning is fully automated. The narrow linewidth obtained with the COBRA-FTMW spectrometer also depends on the free expansion and propagation of the molecular beam in the resonator. Due to the high dilution of the mixture being pulsed into the chamber, the thermodinamical properties of the carrier gas can be approximated to be those of the mixture.[207] Furthermore, the approximation of monoatomic ideal gas is justified by using He, Ne or Ar as carrier gases. From that, the velocity v of the molecular beam can be derived as[207]

$$v = \sqrt{\frac{5R(T_0 - T)}{M}}\,, \tag{6.2}$$

where R is the ideal gas constant, T_0 is the temperature of the gas mixture, T is the temperature of the expansion, and M is the molar mass of the gas.[207] By considering a complete conversion ($T = 0$ K), the resulting velocities are $v_{\mathrm{He}} = 1\,760\ \mathrm{m\,s^{-1}}$, $v_{\mathrm{Ne}} = 790\ \mathrm{m\,s^{-1}}$, and $v_{\mathrm{Ar}} = 560\ \mathrm{m\,s^{-1}}$. Therefore, a

molecular signal of approximately 1 ms yields a flight distance of at least 0.5 m, thus requiring a chamber fitting a resonator of 1 m to avoid excited molecules hit the mirror before emitting.[207] The experiment is repeated from tens to tens of thousands times to aquire a signal, thus an efficient high vacuum is required to avoid the remaining gas interfering with the experiment. An oil diffusion pump assuring a vacuum of the order of 10^{-4} Pa allows a repetition rate of 20–40 Hz.[207]

6.4 Laboratory conditions

In the context of rotational spectroscopy experiments, the expression "laboratory conditions" refers to the set of physical quantities (such as temperature and pressure) and instrumental settings employed in the experiment, together with the procedure used to generate the molecule of interest in the gas phase. As expected, their choice strongly depends on the physical properties of the investigated molecule: primarily, its reactivity has to be taken into consideration. The molecules under study in this thesis can be divided in two groups, according to their availability from chemical vendors:

i. n- and iso-propylamine, 2- and 3-furonitrile, 3-aminoisoxazole;

ii. cyanovinylacetylene, allenylacetylene, and phenylmethanimine.

The former set includes stable species which are obtainable from Sigma-Aldrich at a high level of purity (99%); on the contrary, no chemical vendor can provide the molecules belonging to the latter group, thus requiring their synthesis to be addressed in this thesis.

FM-mmW experiment

For the first set, the FM-mmW experiment has been carried out using commercial samples, which have been used without any further preparation, with the exception of a degasification using liquid N_2. Although globally similar, the optimal experimental conditions vary with respect to the investigated molecule, depending on their physical state, vapor pressure, and dipole moment (which affects the intensity of the lines; see Equation 2.32). The experimental conditions of these five species are summarized in Table 6.1.

Instead, for the molecules of the second group, the synthesis needs to be addressed. If the investigated species is stable and isolable, it can be synthetised in the organic chemistry laboratory, and then brought to the gas

phase in the spectrometer. This is not the case for phenylmethanimine and allenylacetylene, whose isolation is not reported in literature.[208–211] Therefore, an *in situ* generation is mandatory. The same strategy has been chosen for cyanovinylacetylene, although a synthesis of (*Z*)-cyanovinylacetylene was reported in Ref. 212.

Radical and ions are transient species which are typically generated *in situ* using plasma techniques;[213,214] instead, for neutral unsaturated species flash vacuum pyrolysis (FVP) is usually exploited, which is a reliable method employed in several works (see, as example, Refs. 215–218). Within this approach, one needs to find a precursor able to generate the target molecule in the gas phase. The choice can be driven by chemical intuition or, when possible, by the literature. In particular, the choice of 2,3-pyridinedicarboxylic anhydride as precursor of (*E*)-cyanovinylacetylene was already suggested in Ref. 219, while no information was available for allenylacetylene and phenylmethanimine (with the exception of the former being a co-product of benzene electric discharge, see Ref. 220). The same precursor and generation method used for the *E* isomer, have successfully been exploited also for (*Z*)-cyanovinylacetylene.

Instead, two different precursors have been tested for allenylacetylene, namely dipropargylamine and tripropargylamine. A portion of spectrum is reported in Figure 6.3. Both precursors generated the target species in the gas phase (even if in small abundance), but the former has been chosen thanks to the presence of less interfering lines in the rotational spectrum.

For (*E*)- and (*Z*)-phenylmethanimine, given that the thermolysis of hydrobenzamide (HBA, or 1-phenyl-*N*,*N'*-bis(phenylmethylene)-methanediamine) is not effective (*cf.* Chapter 10), FVP has been selected as generation method. According to the available literature on its vacuum dynamic

Table 6.1: Laboratory conditions for the five commercial molecules investigated in this thesis, using the FM-mmW spectrometer. For all these cases, $T_{cell} = 25\,^{\circ}\mathrm{C}$. The last column contains the conditions of the sample reservoir.

Molecule	State	T_s	P_{cell}	Notes
n-propylamine	liquid	25 °C	3–5 mTorr	
iso-propylamine	liquid	25 °C	3–5 mTorr	
2-furonitrile	liquid	0 °C	2–3 mTorr	ice/water bath
3-furonitrile	solid	35 °C	2–3 mTorr	alternately immersed in a warm water bath
3-aminoisoxazole	liquid	25 °C	5 mTorr	

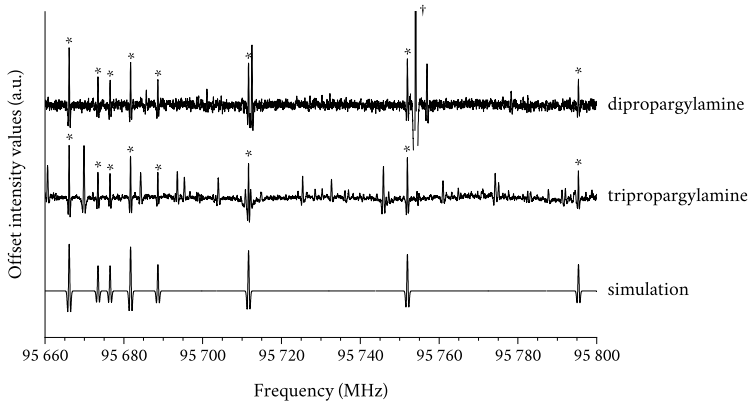


Figure 6.3: Testing of different precursors for allenylacetylene: from top to bottom, dipropargylamine, tripropargylamine and the simulated spectrum. Transitions marked with an asterisk were assigned to allenylacetylene. The strong interfering feature marked with a dagger is out of scale.

preparation through dehydrogenation or loss of methane,[221] two precursors have been tried: respectively, α-methylbenzylamine and benzylamine. In addition, *N*-methylbenzylamine has also been tested. The results of the tests are displayed in Figure 6.4: they show two distinct transitions of (*E*)-phenylmethanimine when both benzylamine and α-methylbenzylamine are used, while none with *N*-methylbenzylamine. Therefore, the choice fell on α-methylbenzylamine because of its lower signal-to-noise ratio. No signal ascribable to (*Z*)-phenylmethanimine has been observed *via* FVP.

All FM-mmW experiments (including the species from the first group) have been performed in dynamical conditions, with a tenuous flow of fresh pyrolysis products continuously assured by the vacuum system. The optimized experimental conditions for the molecules of the second group are collected in Table 6.2.

COBRA-FTMW experiment

As stated at the beginning of this Chapter, the usual experimental course of action may include measurements in the centimeter-wave region, prior to the FM-mmW experiment. Especially, this is the case when no spectroscopic information is available for the species under investigation. Table 6.3 summarizes the corresponding literature, collecting all the publications on the investigated molecules preceding this work. As shown in Table 6.3, no previous characterization was found for iso-propylamine, 3-furonitrile,

phenylmethanimine, and 3-aminoisoxazole.

For the first two molecules, a study in the cm-wave range has been carried out by scholars at the School of Chemistry and Chemical Engineering of the Chongqing University, in China, in the group led by Prof. Dr. Qian Gou*, using a COBRA-like FTMW spectrometer. The estimated accuracy for frequency measurements with this instrument is better than 3 kHz. For the sake of completeness, the same characterization has also been carried out for n-propylamine and 2-furonitrile. For iso- and n-propylamine, a gas mixture (1% in He) with stagnation pressure (P_{stagn}) of 2 bar has been supersonically expanded through the solenoid valve into the cavity of the spectrometer. Instead, furonitrile isomers (obtained from Adamas at 97% purity) have been placed at room temperature in a reservoir inserted in the gas line, where He acted as carrier gas with P_{stagn} in the 1–2 bar range.

Concerning phenylmethanimine and 3-aminoisoxazole, the study in the centimeter-wave region has been carried out using the COBRA-FTMW spectrometer described in Section 6.3. The former has been generated *in situ* by thermolysis of HBA. This *pseudo*-trimer of phenylmethanimine, which releases the monomer upon heating, has been synthesized according to the procedure available in literature[222–226] (reported in Section A.2). The phenylmethanimine isomers have been produced in the gas phase by heating the nozzle head filled with HBA at a temperature of $T_n = 80\,^{\circ}\text{C}$, under a flux of Ne as carrier gas with $P_{stagn} = 1$ bar. The reason behind the different generation technique with respect to the FM-mmW experiment is addressed in Chapter 10.

As mentioned above, 3-aminoisoxazole is commercially available; hence, the nozzle head has been filled with glass wool and a few drops of the sample. A flux of Ar at $P_{stagn} = 1$ bar ensured the transportation of 3-aminoisoxazole in the gas phase.

*See also: `hgxy.cqu.edu.cn/info/1510/4550.htm`.

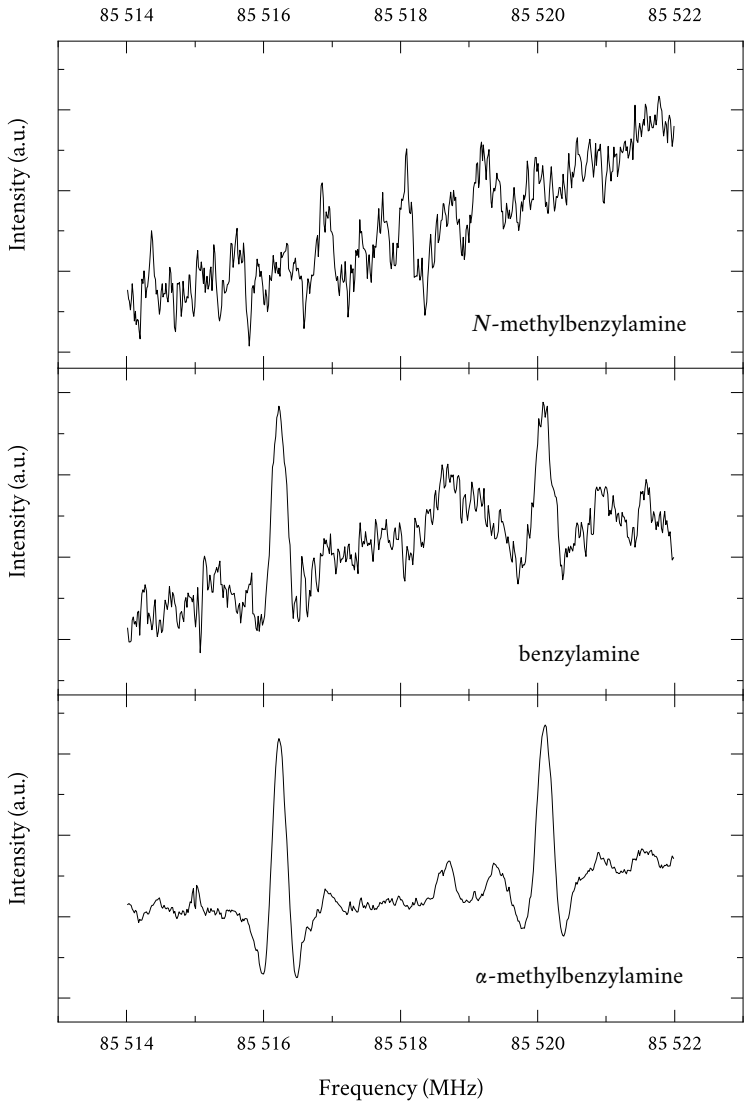


Figure 6.4: Test of different precursors for (*E*)-phenylmethanimine: from top to bottom, *N*-methylbenzylamine, benzylamine, and α-methylbenzylamine.

Table 6.2: Laboratory conditions for the three reactive species and their isomers investigated in this thesis, using the FM-mmW spectrometer. For all these cases, $T_{cell} = 25\,^{\circ}\mathrm{C}$. The last column contains the conditions of the precursor reservoir.

Molecule	Precursor	State	T_{prec}	T_{pyr}	P_{cell}	Notes
cyanovinylacetylene	2,3-pyridinedicarboxylic anhydride	solid	120 °C	950°C	6–7 mTorr	wrapped in heating tape
allenylacetylene	dipropargylamine	liquid	25 °C	800°C	9–12 mTorr	
(*E*)-phenylmethanimine	α-methylbenzylamine	liquid	25 °C	890°C	13 mTorr	

Table 6.3: Collection of the previously available spectroscopic information and references on the investigated species. The last column indicates which spectrometer has been used for the possible study in the centimeter-wave range.

Molecule	Ref(s).	Previous characterization	Cm-wave study
iso-propylamine	227, 228	Up to 50 GHz, only for *trans* isomer	Chongquin, China[a]
n-propylamine		None	Chongquin, China[a]
(*E*)-cyanovinylacetylene	219, 229	Up to 40 GHz	None
(*Z*)-cyanovinylacetylene	212	Up to 15 GHz	None
allenylacetylene	75, 220	Up to 50 GHz	None
2-furonitrile	230, 231	Up to 40 GHz	Chongquin, China[a]
3-furonitrile		None	Chongquin, China[a]
phenylmethanimine		None	COBRA-FTMW, Hannover, Germany
3-aminoisoxazole		None	COBRA-FTMW, Hannover, Germany

[a] The work has been carried out by scholars at the School of Chemistry and Chemical Engineering of the Chongqing University (China), in the group led by Prof. Dr. Qian Gou, using a COBRA-like FTMW spectrometer.

Results and Discussion

Chapter 7

Testing of the TM-SE(-LR) Approach

The validation and testing of the TM-SE(-LR) method introduced in Section 4.2 has been carried out over a dataset of molecules generated by addition/elimination reactions of nucleophylic unsaturated radicals (*e.g.*, CN, C_2H, and phenyl) to alkenes, imines, and aldehydes. Then, the approach has been applied to the species investigated in this thesis.

As a matter of fact, free radical addition to double bonds is a known reaction, extensively studied thanks to its interest in several fields, from organic synthesis to atmospheric chemistry and astrochemistry (see Refs. 232–235 and references therein). Moreover, the addition/elimination products derivated from alkenes and aldehydes are well known, although accurate structures of some of them are not available. In particular, the situation is more involved for imines, which play a pivotal role in Astrochemistry, and whose characterization is hampered by their high reactivity under terrestrial conditions.[216,235–237]

Aside from the underlying chemical motivations, the composition of the dataset has been chosen according to two selection criteria: i. the availability of the rotational constants of the main isotopic species; and ii. the possibility to envisage each molecule as formed by two fragments, whose SE equilibrium structures are available. The molecular dataset is shown in Figure 7.1, while the list of the used reference studies is reported in Table 7.1.

The dataset is composed of 16 species (21, if isomers are considered) which are the result of the combination of seven fragments: methanimine (CH_2NH), formaldehyde (H_2CO), hydrogen cyanide (HCN), acetylene (C_2H_2), ethene (C_2H_4), acetonitrile (CH_3CN), and benzene (C_6H_6). The SE equilibrium structures of the fragments are available from the SE database of

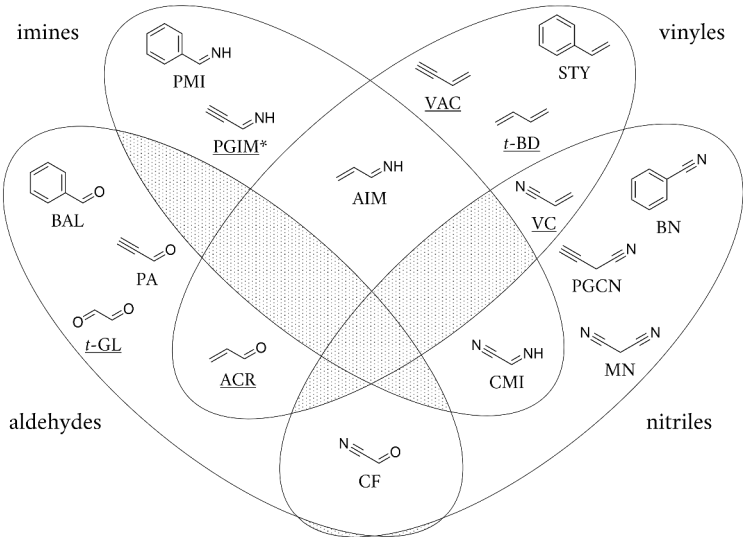


Figure 7.1: Representation of the molecular dataset for the validation and testing of the TM-SE(-LR) method. The species are divided in four groups according to their chemical moieties. The nomenclature used to refer to the dataset items is also shown. The underlined molecules constitute the r_e^{SE} database used for the validation of the method (the asterisk indicates that only the Z isomer of propargylimine is used for the latter).

Refs. 169–172, with the only exceptions being methanimine and acetonitrile, whose structures are respectively taken from Refs. 185 and 238. All the r_e^{SE} geometries are collected in Table A.1. For a subset of this database (indicated as underlined in Figure 7.1), the r_e^{SE} structure is available: these species compose the r_e^{SE} database which will be used for the validation of the method.

As it can be seen from their molecular structures shown in Figure 7.1, all the products of these reactions can be seen as formed by two distinguishible and defined moieties. Moreover (the exceptions being malononitrile, MN, and propargyl cyanide, PGCN), the linking between the moieties consists in a single bond connecting two sp^2 and/or sp hybridised carbon atoms: non-negligible conjugation effects are thus expected. This aspect, together with the high level of flexibility of certain molecules (*i.e.*, the C–C–C frame of the two non-conjugated molecules, MN and PGCN), demonstrates that the selected one is a challenging test suite to validate the TM-SE(-LR) model.

As explained in Section 4.2, for the application of the TM-SE(-LR) approach, a reliable yet cost-effective level of theory needs to be selected.

Table 7.1: Reference works reporting the experimental ground-state rotational constants (B_0^{exp}) and available SE equilibrium structures (r_e^{SE}, in brackets) for the dataset molecules.

Molecule	Ref(s).	Molecule	Ref(s).
cis-acrolein	239 [240]	malononitrile	241
trans-acrolein	242 [240]	phenylmethanimine	243
allylimine	244	propargylcyanide	245
benzaldehyde	246	(*E*)-propargylimine	247
benzonitrile	93	(*Z*)-propargylimine	247 [–][a]
trans-1,3-butadiene	248 [171, 172]	propriolaldehyde	249
cyanoformaldehyde	250	styrene	251
cyanomethanimine	216	vinylacetilene	252 [253]
trans-glyoxal	254 [254]	vinylcyanide	255 [253]

[a] The determination of the r_e^{SE} has been carried out in this work (*cf.* Section A.1).

In this work, the double-hybrid revDSD-PBEP86-D3(BJ) functional has been used, in conjunction with the jun-cc-pVTZ basis. The choice has been driven by a recent benchmark study proving its good performance in terms of structural determinations.[256] This level of theory will be referred to as revDSD, and it will constitute the "low-cost" level of theory which appears in the formulation of the TM-SE and LR methods (see Equations 4.3 and 4.7, respectively). For the latter, the coefficients of the LR correction employed are $a = -0.0067$ and $b = 0.0069$. These values have been determined from a study conducted at the revDSD level including nearly 100 semi-experimental structures.[187] The optimized revDSD geometries of the seven fragments are collected in Table A.1.

Once validated using the available r_e^{SE} structures, the question on how to assess the accuracy of the TM-SE(-LR) equilibrium structures arises. While their comparison with the SE equilibrium structures can provide a validition, the dataset is limited because the TM-SE(-LR) approach is actually designed for those cases in which the standard SE method cannot be applied. Therefore, since rotational constants are directly connected to the molecular structure and their determination can be carried out with great accuracy, they represent the perfect means for this test.

It is of fundamental importance to recall that when moving from the ground-state to the equilibrium rotational constants and *vice versa*, the vibrational corrections are required (see Equation 3.51). In the VPT2 framework, these are expressed according to Equation (3.50), and are determined from anharmonic force field calculations. The method of choice for the

latter is the global-hybrid B3LYP-D3(BJ) functional, in conjunction with the partially augmented double-ζ jun-cc-pVDZ basis set. Hereafter, this level of theory is referred to as B3.

This Chapter is organized in two Sections. Section 7.1 presents the validation of the TM-SE(-LR) outcomes using the r_e^{SE} dataset, a subset whose molecules have available SE equilibrium geometries (see Table 7.1). This allows to compare the computed rotational constants with the B_e derived from the r_e^{SE} (defined as B_e^{SE}). Then, in Section 7.2, the TM-SE(-LR) procedure is tested on: i. the complete dataset; and ii. the species investigated in this thesis. The results are compared with the experimental rotational constants (B_0^{exp}), corrected for the vibrational contributions at the B3 level (therefore giving the semi-experimental equilibrium B_e^{B3} constants).

7.1 Validation on the r_e^{SE} database

At first, the outcomes of the TM-SE(-LR) approach are compared with the r_e^{SE} database. This subset is composed by vinylacetylene (VAC), vinyl cyanide (VC), *trans*-acrolein (*t*-ACR), *cis*-acrolein (*c*-ACR), *trans*-1,3-butadiene (*t*-BD), *trans*-glyoxal (*t*-GL), and (*Z*)-propargylimine (*Z*-PGIM). The list of references used to collect their r_e^{SE} structures is presented in Table 7.1, while all the SE, revDSD and TM-SE(-LR) equilibrium structures are collected in Table A.2. In particular, for *Z*-PGIM, the SE equilibrium structure has been determined in Section A.1 using the data from Refs. 247 and 257.

For the r_e^{SE} subset, the identification of the fragments is straightforward: ethene and acetylene for VAC, ethene and hydrogen cyanide for VC, ethene and formaldehyde for ACR, ethene for *t*-BD, formaldehyde for *t*-GL, and methanimine and acetylene for *Z*-PGIM. Table A.2 shows a very good agreement between the TM-SE(-LR) and r_e^{SE} geometries: deviations are in the order of 1 mÅ for bond lenghts and 0.1° for angles. Few exceptions lie however below the 2 mÅ limit and can be ascribed to conjugative effects. In particular, the revDSD already provides valid results (+3 mÅ and |0.2|° for bonds and angles, respectively), but the improvement brought by the TM-SE(-LR) approach is evident. According to the findings of Table A.2, the accuracy of TM-SE(-LR) can be compared with that from the high-level composite schemes based on CC theory such as the CCSD(T)/CBS+CV one (see Table 4.1). The only uncertain comparison concerns the *Z*-PGIM case, for which a complete determination of the SE equilibrium structure was not possible (*cf.* Section A.1). It is now interesting to address the effect of these small discrepancies on the rotational constants, which are reported in Table 7.2. In fact, the *B*'s are highly sensitive even to slight modifications

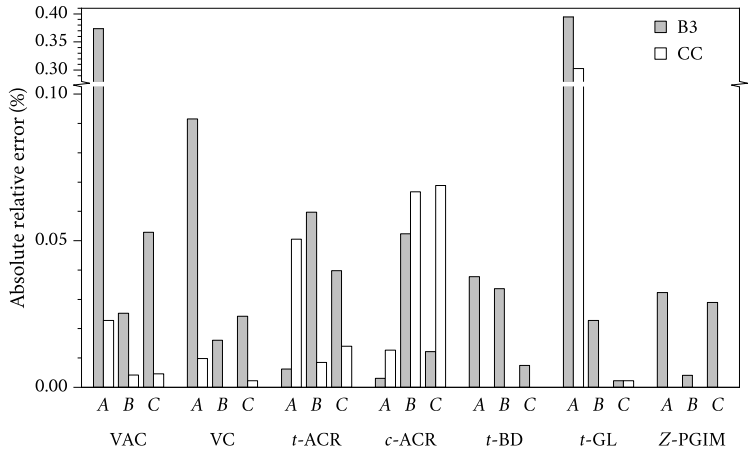


Figure 7.2: Absolute relative error of equilibrium rotational constants derived from the experimental ones and vibrationally corrected at the B3 and fc-CCSD(T)/cc-pVTZ levels of theory (B_e^{B3} and B_e^{CC}, respectively; the latter are available only for VAC, VC, ACR, and *t*-GL) with respect to equilibrium rotational constants directly derived from the SE equilibrium structures (B_e^{SE}).

of the geometry: given the typical values for A of 50 GHz, and 15 GHz for B and C, variations in the order of 0.5 mÅ for bond lenghts and 0.1° for angles may lead to changes up to ±50 MHz on A and ±15 MHz on B and C (an estimated error of 0.1% has been applied, *cf.* Equation 4.1). It is therefore self-evident the importance of accurate structures for the prediction of the rotational constants for a spectroscopically "unknown" molecule.

In Table 7.2, B_e^{SE} refer to the rotational constants directly derived from the SE equilibrium structures, while the B_e^{B3}'s are derived from the experimental rotational constants corrected for the vibrational contributions at the B3 level. Analogously, for some molecules (VAC, VC, ACR and *t*-GL) the fc-CCSD(T)/cc-pVTZ vibrational contributions are available from the literature[240,253,254] and have been used for the determination of the semi-experimental rotational constants, which are denoted as B_e^{CC}'s. The comparison between B_e^{SE}'s and B_e^{B3}'s is of fundamental importance: indeed, when the TM-SE(-LR) procedure is extended to the complete dataset, the latter type of SE rotational constants will be used as the reference value. The outcome of this comparison is also summarized in Figure 7.2. In principle, the two approaches should lead to identical results. However, average discrepancies in the order of 0.04% and 0.06% are noted for the B_e^{CC}'s and the B_e^{B3}'s, respectively. For the B_e^{B3} set, the largest deviations are 0.37% for VAC and 0.40% for *t*-GL, and they are both observed on the A rotational

Table 7.2: Comparison of computed equilibrium rotational constants with semi-experimental counterparts (B_e^{SE}) for the r_e^{SE} subset (namely the underlined species in Figure 7.1). The relative difference with respect to B_e^{SE} is reported in brackets. All values are in MHz.

	revDSD	TM-SE	TM-SE-LR	TM-SE-LR+c[a]	B_e^{SE}	B_e^{B3b}		B_e^{CCc}	
VAC									
A_e	50 611.7 [+0.39%]	50 741.1 [+0.65%]	50 751.6 [+0.67%]	50 552.9 [+0.28%]	50 413.9	50 602.3	*+188.4*	50 425.4	*+11.5*
B_e	4 730.2 [−0.59%]	4 738.2 [−0.42%]	4 745.9 [−0.26%]	4 752.6 [−0.12%]	4 758.2	4 759.4	*+1.2*	4 758.0	*−0.2*
C_e	4 325.9 [−0.50%]	4 333.5 [−0.33%]	4 340.1 [−0.18%]	4 344.2 [−0.08%]	4 347.8	4 350.1	*+2.3*	4 347.6	*−0.2*
VC									
A_e	50 115.1 [+0.35%]	50 250.1 [+0.62%]	50 262.7 [+0.64%]	50 071.7 [+0.27%]	49 939.0	49 893.3	*−45.6*	49 943.9	*+4.9*
B_e	4 955.5 [−0.64%]	4 967.0 [−0.41%]	4 975.5 [−0.24%]	4 982.8 [−0.09%]	4 987.3	4 986.5	*−0.8*	4 987.3	*+0.0*
C_e	4 509.6 [−0.55%]	4 520.2 [−0.31%]	4 527.4 [−0.15%]	4 531.9 [−0.06%]	4 534.4	4 533.3	*−1.1*	4 534.5	*+0.1*
t-ACR									
A_e	47 786.8 [−0.15%]	47 934.4 [+0.15%]	47 996.9 [+0.28%]	47 899.1 [+0.08%]	47 860.9	47 857.9	*−2.1*	47 836.7	*−23.3*
B_e	4 660.4 [−0.58%]	4 667.9 [−0.42%]	4 675.5 [−0.26%]	4 680.3 [−0.16%]	4 687.8	4 685.0	*−2.8*	4 688.2	*+0.4*
C_e	4 246.3 [−0.55%]	4 253.7 [−0.37%]	4 260.5 [−0.21%]	4 263.7 [−0.14%]	4 269.6	4 267.9	*−1.7*	4 270.2	*+0.6*
c-ACR									
A_e	22 938.3 [+0.19%]	23 009.8 [+0.50%]	23 010.0 [+0.50%]	22 971.9 [−0.10%]	22 895.7	22 896.4	*+0.7*	22 892.8	*−2.9*
B_e	6 228.5 [−1.12%]	6 235.5 [−1.01%]	6 250.2 [−0.78%]	6 263.0 [−0.57%]	6 299.1	6 295.8	*−3.2*	6 303.3	*+4.2*
C_e	4 898.4 [−0.83%]	4 906.0 [−0.69%]	4 915.1 [−0.50%]	4 921.3 [−0.38%]	4 940.0	4 939.4	*−0.6*	4 943.4	*+3.4*
t-BD									
A_e	42 069.0 [−0.10%]	42 190.6 [+0.19%]	42 239.2 [+0.30%]	42 123.9 [+0.03%]	42 111.2	42 127.1	*+16.0*		
B_e	4 441.7 [−0.46%]	4 445.0 [−0.38%]	4 451.8 [−0.23%]	4 459.7 [−0.05%]	4 462.1	4 460.6	*−1.4*		
C_e	4 017.5 [−0.42%]	4 021.3 [−0.33%]	4 027.3 [−0.18%]	4 032.7 [−0.05%]	4 034.6	4 034.3	*−0.3*		

(continues)

(Table 7.2)

	revDSD	TM-SE	TM-SE-LR	TM-SE-LR+c[a]	B_e^{SE}	B_e^{B3b}		B_e^{CCc}	
t-GL									
A_e	55 888.5 [−0.51%]	56 092.7 [−0.15%]	56 183.9 [+0.01%]	55 973.2 [−0.36%]	56 177.3	55 955.7	*−221.6*	56 007.1	*−170.2*
B_e	4 781.5 [−0.80%]	4 792.6 [−0.57%]	4 802.1 [−0.37%]	4 811.6 [−0.17%]	4 819.9	4 821.0	*+1.1*	4 819.9	*+0.0*
C_e	4 404.7 [−0.68%]	4 415.4 [−0.53%]	4 423.9 [−0.34%]	4 430.7 [−0.19%]	4 439.1	4 439.0	*−0.1*	4 439.0	*−0.1*
Z-PGIM									
A_e	54 707.3 [−0.57%]	54 894.9 [−0.23%]	54 904.5 [−0.21%]	54 690.0 [−0.61%]	55 022.9	55 040.7	*+17.8*		
B_e	4 850.0 [−0.68%]	4 858.9 [−0.50%]	4 867.2 [−0.33%]	4 873.9 [−0.19%]	4 883.4	4 883.2	*−0.2*		
C_e	4 455.1 [−0.68%]	4 463.8 [−0.48%]	4 470.8 [−0.33%]	4 475.1 [−0.23%]	4 485.5	4 484.2	*−1.3*		

[a] The "+c" refers to a decrease of 0.2° on the linkage angle. [b] Equilibrium rotational constant derived from experimental B_0 and vibrational corrections at the B3 level. The difference ($B_e^{B3} - B_e^{SE}$) is reported in italics. [c] Equilibrium rotational constant derived from experimental B_0 and vibrational corrections at the fc-CCSD(T)/cc-pVTZ level of theory (data from Ref. 253 for VAC and VC, from Ref. 240 for ACR, and from Ref. 254 for *t*-GL). The difference ($B_e^{CC} - B_e^{SE}$) is reported in italics.

constant. The only exception in the B_e^{CC} set is a difference of 0.30% for the A constant of t-GL. By excluding these three outliers, the average deviations reduce to 0.02% for B_e^{CC}'s and 0.03% for B_e^{B3}'s. This comparison provides the order of magnitude of the systematic error arising from the use of the B_e^{B3} rotational constants; a conservative estimate of 0.1% can be derived. Despite B3 vibrational corrections affect (even if marginally) the accuracy of the SE rotational constants, the literature on this topic[169,170,183] demonstrates that their effects are entirely negligible on the structural determination.

Table 7.2 points out small deviations for all test cases when TM-SE(-LR) outcomes are compared with B_e^{SE}'s. In addition, an overall improvement is noticed when moving from revDSD to TM-SE, and, even more, to TM-SE-LR. The absolute mean deviations from the B_e^{SE} are 0.54% at the revDSD level, 0.44% for TM-SE, and 0.33% for TM-SE-LR. As noted above, and confirmed by the literature,[258] the agreement on the rotational constant A is the most critical. In few cases, it worsens when passing from revDSD to TM-SE, and the LR contribution is unable to correct the discrepancy.

By inspecting in detail the structural parameters, this phenomenon seems to be ascribable to the missing LR correction for the linkage angle.* This may affect the overall precision of the spectral prevision of an "unknown" molecule; however, once the latter is analyzed and a set of rotational constants is derived, highly-accurate equilibrium geometries may be determined thanks to LSF procedures on the linkage parameters.

Given these findings, tests have been carried out to understand if a systematic correction may be applied to the linkage angle: the outcomes are collected in the TM-SE-LR+c column of Table 7.2. It appears that lowering the linkage angle by 0.2° improves the agreement with B_e^{SE}, with the exceptions being the A constants of t-GL and Z-PGIM. The mean deviation decreases to 0.2%. This suggests that the LR correction is able to recover only partially for the conjugative effects due to double and triple bonds connected through the C–C bond. While the "+c" term (*i.e.*, the decrease of the linkage angle by 0.2°) works well for most imines, the extension to the entire dataset showed improvements only for molecules containing the vinyl moiety (*cf.* Table 7.3).

*LR corrections for angles are available, but inter-fragments angles tend to behave differently.

7.2 Application and performance

Having evaluated the impact of using the B_e^{B3} with respect to the B_e^{SE}, the TM-SE(-LR) method has been extended to the complete dataset represented in Figure 7.1: the results are collected in Table 7.3. As mentioned above, the revDSD, TM-SE, and TM-SE-LR(+c) equilibrium rotational constants are compared with the B_e^{B3} ones.

The results are in line with those reported in Table 7.2. In almost all cases, the revDSD level provides valid results, with the mean deviation within 1%. A noticeable improvement is found when the TM-SE method is used (the error lowers to 0.33%), with a further reduction of the discrepancy provided by the introduction of the LR correction, leading to an average relative deviation of 0.24%. In particular, the TM-SE approach is found to perform remarkably well on rigid molecules, namely those containing the benzene moiety, for which the average discrepancy of the TM-SE and TM-SE-LR equilibrium constants with respect to the B_e^{B3}'s is 0.18% and 0.11%, respectively.

As anticipated above, the decrease of the linkage angle by 0.2° (the "+c" term) is significant only for the species bearing the vinyl moiety (*i.e.*, styrene, vinylcyanide, vinylacetylene, acrolein, allylimine, and *trans*-butadiene). Indeed, for these molecules, the TM-SE-LR+c equilibrium rotational constants show an average discrepancy of 0.14% with respect to the B_e^{B3} ones: a noticeable improvement if compared with the average deviations of the TM-SE and TM-SE-LR methods, which are 0.36% and 0.28%, respectively. Instead, on average, the introduction of the "+c" correction for the remaining species worsens the outcome (from 0.22% of the TM-SE-LR method to 0.28% of the TM-SE-LR+c). Given that, the "+c" correction will not be used in the application of the TM-SE(-LR) method to the species investigated in this thesis.

Based on the systematic error of 0.1% on the B_e^{B3} estimated above, the relative error on the TM-SE-LR rotational constants can be as small as 0.1%. This implies a high accuracy for the $r_e^{TM\text{-}SE\text{-}LR}$ structures, whose mean average deviations are thus expected to be below 1 mÅ on bond lengths and 0.1° on angles, an outstanding result in line with the CCSD(T)/CBS+CV performance (see Table 4.1), retaining the cost-effectiveness of the revDSD level of theory.

Table 7.3: Comparison of computed equilibrium rotational constants with B3 vibrationally corrected experimental counterparts (B_e^{B3}) for the entire dataset. The relative difference with respect to B_e^{B3} is reported in brackets. All values are in MHz.

	revDSD	TM-SE	TM-SE-LR	TM-SE-LR+c[a]	B_e^{B3b}	B_0^{exp}	$\Delta B_{0,\mathrm{vib}}^{\mathrm{B3c}}$
Z-PGIM							
A	54 707.3 [−0.61%]	54 894.9 [−0.27%]	54 904.5 [−0.25%]	54 690.0 [−0.64%]	55 040.8	54 640.1468(45)	400.6
B	4 850.0 [−0.68%]	4 858.9 [−0.50%]	4 867.2 [−0.33%]	4 873.9 [−0.19%]	4 883.2	4 862.362758(60)	20.9
C	4 455.1 [−0.67%]	4 463.8 [−0.48%]	4 470.8 [−0.32%]	4 475.1 [−0.23%]	4 485.2	4 458.249970(55)	27.0
E-PGIM							
A	63 274.6 [−0.32%]	63 397.5 [−0.12%]	63 412.3 [−0.10%]	63 134.1 [−0.54%]	63 476.8	63 099.2207(22)	377.6
B	4 752.0 [−0.73%]	4 761.0 [−0.55%]	4 768.9 [−0.38%]	4 775.2 [−0.25%]	4 787.1	4 766.557614(55)	20.5
C	4 420.0 [−0.70%]	4 428.4 [−0.51%]	4 435.4 [−0.36%]	4 439.4 [−0.27%]	4 451.3	4 425.560983(58)	25.7
Z-CMI							
A	54 326.8 [−0.58%]	54 527.8 [−0.21%]	54 537.2 [−0.20%]	54 330.0 [−0.57%]	54 644.0	54 193.405(32)	450.6
B	5 057.5 [−0.75%]	5 070.7 [−0.49%]	5 079.3 [−0.32%]	5 086.7 [−0.18%]	5 095.8	5 073.86584(15)	22.0
C	4 626.7 [−0.74%]	4 639.3 [−0.47%]	4 646.6 [−0.31%]	4 651.2 [−0.21%]	4 661.0	4 632.38905(14)	28.7
E-CMI							
A	62 939.3 [−0.26%]	63 084.0 [−0.03%]	63 100.1 [+0.00%]	62 830.2 [−0.43%]	63 101.5	62 700.392(22)	401.1
B	4 951.0 [−0.88%]	4 966.8 [−0.56%]	4 975.1 [−0.39%]	4 982.0 [−0.26%]	4 994.8	4 972.04534(22)	22.7
C	4 590.0 [−0.83%]	4 604.3 [−0.52%]	4 611.5 [−0.36%]	4 615.9 [−0.27%]	4 628.3	4 600.29561(23)	28.0
Z-PMI							
A	5 229.2 [−0.26%]	5 247.6 [+0.09%]	5 247.8 [+0.09%]	5 246.0 [+0.06%]	5 242.9	5 200.81278(16)	42.0
B	1 550.6 [−0.53%]	1 554.0 [−0.31%]	1 556.0 [−0.18%]	1 557.2 [−0.11%]	1 558.8	1 548.969349(92)	9.9
C	1 196.0 [−0.50%]	1 198.9 [−0.25%]	1 200.1 [−0.15%]	1 200.8 [−0.10%]	1 201.9	1 194.842313(78)	7.1

(continues)

(Table 7.3)

	revDSD	TM-SE	TM-SE-LR	TM-SE-LR+c[a]	B_e^{B3b}	B_0^{exp}	$\Delta B_{0,\mathrm{vib}}^{\mathrm{B3c}}$
E-PMI							
A	5 241.9 [−0.30%]	5 260.2 [+0.05%]	5 260.3 [+0.05%]	5 258.5 [+0.02%]	5 257.5	5 217.29202(11)	40.2
B	1 567.1 [−0.49%]	1 570.4 [−0.28%]	1 572.3 [−0.16%]	1 573.6 [−0.08%]	1 574.8	1 565.283633(28)	9.5
C	1 206.4 [−0.46%]	1 209.4 [−0.21%]	1 210.5 [−0.12%]	1 211.2 [−0.07%]	1 212.0	1 204.540307(14)	7.4
CF							
A	66 979.5 [−0.87%]	67 294.3 [−0.40%]	67 313.0 [−0.38%]	67 018.0 [−0.81%]	67 566.7	67 469.6749(29)	97.0
B	4 989.9 [−0.69%]	5 004.8 [−0.39%]	5 014.4 [−0.20%]	5 021.3 [−0.06%]	5 024.5	5 010.18856(34)	14.3
C	4 643.9 [−0.70%]	4 658.3 [−0.39%]	4 666.7 [−0.21%]	4 671.3 [−0.11%]	4 676.6	4 656.60175(29)	20.0
PGCN							
A	19 652.8 [−0.99%]	19 736.2 [−0.57%]	19 751.8 [−0.49%]	19 662.4 [−0.94%]	19 848.6	19 820.080(70)	28.6
B	2 902.8 [−0.36%]	2 912.4 [−0.03%]	2 916.1 [+0.09%]	2 922.8 [+0.32%]	2 913.4	2 909.6062(12)	3.8
C	2 569.3 [−0.43%]	2 578.1 [−0.10%]	2 581.2 [+0.03%]	2 584.9 [+0.17%]	2 580.5	2 573.2123(12)	7.3
VC							
A	50 115.1 [+0.44%]	50 250.1 [+0.72%]	50 262.7 [+0.74%]	50 071.7 [+0.36%]	49 893.3	49 850.6974(20)	42.6
B	4 955.5 [−0.62%]	4 967.0 [−0.39%]	4 975.5 [−0.22%]	4 982.8 [−0.07%]	4 986.5	4 971.163651(24)	15.3
C	4 509.6 [−0.52%]	4 520.2 [−0.29%]	4 527.4 [−0.13%]	4 531.9 [−0.03%]	4 533.3	4 513.877260(25)	19.4
MN							
A	20 674.4 [−1.16%]	20 746.9 [−0.81%]	20 784.8 [−0.63%]	20 690.6 [−1.08%]	20 916.4	20 882.77323(53)	33.6
B	2 934.7 [−0.38%]	2 942.2 [−0.12%]	2 949.7 [+0.13%]	2 956.4 [+0.36%]	2 945.8	2 942.304967(75)	3.5
C	2 611.6 [−0.47%]	2 618.5 [−0.20%]	2 625.1 [+0.05%]	2 628.9 [+0.19%]	2 623.8	2 616.727165(66)	7.1
E-AIM							
A	46 182.1 [−0.15%]	46 286.3 [+0.07%]	46 341.1 [+0.19%]	46 254.9 [+0.01%]	46 252.4	45 773.628(18)	478.8

(continues)

(Table 7.3)

	revDSD	TM-SE	TM-SE-LR	TM-SE-LR+c[a]	B_e^{B3b}	B_0^{exp}	$\Delta B_{0,\mathrm{vib}}^{\mathrm{B3c}}$
B	4 568.6 [−0.42%]	4 573.0 [−0.32%]	4 580.0 [−0.17%]	4 584.3 [−0.07%]	4 587.8	4 560.916(4)	26.8
C	4 157.3 [−0.41%]	4 161.8 [−0.30%]	4 168.1 [−0.15%]	4 170.9 [−0.08%]	4 174.4	4 148.242(3)	26.1
Z-AIM							
A	44 166.3 [−0.11%]	44 307.8 [+0.21%]	44 361.7 [+0.33%]	44 286.8 [+0.16%]	44 215.4	43 759.52(18)	455.9
B	4 573.3 [−0.45%]	4 577.6 [−0.36%]	4 584.5 [−0.21%]	4 588.8 [−0.11%]	4 593.9	4 564.581(39)	29.4
C	4 144.1 [−0.43%]	4 148.9 [−0.32%]	4 155.1 [−0.17%]	4 157.9 [−0.10%]	4 162.1	4 134.423(30)	27.6
STY							
A	5 195.5 [−0.20%]	5 213.8 [+0.15%]	5 214.0 [+0.16%]	5 211.9 [+0.12%]	5 205.8	5 163.385(11)	42.4
B	1 546.9 [−0.53%]	1 550.1 [−0.32%]	1 552.1 [−0.20%]	1 553.4 [−0.11%]	1 555.1	1 545.1699(7)	9.9
C	1 192.0 [−0.52%]	1 194.9 [−0.28%]	1 196.0 [−0.19%]	1 196.7 [−0.13%]	1 198.3	1 191.2240(7)	7.1
VAC							
A	50 611.7 [+0.02%]	50 741.1 [+0.27%]	50 751.6 [+0.30%]	50 552.9 [−0.10%]	50 602.3	50 300.158(55)	302.2
B	4 730.2 [−0.61%]	4 738.2 [−0.45%]	4 745.9 [−0.28%]	4 752.6 [−0.14%]	4 759.4	4 744.94254(22)	14.4
C	4 325.9 [−0.56%]	4 333.5 [−0.38%]	4 340.1 [−0.23%]	4 344.2 [−0.14%]	4 350.1	4 329.77304(23)	20.3
BAL							
A	5 256.4 [−0.31%]	5 275.8 [+0.05%]	5 276.0 [+0.06%]	5 274.1 [+0.02%]	5 273.0	5 234.364365(54)	38.6
B	1 565.3 [−0.53%]	1 569.1 [−0.29%]	1 571.1 [−0.16%]	1 572.4 [−0.08%]	1 573.7	1 564.274377(22)	9.4
C	1 206.1 [−0.49%]	1 209.4 [−0.22%]	1 210.6 [−0.12%]	1 211.3 [−0.07%]	1 212.1	1 204.681915(16)	7.4
PA							
A	67 635.2 [−0.65%]	67 933.1 [−0.21%]	67 949.0 [−0.18%]	67 641.0 [−0.64%]	68 074.6	68 035.2557(13)	39.4
B	4 808.4 [−0.65%]	4 819.4 [−0.43%]	4 828.3 [−0.24%]	4 834.6 [−0.11%]	4 840.1	4 826.22365(10)	13.8
C	4 489.2 [−0.65%]	4 500.1 [−0.41%]	4 507.9 [−0.24%]	4 512.1 [−0.15%]	4 518.8	4 499.592240(99)	19.2

(continues)

(Table 7.3)

	revDSD	TM-SE	TM-SE-LR	TM-SE-LR+c[a]	B_e^{B3b}	B_0^{exp}	$\Delta B_{0,vib}^{B3c}$
t-ACR							
A	47 786.8 [−0.15%]	47 934.4 [+0.16%]	47 996.9 [+0.29%]	47 899.1 [+0.09%]	47 857.9	47 353.729(9)	504.1
B	4 660.4 [−0.53%]	4 667.9 [−0.37%]	4 675.5 [−0.20%]	4 680.3 [−0.10%]	4 685.0	4 659.4894(4)	25.6
C	4 246.3 [−0.51%]	4 253.7 [−0.33%]	4 260.5 [−0.18%]	4 263.7 [−0.10%]	4 267.9	4 242.7034(4)	25.2
c-ACR							
A	22 938.3 [+0.18%]	23 009.8 [+0.50%]	23 010.0 [+0.50%]	22 971.9 [+0.33%]	22 896.4	22 831.650(16)	64.7
B	6 228.5 [−1.08%]	6 235.5 [−0.97%]	6 250.2 [−0.73%]	6 263.0 [−0.53%]	6 295.8	6 241.0470(29)	54.7
C	4 898.4 [−0.84%]	4 906.0 [−0.68%]	4 915.1 [−0.50%]	4 921.3 [−0.37%]	4 939.4	4 902.2063(28)	37.2
BN[d]							
A	5 678.3 [−0.34%]	5 698.9 [+0.02%]	5 698.9 [+0.02%]		5 697.8	5 655.26522(59)	42.5
B	1 546.7 [−0.39%]	1 551.4 [−0.09%]	1 553.1 [+0.02%]		1 552.8	1 546.875836(63)	5.9
C	1 215.6 [−0.38%]	1 219.4 [−0.07%]	1 220.5 [+0.02%]		1 220.3	1 214.404061(48)	5.9
t-BD							
A	42 069.0 [−0.14%]	42 190.6 [+0.15%]	42 239.2 [+0.27%]	42 123.9 [−0.01%]	42 127.1	41 682.6577(21)	444.5
B	4 441.7 [−0.43%]	4 445.0 [−0.35%]	4 451.8 [−0.20%]	4 459.7 [−0.02%]	4 460.6	4 433.5047(30)	27.1
C	4 017.5 [−0.42%]	4 021.3 [−0.32%]	4 027.3 [−0.17%]	4 032.7 [−0.04%]	4 034.3	4 008.0423(60)	26.2
t-GL							
A	55 888.5 [−0.12%]	56 092.7 [+0.25%]	56 183.9 [+0.41%]	55 973.2 [+0.03%]	55 955.7	55 290.6121(510)	665.1
B	4 781.5 [−0.82%]	4 792.6 [−0.59%]	4 802.1 [−0.40%]	4 811.6 [−0.20%]	4 821.0	4 798.0371(69)	23.0
C	4 404.7 [−0.78%]	4 415.4 [−0.54%]	4 423.9 [−0.34%]	4 430.7 [−0.19%]	4 439.0	4 416.8983(69)	22.2

[a] The "+c" refers to a decrease of 0.2° on the linkage angle. [b] Equilibrium rotational constant derived from experimental B_0^{exp} and vibrational corrections at the B3 level. [c] Vibrational contributions at the B3 level. [d] Given the symmetry point group of benzonitrile (C_{2v}), the "+c" term cannot be applied without breaking the molecular symmetry and thus it has been ignored.

Application to the investigated species

Finally, the TM-SE(-LR) method has been applied to the species investigated in this thesis, namely (*E*)- and (*Z*)-cyanovinylacetylene, allenylacetyelene, 2- and 3-furonitrile, and 3-aminoisoxazole, in addition to the isomers of phenylmethanimine, which have already been discussed in Table 7.3. This approach has not been applied for iso- and n-propylamine, due to their high flexibility (as demonstrated above, this approach is especially suited for rather rigid molecules).

Like those of the dataset reported in Figure 7.1, the molecules of this set can be seen as formed by two (or three) distinct fragments: acetylene, ethene and hydrogen cyanide for cyanovinylacetylene, allene and acetylene for allenylacetyelene, furan and hydrogen cyanide for the two furonitrile isomers, and isoxazole and ammonia for 3-aminoisoxazole. In particular, cyanovinylacetylene can also be seen as the union of vinylcyanide and acetylene as well as vinylacetylene and hydrogen cyanide. Most of the listed fragments have already been used in the formation of the molecules of the dataset of Figure 7.1, and only a few ones need to be added: allene ($H_2C_3H_2$), furan (C_4H_4O), isoxazole (C_3H_3NO), and ammonia (NH_3). As explained in detail above, the TM-SE(-LR) method relies on the availability of an accurate structural determination for a fragment of the target molecule. For the two heterocycles and for ammonia, the r_e^{SE} structure is available in literature (from the SE database of Refs. 169–172 for furan and ammonia, and from Ref. 187 for isoxazole), while it has been determined in this work for allene (using the data collected in Ref. 259, see Section A.1). For all the investigated species, the low-cost method (see again Equations 4.3 and 4.5) is the revDSD one, as defined above. All the r_e^{SE} structures of the fragments can be found in Tables A.1, A.2, and A.3.

The results of the TM-SE(-LR) approach for the species investigated in this thesis are collected in Table 7.4, while the molecular structures are reported in the corresponding Tables in Section A.3. Similarly to Table 7.3, the outcomes are compared with the B_e^{B3} derived from ground-state experimental rotational constants B_0^{exp} (which have been determined by means of the laboratory works detailed in Chapters 8–12) corrected for vibrational contributions at the B3 level ($\Delta B_{0,\mathrm{vib}}^{\mathrm{B3}}$).

At this point, a note on cyanovinylacetylene isomers is deserved. As mentioned above, the former can be seen as sum of different fragments: i. acetylene, ethene, and hydrogen cyanide; ii. vinylacetylene and hydrogen cyanide; or iii. vinylcyanide and acetylene. In this work, all three options have been tested, and the results are not significantly different. Indeed, for the TM-SE approach, the average discrepancy between the $B_e^{\mathrm{TM\text{-}SE}}$

Table 7.4: Comparison of computed equilibrium rotational constants with B3 vibrationally corrected experimental counterparts for the species investigated in this thesis. The relative difference with respect to B_e^{B3} is reported in brackets. All values are in MHz.

	revDSD	TM-SE	TM-SE-LR	B_e^{B3a}	B_0^{exp}	$\Delta B_{0,\mathrm{vib}}^{\mathrm{B3b}}$
(*E*)-cyanovinylacetylene[c]						
A	46 602.3 [+0.52%]	46 725.6 [+0.79%]	46 730.0 [+0.80%]	46 360.9	46 248.0(2)	112.9
B	1 467.4 [−0.56%]	1 470.0 [−0.38%]	1 472.9 [−0.19%]	1 475.6	1 472.15368(4)	3.5
C	1 422.6 [−0.52%]	1 425.2 [−0.34%]	1 427.9 [−0.16%]	1 430.1	1 426.24867(4)	3.9
(*Z*)-cyanovinylacetylene[c]						
A	7 124.0 [+0.13%]	7 142.9 [+0.39%]	7 156.8 [+0.59%]	7 115.1	7 098.2006(2)	16.9
B	2 651.6 [−1.05%]	2 656.0 [−0.89%]	2 659.8 [−0.75%]	2 679.8	2 682.59013(6)	−2.8
C	1 932.4 [−0.71%]	1 936.1 [−0.52%]	1 939.1 [−0.37%]	1 946.3	1 943.54387(5)	2.7
allenylacetylene						
A	25 809.3 [−0.75%]	25 885.6 [−0.45%]	25 904.7 [−0.38%]	26 003.4	25 963.71(9)	39.7
B	2 611.5 [−0.39%]	2 615.4 [−0.25%]	2 618.3 [−0.13%]	2 621.8	2 616.3763(1)	5.5
C	2 411.0 [−0.40%]	2 414.8 [−0.25%]	2 417.5 [−0.13%]	2 420.7	2 412.5732(1)	8.1
2-furonitrile						
A	9 247.9 [−0.44%]	9 299.2 [+0.11%]	9 299.2 [+0.11%]	9 288.5	9 220.2506(1)	68.3
B	2 026.3 [−0.50%]	2 032.4 [−0.20%]	2 034.9 [−0.08%]	2 036.5	2 029.27372(2)	7.2
C	1 662.1 [−0.49%]	1 667.9 [−0.15%]	1 669.6 [−0.04%]	1 670.3	1 662.64314(2)	7.7
3-furonitrile						
A	9 330.7 [−0.44%]	9 365.0 [−0.08%]	9 365.0 [−0.08%]	9 372.4	9 296.5489(1)	75.9
B	1 937.8 [−0.44%]	1 944.8 [−0.09%]	1 947.2 [+0.03%]	1 946.5	1 940.26649(2)	6.2
C	1 604.6 [−0.45%]	1 610.3 [−0.09%]	1 612.0 [+0.01%]	1 611.8	1 604.63205(2)	7.2

(continues)

(Table 7.4)

	revDSD	TM-SE	TM-SE-LR	B_e^{B3a}	B_0^{exp}	$\Delta B_{0,\mathrm{vib}}^{\mathrm{B3b}}$
3-aminoisoxazole						
A	9 387.2 [−0.43%]	9 424.6 [−0.03%]	9 424.6 [−0.03%]	9 427.7	9 356.3832(1)	71.3
B	3 677.6 [−0.48%]	3 684.3 [−0.30%]	3 688.7 [−0.18%]	3 695.5	3 669.78547(7)	25.7
C	2 651.1 [−0.44%]	2 657.5 [−0.20%]	2 659.8 [−0.12%]	2 662.9	2 642.30281(5)	20.6
Z-phenylmethanimine						
A	5 229.2 [−0.26%]	5 247.6 [+0.09%]	5 247.8 [+0.09%]	5 242.9	5 200.81278(16)	42.0
B	1 550.6 [−0.53%]	1 554.0 [−0.31%]	1 556.0 [−0.18%]	1 558.8	1 548.969349(92)	9.9
C	1 196.0 [−0.50%]	1 198.9 [−0.25%]	1 200.1 [−0.15%]	1 201.9	1 194.842313(78)	7.1
E-phenylmethanimine						
A	5 241.9 [−0.30%]	5 260.2 [+0.05%]	5 260.3 [+0.05%]	5 257.5	5 217.29202(11)	40.2
B	1 567.1 [−0.49%]	1 570.4 [−0.28%]	1 572.3 [−0.16%]	1 574.8	1 565.283633(28)	9.5
C	1 206.4 [−0.46%]	1 209.4 [−0.21%]	1 210.5 [−0.12%]	1 212.0	1 204.540307(14)	7.4

[a] Equilibrium rotational constant derived from experimental B_0^{exp} and vibrational corrections at the B3 level. [b] Vibrational contributions at the B3 level. [c] The fragments used for the TM-SE approach are acetylene, ethene, and hydrogen cyanide.

and the B_e^{B3} varies by 0.06% at most when changing the set of fragments. For the TM-SE-LR approach, the results are even closer one to the other (0.02%). Therefore, given the uncertainty due to the use of B_e^{B3} rotational constants of 0.1% demonstrated above, it can be derived that the choice of the fragments for cyanovinylacetylene isomers does not affect the accuracy of the $B_e^{\mathrm{TM\text{-}SE(\text{-}LR)}}$ values.

From the inspection of Table 7.4, it is noted that the results can be classified in two groups. The first group includes the isomers of cyanovinylacetylene and allenylacetyelene, while the second group collects the remaining species (2- and 3-furonitrile and 3-aminoisoxazole). For the sake of completeness, the results of phenylmethanimine isomers are also included in the second group as well as in Table 7.4. For the molecules belonging to the first set, only a small improvement is found when moving from the revDSD equilibrium rotational constants to the TM-SE ones (on average, 0.09%), and to TM-SE-LR structures (0.17%). As a matter of fact, all three species present a comparable planar (or almost planar) structure, formed by linear (or strongly prolate-shaped, like ethene) unsaturated moieties connected by a single bond, therefore retaining a certain degree of flexibility, in particular in the linkage angle.

The results are indeed different for the second group, which collects molecules sharing a mono-substituted aromatic ring. For these species, given an average discrepancy between the revDSD equilibrium rotational constants and B_e^{B3} of 0.44%, a relevant improvement (the average discrepancy reduces to 0.16%) is observed when the TM-SE is applied. Moreover, the error lowers to 0.10% if LR contribution is included. In conclusion, by taking into account the systematic error affecting this analysis, these results show the reliability of the TM-SE(-LR) approach for unsaturated systems, and a superb performance for rather rigid molecules.

In order to understand the effect of the improvement given by the TM-SE(-LR) approach in the prediction of rotational spectra, Figure 7.3 shows a comparison between different simulations and the exact frequency of a transition for four different species, namely (*Z*)- and (*E*)-phenylmethanimine, 3-aminoisoxazole, and 3-furonitrile. These species have been selected among the group of investigated molecules due to their lacking of a previous spectroscopic characterization: the accurate prediction of line positions in the cm-wave frequency range is of great help in the first phases of the experiment. To simulate the typical laboratory course of action at best, a strong transition with a low K_a value has been selected for each species, in the optimal frequency range for the COBRA-FTMW spectrometer.

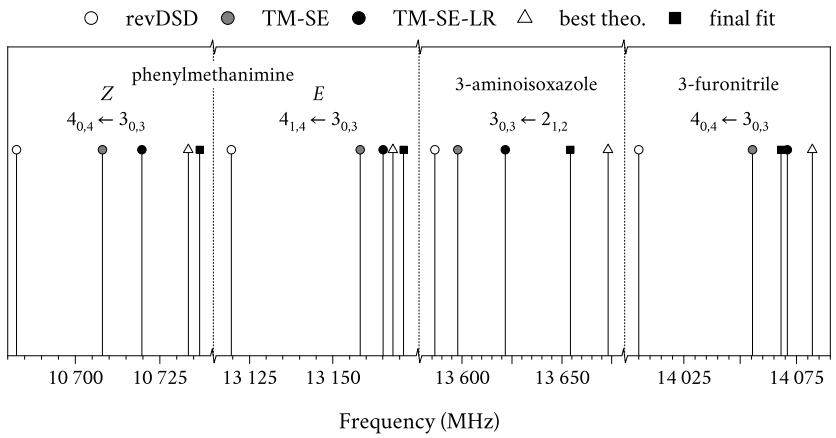


Figure 7.3: Comparison between different simulated rotational transition frequencies (predicted using equilibrium rotational constants at the revDSD, TM-SE, and TM-SE-LR, augmented by B3 vibrational corrections, and "best theoretical" levels, see text; quartic centrifugal distortion constants in all cases at the "best theoretical" level) and the line position derived from the experimental spectroscopic parameters (any hyperfine structure is neglected). For the "best theoretical" level definition and the experimental spectroscopic parameters, the reader is referred to the corresponding Chapters.

In Figure 7.3, the simulations are built using: i. the revDSD, TM-SE, and TM-SE-LR equilibrium rotational constants augmented by B3 vibrational corrections; and ii. the "best theoretical" level. In particular, they all include the computed quartic centrifugal distortion constants evaluated at the "best theoretical" level, which is based on the "cheap" composite scheme, or its variants, for the equilibrium rotational constants, and on (double-)hybrid DFT functionals for the vibrational corrections and the quartic centrifugal distortion terms.

A note is warranted for what concerns the definition of the exact frequency for the transitions shown in Figure 7.3. The value has been determined using the experimental spectroscopic parameters determined from the analysis of the rotational spectra investigated in the corresponding Chapter. For the sake of clarity, the hyperfine structure has been neglected.

Being the "final fit" values (represented by black squares) the target, Figure 7.3 points out a clear improvement when moving from revDSD (white circles) to TM-SE (gray circles), and, further, to TM-SE-LR (black circles). Indeed, on average, the revDSD transitions lie *ca.* 60 MHz away from the exact position; this error halves when the TM-SE approach is applied, and is reduced to a quarter when the LR correction is added.

Furthermore, the average error associated with the line position for the TM-SE-LR (*ca.* 15 MHz) is of the same order of magnitude as the one achievable with the "best theoretical" one (*ca.* 10 MHz, indicated by white triangles in Figure 7.3). Concluding, it is necessary to recall that the TM-SE-LR approach has the same computational cost of the "low-cost" method chosen in the formulation of Equations (4.3) and (4.5), and its performance is comparable with CC-based composite schemes.

Chapter 8

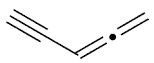
Allenylacetylene

The astrochemical relevance of allenylacetylene (or 3,4-pentadiene-1-yne, shown in Figure 8.1) has been introduced in Section 1.3. In this Chapter,* the outcomes of a detailed theoretical and experimental characterization of this unsaturated pure hydrocarbon are presented.

8.1 Computational characterization

An accurate computational characterization of the spectroscopic parameters of allenylacetylene has been carried out in order to support the recording of its rotational spectrum at higher frequency. Indeed, the previous studies[75,220] included only a partial set of quartic centrifugal distortion terms.

In more detail, the equilibrium rotational structure of allenylacetylene has been determined using the jun-ChS model, while the B2PLYP-D3(BJ)/jun-cc-pVTZ level of theory (hereafter, jB2) has been chosen for the vibrational corrections. The latter level of theory has also been used for the determination of quartic and sextic centrifugal distortion terms, as well as electric dipole moment components. All optimized geometries are



allenylacetylene

Figure 8.1: Molecular structure of allenylacetylene.

*The contents of this Chapter are taken from Ref. 260 after re-elaboration.

collected in Tables A.4, while the spectroscopic parameters are collected in Table 8.1.

8.2 Experimental study

Allenylacetylene (Figure 8.1) is a pure hydrocarbon belonging to the C_s point group, with the two hydrogen atoms of the allene moiety laying outside the symmetry plan. From a spectroscopic point of view, it is a nearly prolate asymmetric-top ($\kappa = -0.98$) with only one significant electric dipole component ($\mu_a = 0.72$ D). Rotational transitions of allenylacetylene were identified by McCarthy *et al.*[220] in the plasma-induced decomposition of benzene using FMTW spectroscopy. Then, 19 lines were detected in TMC-1 by Chernicharo *et al.*[75] in the 31–50 GHz range, thus obtaining improved rotational constants and three quartic centrifugal distortion parameters, which are collected in the "previous results" column of Table 8.1.

In this work, the rotational spectrum of allenylacetylene has been recorded using the FM-mmW spectrometer (*cf.* Section 6.2); the laboratory conditions are summarized in Section 6.4. Thanks to the availability of the spectroscopic parameters, the assignment of the rotational transitions in the 75–115 GHz range in the FM-mmW experiment has been straightforward. However, from 225 GHz, the line positions have been found to be rather inaccurate: deviations as large as *ca.* 3 MHz have been found for $K_a = 0$ transitions, with the discrepancy rapidly increasing with J. Moreover, the centrifugal distortion description was not improved by the inclusion of these lines in the spectral analysis, as no transitions with

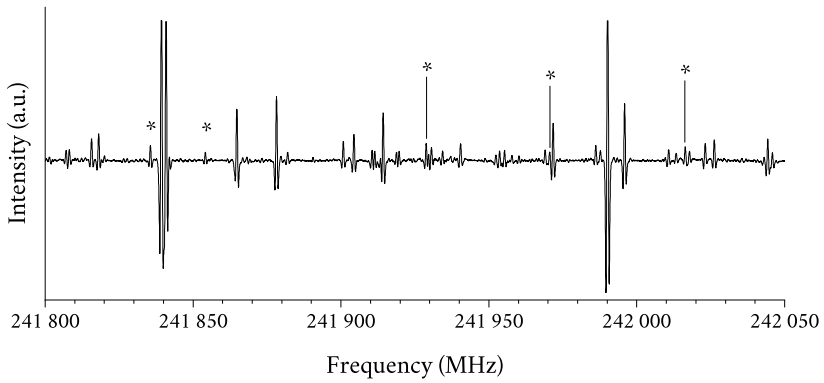


Figure 8.2: Portion of the FM-mmW experimental spectrum of allenylacetylene. Transitions with an asterisk have been included in the fit.

$K_a \geq 1$ could be unambiguously assigned. To overcome this problem, the accuracy of the prediction need to be addressed. Indeed, in addition to the full set of computed sextic centrifugal distortion constants, three octic terms have been used; the latter have been experimentally determined for cyanoallene,[261,262] which is isoelectronic to allenylacetylene. This led to the identification of almost 200 lines in the 225–315 GHz range, with the maximum value of K_a equal to 13.

The rotational Hamiltonian is reported in Equation (2.58). The fit procedure included all the literature data, with an uncertainty of 2 kHz and 10–20 kHz being assigned to the previous FTMW[220] and astronomical[75] experiments, respectively. For the FM-mmW data, an average uncertainty of 35 kHz has been assumed, given the low S/N and the presence of strong interfering lines. Indeed, a portion of the experimental spectrum has been reported in Figure 8.2, from which it is evident the low intensity of allenylacetylene transitions. Despite the FVP conditions have been optimized, the resulting spectrum has been characterized by a global weakness with respect to other rotational features from pyrolysis co-products.

The resulting spectroscopic parameters are reported in Table 8.1, together with the theoretical and previous values. From the comparison with the literature results, a significant improvement is noted for all rotational constants, and in particular for A, whose absolute error is almost one order of magnitude smaller. A complete set of quartic, four sextic and one octic centrifugal distortion terms have been determined. Conversely, H_J, h_2, and h_3 could not be fitted, and their values have been constrained to the theoretical counterparts. Their inclusion in the fit procedure does not significantly affect the outcomes; it is also noted that cyanoallene centrifugal distortion constants are no longer needed in the final analysis.

The fit statistics show a rms error close to 30 kHz, with a standard deviation (σ) below unity: the derived spectroscopic parameters are able to reproduce the rotational transitions within their experimental uncertainty. Finally, the comparison between the theoretical and experimental data points out a good agreement: the average deviation on the rotational constants is close to 0.3%, while discrepancies of *ca.* 9% are found on the quartic centrifugal distortion terms (this being mostly related to the small magnitude of these parameters).

Table 8.1: Ground-state spectroscopic parameters of allenylacetylene (*S*-reduction, I^r representation). Values in parentheses denote one standard deviation and apply to the last digit of the constants.

		allenylacetylene		
		exp.[a]	theo.[b]	prev.[c]
A	MHz	25 963.71(9)	25 806.9	25 961.2(8)
B	MHz	2 616.3763(1)	2 615.1	2 616.3762(2)
C	MHz	2 412.5732(1)	2 410.4	2 412.5733(2)
D_J	kHz	1.14788(4)	1.09	1.157(1)
D_{JK}	kHz	−85.469(5)	−81.8	−85.49(3)
D_K	MHz	2.158(5)	2.05	
d_1	kHz	−0.28715(4)	−0.267	−0.286(1)
d_2	Hz	−8.00(2)	−6.10	
H_J	mHz	3.950(6)	3.27	
H_{JK}	Hz	−0.1645(4)	−0.103	
H_{KJ}	Hz	−14.34(9)	−15.0	
H_K	kHz		0.505	
h_1	mHz	1.568(8)	1.29	
h_2	µHz		69.4	
h_3	µHz		17.5	
L_{KKJ}	mHz	5.1(4)		
$\lvert\mu_a\rvert$	D	Y[d]	0.72	
$\lvert\mu_b\rvert$	D	N	0.01	
# lines		313		33
J^{max}, K_a^{max}		64, 13		10, 3
rms error	kHz	30.8		7.7
σ		0.92		0.59

[a] Missing values have been kept fixed at the theoretical counterpart in the fit procedure. [b] Jun-ChS equilibrium rotational constants augmented by vibrational corrections at the jB2 level. Quartic and sextic centrifugal distortion constants at the jB2 level of theory. [c] Re-fitted using the *S*-reduction from Ref. 75. [d] For each μ_g dipole moment component, "Y" and "N" refer to detected and non-detected *g*-type transitions, respectively.

Chapter 9

(*E*)- and (*Z*)-Cyanovinylacetylene

The astrochemical importance of the cyanovinylacetylene (or 2-penten-4-ynenitrile, represented in Figure 9.1) isomers is reported in Section 1.3. In this Chapter,* a comprehensive investigation of the ground-state rotational spectra of the cyanovinylacetylene isomers is reported, complemented by high-level quantum-chemical computations and the energetic characterization of the $[H_3C_5N]$ family of isomers.

9.1 Computational characterization

Prior to this work, a complete spectroscopic characterization for cyanovinylacetylene isomers (see Figure 9.1) was not available in literature, with some crucial spectroscopic parameters still missing. Indeed, the centrifugal distortion characterization was limited to the quartic terms: the inclusion of high-order terms can be of great help in the spectral analysis.

Therefore, to guide and support the experiment, high-level quantum

(*E*)-cyanovinylacetylene (*Z*)-cyanovinylacetylene

Figure 9.1: Molecular structure of cyanovinylacetylene isomers.

*The contents of this Chapter are taken from Ref. 260 after re-elaboration.

chemical computations have been carried out. At first, the jB2 level of theory has been used to explore the PES: two minimum energy structures have been found, corresponding to the two isomers reported in Figure 9.1. In addition, the jun-ChS composite scheme has been employed for the refinement of the equilibrium structure; all jun-ChS and jB2 geometries are collected in Table A.5. The equilibrium rotational constants have been straightforwardly derived from the jun-ChS geometries, and have been vibrationally corrected at the jB2 level of theory. As a byproduct of the anharmonic calculations, the correspondig ZPE contribution to energetics, as well as quartic and sextic centrifugal distortion constants have been derived. First-order properties such as electric dipole moment components and nuclear quadrupole coupling constants have been determined at the same level of theory. The computed spectroscopic parameters are reported in Tables 9.2 and 9.3.

Moreover, the energetic characterization of the $[H_3C_5N]$ family of isomers has been carried out. Thus, the study also concerns vinylcyanoacetylene (H_2CHC_3N, 4-penten-2-ynenitrile), in addition to the two cyanovinylacetylene isomers reported in Figure 9.1. The results are in disagreement with those obtained in Ref. 103, where vinylcyanoacetylene is indicated as the most stable isomer. This is the case also if the jun-ChS electronic energies (evaluated using the jB2 geometries as reference, see Table A.6 for the jB2 structure of vinylcyanoacetylene) are considered. However, when the harmonic ZPE corrections at the jB2 level are added, the relative stability order changes with respect to that reported in Ref. 103. In fact, vinylcyanoacetylene is found to lie roughly 1 kJ mol^{-1} higher in energy then the most stable (*E*)-cyanovinylacetylene. The *Z* isomer, instead, remains the highest in energy (1.4 kJ mol^{-1} above the *E* one). Considering the abundances of these species detected towards TMC-1,[103] the column densities reported (namely, 2×10^{11} cm^{-2} and 3×10^{11} cm^{-2} for vinylcyanoacetylene and (*E*)-cyanovinylacetylene, respectively) tend to suggest that the abundances of the $[H_3C_5N]$ isomers are in agreement with their thermodynamic stability, thus following the minimum energy principle.[263–265]

9.2 Experimental study

As shown in Figure 9.1, cyanovinylacetylene can exist in two isomeric forms, *E* and *Z*, depending on the orientation of the substituents with respect to the C=C double bond. Both geometrical isomers are planar (indeed, they belong to the C_s symmetry point group) asymmetric-top rotors, with $\kappa = -0.99$ and -0.71 for the *E* and *Z* forms, respectively. In this work,

the rotational spectra of cyanovinylacetylene isomers have been recorded using the FM-mmW spectrometer described in Section 6.2; the laboratory conditions are detailed in Section 6.4. The former isomer is predicted to have a predominant *a*-type spectrum (see Table 9.2), while the latter is characterized by both *a*- and *b*-types spectra of similar intensity (see Table 9.3). As shown in Table 6.3, in previous works, the rotational spectra of (*E*)- and (*Z*)-cyanovinylacetylene have been studied up to *ca.* 40 GHz[219,229] and up to 15 GHz, respectively.[212]

To support the measurements, the available experimental data (taken from the CDMS[11,12]) have been complemented by the computed terms collected in Tables 9.2 and 9.3. Indeed, only an incomplete set of quartic centrifugal distortion constants were determined in the previous works on (*E*)-cyanovinylacetylene;[219,229] moreover, A and d_1 were fitted with large uncertainties (see the "previous results" column in Table 9.2). Instead, for (*Z*)-cyanovinylacetylene, all rotational and quartic terms were already available, as shown in Table 9.3.[212] However, the importance of the inclusion of high-order centrifugal terms in the prediction can be seen in Figure 9.2 for the *Z* isomer. Indeed, it is noted that the previous line catalog (from CDMS) was not able to correctly reproduce the rotational spectrum already at 288 GHz. Not only the rotational transitions are shifted in frequency, but the resulting pattern is completely different. Instead, the incorporation of

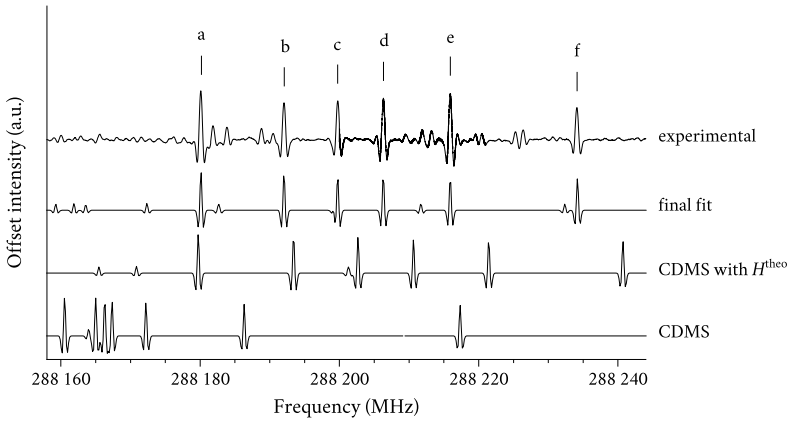


Figure 9.2: From top to bottom, for (*Z*)-cyanovinylacetylene: i. experimental FM-mmW absorption spectrum; ii. prediction based on final spectroscopic parameters; iii. prediction based on CDMS data augmented by computed sextic centrifugal distortion terms (jB2 level); and iv. prediction based solely on CDMS data. The transitions identified by letters have been have been assigned, and they are collected in Table 9.1.

Table 9.1: List of a few rotational transitions of (*Z*)-cyanovinylacetylene, where the letters refer to the labeling used in Figure 9.2. The observed frequencies are expressed in MHz.

J	K_a	K_c	J'	K'_a	K'_c	obs. ν[a]	
74	0	74	73	0	73	288 180.136	a
73	2	72	72	1	71	288 192.052	b
73	1	72	72	2	71		b
73	1	72	72	1	71		b
73	2	72	72	2	71		b
72	3	70	71	2	69	288 199.753	c
72	2	70	71	3	69		c
72	2	70	71	2	69		c
72	3	70	71	3	69		c
71	4	68	70	3	67	288 206.330	d
71	3	68	70	4	67		d
71	3	68	70	3	67		d
71	4	68	70	4	67		d
70	4	66	69	5	65	288 215.936	e
70	5	66	69	4	65		e
70	4	66	69	4	65		e
70	5	66	69	5	65		e
69	6	64	68	5	63	288 234.107	f
69	5	64	68	6	63		f
69	5	64	68	5	63		f
69	6	64	68	6	63		f

[a] The experimental uncertainty has been set to 30 kHz, while an empty frequency value denotes that the transition is blended with the previous one.

the computed sextic centrifugal distortion constants leads to a reasonably good predition of the rotational spectrum. The transitions represented in Figure 9.2 are collected in Table 9.1.

The hyperfine structure due to the presence of the ^{14}N quadrupolar nucelus is resolved for both isomers in the low-frequency studies,[212, 229] while it is completely collapsed in the FM-mmW experiment. Therefore, its effects have been ignored in this analysis. In this work, about 1 000 and 400 lines have been recorded and analyzed for (*E*)- and (*Z*)-cyanovinylacetylene, respectively, in the 80–115 GHz and 245–400 GHz ranges. Both *a*- and *b*-type transitions have been included for the latter, while only *a*-type ones for the former. The data have been merged with the rotational transitions available from the literature[212, 219, 229] in a global fit. The original uncertaineties have been retained for the literature data, while an uncertainty of 10–30 kHz depending on the line width has been assigned to the newly recorded frequencies.

Table 9.2: Ground-state spectroscopic parameters of (*E*)-cyanovinylacetylene (*S*-reduction, I^r representation). Values in parentheses denote one standard deviation and apply to the last digit of the constants.

		(*E*)-cyanovinylacetylene		
		exp.[a]	theo.[b]	prev.[c]
A	MHz	46 248.0(2)	46 204.0	46 226.0(18)
B	MHz	1 472.15368(4)	1 472.9	1 472.1538(2)
C	MHz	1 426.24867(4)	1 426.9	1 426.2484(1)
D_J	Hz	96.824(2)	91.5	95.8(7)
D_{JK}	kHz	−14.7991(2)	−14.4	−14.81(1)
D_K	MHz	1.700(24)	1.62	
d_1	Hz	−7.0554(9)	−6.58	−7.2(8)
d_2	Hz	−0.181(1)	−0.148	
H_J	µHz	34.50(6)	30.6	
H_{JK}	mHz	−7.465(6)	−6.76	
H_{KJ}	mHz	−345.5(8)	−369.9	
H_K	mHz		0.104	
h_1	µHz		5.29	
h_2	µHz		0.232	
h_3	µHz		0.0312	
χ_{aa}	MHz	−3.90(2)	−4.06	−3.90(1)
χ_{bb}	MHz	1.80(3)	1.93	1.82(3)
$\lvert\mu_a\rvert$	D	Y[d]	4.06	
$\lvert\mu_b\rvert$	D	N	0.61	
# lines		1 061		100
J^{max}, K_a^{max}		135, 17		13, 9
rms error	kHz	28.7		39.8
σ		1.01		0.81

[a] Missing values have been kept fixed at the theoretical counterpart in the fit procedure. [b] Jun-ChS equilibrium rotational constants augmented by vibrational corrections at the jB2 level. Quartic and sextic centrifugal distortion constants as well as nuclear quadrupole coupling constants at the jB2 level of theory. [c] Re-fitted using the *S*-reduction from Ref. 229. [d] For each μ_g dipole moment component, "Y" and "N" refer to detected and non-detected *g*-type transitions, respectively.

Table 9.3: Ground-state spectroscopic parameters of (*Z*)-cyanovinylacetylene (*S*-reduction, I^r representation). Values in parentheses denote one standard deviation and apply to the last digit of the constants.

		(*Z*)-cyanovinylacetylene		
		exp.[a]	theo.[b]	prev.[c]
A	MHz	7 098.2006(2)	7 080.7	7 098.2003(4)
B	MHz	2 682.59013(6)	2 693.0	2 682.5904(2)
C	MHz	1 943.54387(5)	1 947.9	1 943.5440(1)
D_J	kHz	3.2784(3)	3.00	3.290(5)
D_{JK}	kHz	−21.131(3)	−20.5	−21.21(2)
D_K	kHz	44.22(3)	45.4	44.43(9)
d_1	kHz	−1.2796(1)	−1.15	−1.281(2)
d_2	kHz	−0.07172(3)	−0.0598	−0.068(2)
H_J	mHz	11.24(5)	16.4	
H_{JK}	mHz	33.0(1)	−91.3	
H_{KJ}	µHz		−0.164	
H_K	µHz		0.842	
h_1	mHz	6.51(2)	8.48	
h_2	mHz	1.619(5)	1.01	
h_3	mHz	0.3043(9)	0.220	
χ_{aa}	MHz	0.249(1)	0.190	0.2490(8)
χ_{bb}	MHz	−2.424(1)	−2.44	−2.4237(8)
$\lvert\mu_a\rvert$	D	Y[d]	2.58	
$\lvert\mu_b\rvert$	D	Y	2.72	
# lines		493		90
J^{max}, K_a^{max}		103, 14		5, 2
rms error	kHz	24.7		1.8
σ		0.91		0.66

[a] Missing values have been kept fixed at the theoretical counterpart in the fit procedure. [b] Jun-ChS equilibrium rotational constants augmented by vibrational corrections at the jB2 level. Quartic and sextic centrifugal distortion constants as well as nuclear quadrupole coupling constants at the jB2 level of theory. [c] Re-fitted using the *S*-reduction from Ref. 212. [d] For each μ_g dipole moment component, "Y" refers to detected *g*-type transitions.

The results are collected in Tables 9.2 and 9.3, together with the computed values. The agreement between theoretical and experimental values is somewhat different for the E and Z isomers: excellent for the former, with an average discrepancy of 0.06%, and good for the latter, which shows 5-times larger average deviation, close to 0.3%. This behavior is replicated in the comparison of the sextic centrifugal distortion terms (mean deviation of *ca.* 9% and *ca.* 35% for E and Z forms, respectively), while quartic terms are in rather good agreement in both cases (8% for the E isomer and 9% for the Z one). Indeed, for (*Z*)-cyanovinylacetylene, the worst agreement is found in the value of H_{JK}, which disagrees in both size and sign with the corresponding experimental result, which is however poorly determined.

Nevertheless, the comparison with the previous experimental results highlights a clear improvement in the accuracy of the spectroscopic parameters, as well as an exaustive characterization. The global fit for (*E*)-cyanovinylacetylene is well conditioned, with a final standard deviation (σ) close to unity and a rms error of 28.7 kHz; similar outcomes are found for the Z isomer, with σ below unity and rms equal to 24.7 kHz. Based on these fit statistics, we can affirm that the global fits are able to reproduce well the experimental data.

Chapter 10

(*E*)- and (*Z*)-Phenylmethanimine

As introduced in Section 1.3, phenylmethanimine (also known as benzylimine, see Figure 10.1) is a non-standard reactive molecule of potential astrochemical relevance. In addition, imines are recognized as crucial intermediates in organic synthesis, due to their extensive use in the preparation of N-containing compounds,[266] and as naturally or biologically active molecules.[267] However, the instability of these species often requires an *in situ* generation inside the spectrometer, hence resorting to poorly predictable and rather harsh techniques. In this Chapter, this issue has been addressed by developing an easy synthetic route, *i.e.*, the thermally tunable formation of imines. In this Chapter,* the results of its testing both in the gas phase (leading to the analysis of the first rotational spectrum of phenylmethanimine) and in solution (resorting to proton nuclear magnetic resonance, ^{1}H-NMR) have been reported. In addition, the outcomes of an accurate computational and energetic characterization of phenylmethanimine are collected.

(*E*)-phenylmethanimine (*Z*)-phenylmethanimine

Figure 10.1: Molecular structure of phenylmethanimine isomers.

*The contents of this Chapter are taken from Ref. 243 after re-elaboration.

10.1 Computational characterization

An accurate quantum-chemical characterization is the essential starting point for guiding the rotational spectroscopy experiment of a new species. For phenylmethanimine, a preliminary investigation of the PES has been conducted at the B3LYP-D3(BJ)/SNSD level of theory (from here on referred to as B3S): this located the two minima shown in Figure 10.1, namely the *E* and *Z* isomers, according to the relative position of the hydrogen atom and the phenyl group with respect to the C=N double bond. Further calculations at the same level of theory have been performed in order to understand the isomerization mechanism. As reported in Ref. 268, the ratio between in-plane inversion and rotation of the iminic N–H bond depends on the substituents of the R^1R^2CNX system. In this case, R^1 = H, R^2 = Ph, and X = H, which is a missing case in the study of Ref. 268. However, based on the B3S results, the isomerization process is well described by an in-plane inversion motion as found for ethanimine[217] and C-cyanomethanimine.[269] Thus, the transition state is characterized by the *quasi*-linearity of the C=N–H moiety. As a matter of fact, no transition states have been found along the rotation of the N–H bond around the C=N axis.

Subsequently, the geometries of the three stationary points have been refined at the B2PLYP-D3(BJ)/maug-cc-pVTZ level of theory (hereafter, mB2), with the harmonic force field being also computed at the same level of theory. All of them show a planar structure and, thus, belong to the C_s point group. At this point, to confirm that the imaginary frequency of the transition state corrisponds to the isomerization coordinate, an intrinsic reaction coordinate[270] (IRC) calculation has been performed at the mB2 level of theory. In order to obtain an accurate estimate of the rotational constants, the geometry of each minimum has been further refined by resorting to the ChS composite scheme. The mB2 and ChS optimized geometries are collected in Tables A.7 and A.8.

The ChS approach has also been used to derive accurate electronic energies on top of the mB2 optimized geometries: the *Z* isomer is found to lie 6.5 kJ mol^{-1} (6.3 kJ mol^{-1} if harmonic ZPE at the mB2 level is considered) higher in energy than the most stable *E* one. The barrier for the isomerization process is 119.9 kJ mol^{-1} (112.2 kJ mol^{-1} with harmonic mB2 ZPE included). A qualitative comparison with different imines, such as ethanimine[217] and C-cyanomethanimine,[269] can be drawn even if different computational schemes have been applied. The values are in a general agreement: a smaller energy difference between the isomers has been found

(2.8 kJ mol^{-1} for ethanimine and 2.0 kJ mol^{-1} for C-cyanomethanimine), probably justified by the size of the substituent on the CHNH moiety (methyl and cyano, respectively). For the isomerization barrier, similar values are found (*ca.* 116 kJ mol^{-1} for ethanimine and *ca.* 111 kJ mol^{-1} for C-cyanomethanimine). In all cases, the isomerization occurs on the molecular plane, through a *quasi*-linear transition state.

To complete the spectroscopic characterization, the equilibrium rotational constants—straightforwardly derived from the ChS equilibrium structures—have been corrected for the vibrational contributions at the B3S level. First-order properties, such as dipole moment components and nuclear quadrupole coupling constants, have been computed at the mB2 level of theory; quartic centrifugal distortion constants have been derived at the same level. The quantum-chemical spectroscopic parameters are collected in Table 10.1.

10.2 Experimental study

The first rotational spectroscopy study on phenylmethanimine (Figure 10.1) has been performed in the 3–26 GHz range using the COBRA-FTMW spectrometer described in Section 6.3, starting from the solid precursor HBA, whose synthesis and characterization is addressed in Section A.2. The laboratory conditions are summarized in Section 6.4.

Both phenylmethanimine isomers are nearly-prolate asymmetric-top rotors with $\kappa = -0.82$. For (*Z*)-phenylmethanimine, the computed dipole moment components (see Table 10.1) predict a strong *a*-type spectrum and a *ca.* 400-times weaker *b*-type one, while the spectrum of (*E*)-phenylmethanimine should be dominated by *b*-type transitions, with *a*-type ones only four to five times weaker.

Thanks to the high accuracy of the quantum-chemical simulations, it has been possible to unequivocally identify both isomers of phenylmethanimine in the gas phase. Indeed, more than one hundred transition frequencies have been registered and assigned, with an example of them shown in Figure 10.2. Both *a*- and *b*-type rotational transitions have been observed and assigned for the *E* and *Z* isomers, although only a few *b*-type components have been fitted for (*Z*)-phenylmethanimine, in accordance with its low predicted dipole moment along the *b* axis.

The global rotational Hamiltonian can be expressed as the sum of two terms

$$\hat{\mathscr{H}}_{\mathrm{rot}} = \hat{\mathscr{H}}^{S}_{\mathrm{rot}} + \hat{\mathscr{H}}_{\mathrm{hfs}}\,, \tag{10.1}$$

Table 10.1: Ground-state spectroscopic parameters of the *E* and *Z* isomers of phenylmethanimine (*S*-reduction, I^r representation). Values in parentheses denote one standard deviation and apply to the last digits of the constants.

		(*E*)-phenylmethanimine		(*Z*)-phenylmethanimine	
		exp.	theo.[a]	exp.[b]	theo.[a]
A	MHz	5 217.2920(1)	5 217.0	5 200.8128(2)	5 200.4
B	MHz	1 565.28363(3)	1 564.7	1 548.96935(9)	1 548.8
C	MHz	1 204.54031(1)	1 204.1	1 194.84231(8)	1 194.3
D_J	Hz	57.8(1)	55.7	56.4(4)	52.8
D_{JK}	kHz	0.168(4)	0.181	0.145(5)	0.157
D_K	kHz	0.77(1)	0.707		0.706
d_1	Hz	−16.94(6)	−16.3	−15.4(4)	−15.1
d_2	Hz	−3.16(2)	−3.00		−2.68
$\frac{3}{2}\chi_{aa}$	MHz	1.527(1)	1.61	−5.887(2)	−6.05
$\frac{1}{4}(\chi_{bb}-\chi_{cc})$	MHz	−1.7063(4)	−1.83	−0.495(2)	−0.555
$\lvert\mu_a\rvert$	D	Y[c]	0.77	Y	3.09
$\lvert\mu_b\rvert$	D	Y	1.66	Y	0.15
# lines		180		118	
J^{max}, K_a^{max}		43, 4		9, 3	
rms error	kHz	4.56		1.38	
σ		0.65		0.69	

[a] ChS equilibrium rotational constants augmented by vibrational corrections at the B3S level. Quartic centrifugal distortion and nuclear quadrupole coupling constants as well as dipole moment components at the mB2 level. [b] Missing values have been kept fixed at the theoretical value in the fit procedure. [c] For each μ_g dipole moment component, "Y" refers to detected *g*-type transitions.

where $\hat{\mathcal{H}}^S_{\text{rot}}$ is explained in Equation (2.58) and $\hat{\mathcal{H}}_{\text{hfs}}$ is the hyperfine-structure Hamiltonian (see Equation 2.65) due to the presence of a nitrogen atom; the coupling scheme of Equation (2.61) has been used. As a matter of fact, the hyperfine structure has been of great help for discerning between phenylmethanimine and other thermolysis by-products.

In addition, the spectral analysis of phenylmethanimine has been extended in the 83–100 GHz range using the FM-mmW spectrometer described in Section 6.2, with a different production method (*viz.* FVP, as explained in Section 6.4) with respect to that used in the FTMW experiment; the underlying reason will be described in the next Section. Within this experiment, a small set of 27 transitions ascribable to (*E*)-phenylmethanimine has been measured, showing a collapsed hyperfine structure. No transitions of the *Z* isomer have been found: this may suggest that only (*E*)-

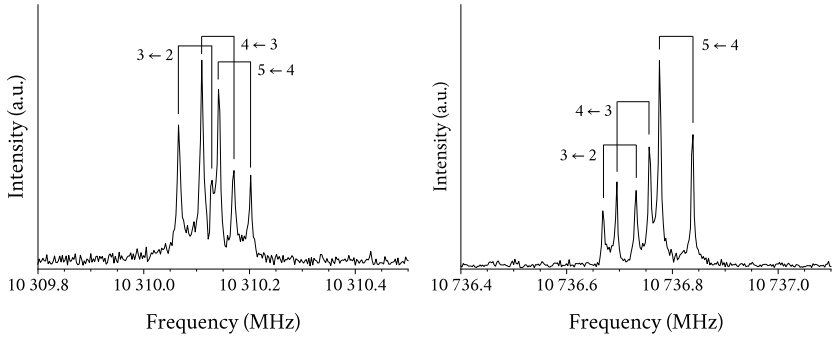


Figure 10.2: FTMW spectra of the $4_{1,4} \leftarrow 3_{1,3}$ transition of (*E*)-phenylmethanimine (left panel) and the $4_{0,4} \leftarrow 3_{0,3}$ transition of (*Z*)-phenylmethanimine (right panel). The $F \leftarrow F'$ quantum numbers are reported and the splitting due to Doppler effect is shown.

phenylmethanimine is generated *via* FVP, but this aspect has not been further investigated.

Uncertainities of 2 kHz have been assigned to FTMW lines, and in the 10–20 kHz range for the FM-mmW ones. The spectroscopic parameters derived from the fitting of the observed transitions are reported in Table 10.1: an excellent agreement between theoretical and experimental values is noted. Indeed, the average error on the effetive rotational constants is 0.03%, with a maximum discrepancy of 0.05%. A good agreement is also found for the quartic centrifugal distortion terms and nuclear quadrupole coupling constants. In the analysis of the *Z*-isomer rotational spectrum, the quartic centrifugal distortion constants D_K and d_2 could not be properly fitted, and thus have been kept fixed at the corresponding calculated values. Finally, the fit statistics points out the achievement of the experimental accuracy for both isomers: the rms error is below 5 kHz for (*E*)-phenylmethanimine and below 2 kHz for the *Z* form, while the standard deviation (σ) is in both cases below unity.

10.3 Thermal decomposition of hydrobenzamide

The successful generation and characterization of phenylmethanimine in the gas phase raised the need of a better understanding of the mechanism for its formation, and the possible extension of the methodology to the condensed phase (*i.e.*, in solution) or, more generally, to other N-unsubstituted imines (thus avoiding more expensive and/or more chemically demanding procedures). Therefore, the thermal behavior of hydrobenzamide has been

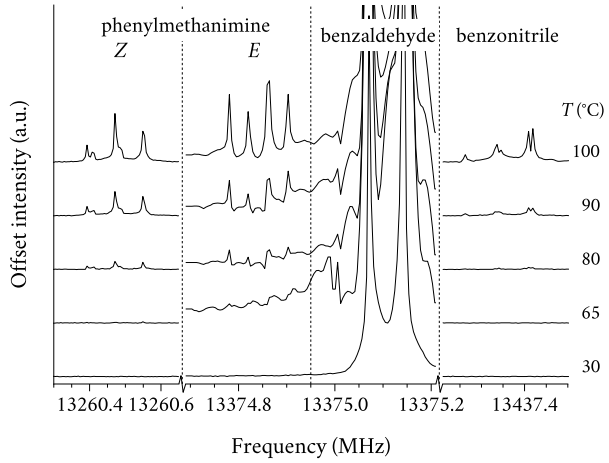


Figure 10.3: FTMW spectra of the $5_{0,5} \leftarrow 4_{0,4}$ transition for phenylmethanimine, benzaldehyde, and benzonitrile. The huge abundance of benzaldehyde in the gas phase leads to out-of-chart intensities at any temperature.

explored from room temperature up to 100–110 °C by means of rotational and ^{1}H-NMR spectroscopies. For the latter, a description of the instrument and working conditions can be found in Section A.2.

In particular, the $5_{0,5} \leftarrow 4_{0,4}$ transition of four molecular species, namely benzaldehyde, benzonitrile, (*E*)- and (*Z*)-phenylmethanimine, has been recorded with the COBRA-FTMW spectrometer. To provide a reliable comparison in terms of intensity of the transitions, the field amplitude of the microwave excitation pulse has been scaled reciprocal to the value of the dipole moment component along the *a* axis, and a 2 000 averaging signal acquisition has been chosen. The recordings have been started ten minutes after the temperature set up, in order to guarantee the reaching of the thermal equilibrium.

The recorded spectra are represented in Figure 10.3, which reveal the presence of benzaldehyde at all temperatures. The signals ascribable to phenylmethanimine isomers are observed only above 80 °C, together with the benzonitrile ones. For both phenylmethanimine and benzonitrile, an increase in the nozzle head temperature is followed by an increase in the transition intensity, thus improving the S/N. Although not shown in Figure 10.3, the 12 321.0 MHz transition of the water dimer[271] has been observed for each temperature increment. This confirms the constant presence of water in the experimental apparatus, which very likely plays a key role in the investigated thermolytic process.

An important note on the relative populations of the two phenyl-

methanimine isomers is deserved: in principle, this information should be derivable from the spectra shown in Figure 10.3; however, the (*E*)-phenylmethanimine transition reported lies very close to the strong feature of benzaldehyde, which clearly modifies the baseline of the (*E*)-phenylmethanimine transition. As a consequence, the S/N of the *Z* and *E* forms cannot be quantitatively compared, thus preventing the derivation of the relative population. A qualitative conclusion can however be drawn; indeed, during the recording of the spectra, the intensities have been found to be comparable. This is probably due to a cooperation of population distribution and differences in the electric dipole moment components.

Moving to the ^{1}H-NMR characterization, according to the literature,[272] phenylmethanimine shows two distinct sets of signals for the imino hydrogen, each with two doublets, and with different coupling constants ($J = 16$ Hz and $J = 25$ Hz for the *E* and *Z* isomers, respectively). In our analysis, the sample was heated to the desired temperature and the ^{1}H-NMR spectrum was acquired after ten minutes, thus avoiding thermalization processes during the acquisition. The formation of both benzaldehyde and (*E*)-phenylmethanimine has been observed as a consequence of the step-by-step temperature ramp from 25 °C to 110 °C, as shown in the top panel of Figure 10.4. However, no NMR signals denoting the possible formation of benzonitrile or (*Z*)-phenylmethanimine have been observed, even when the experiments have been conducted at higher temperature. As far as the *Z* isomer is concerned, however, our experimental resolution did not allow for unequivocally excluding its presence.

This suggests that the rotational and NMR experiments are characterized by either: i. different hydrolytic mechanisms; or ii. additional pathways only available in the gas-phase measurements. Moreover, to confirm the absence of benzonitrile in solution after the thermal treatment of hydrobenzamide, a trace amount of the former has been added to the NMR samples previously heated at 110 °C, and the spectrum has been recorded again. This analysis revealed the presence of new signals (*cf.* Section A.2) due to the presence of benzonitrile. This confirms that no benzonitrile is formed *via* hydrolysis of HBA in solution; therefore, in the gas phase, its generation might occur through pathways that rely on metallic catalytic surfaces, such as fast dehydrogenation of phenylmethanimine.[273,274] Within this context, it is proven that metal catalysis can take place in the nozzle head of FTMW experiments,[275] whereas this is not the case in the NMR tube.

In detail, Figure 10.4 shows the changes upon heating of the diagnostic NMR signals (top panel) and the temperature dependence of their normalized integral values (bottom panel). This analysis was performed

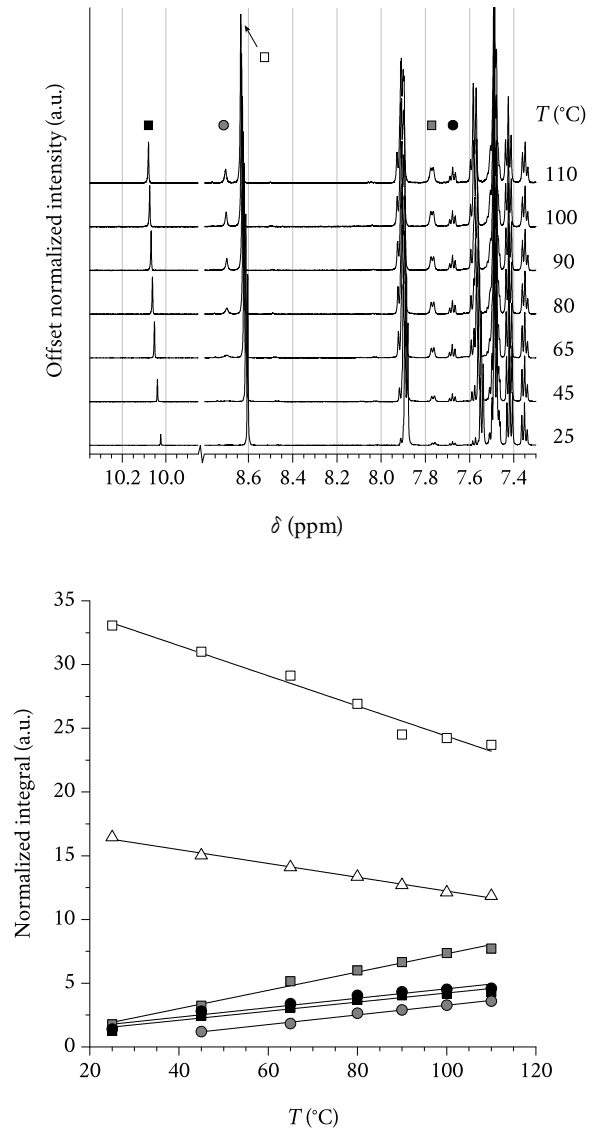


Figure 10.4: ^{1}H-NMR signals at different temperatures for HBA and its thermolysis products (top panel) and temperature-dependence of their normalized integral values (bottom panel). The integral value of the residual peak of 1,1,2,2-tetrachloroethane-d_2 has been arbitrarily fixed to 10. In both panels, white markers represent HBA (□: CHN, △: aminal H), gray markers represent (*E*)-phenylmethanimine (■: *ortho*-protons, ●: CHN), and black ones represent benzaldehyde (■: CHO, ●: *para*-proton).

after an arbitrary assignment to the integral value of the residual peak of 1,1,2,2-tetrachloroethane-d_2 (fixed to 10). The traces with white markers are characterized by a negative slope, and they correspond to the starting material (hydrobenzamide). As expected, the signal relative to the two CHN protons of HBA (identified by white triangles in Figure 10.4) decreases twice as fast as the single-proton aminal signal (white squares). The traces with gray and black markers, instead, are attributable to forming species, and are indeed characterized by a positive slope. Two signals (black squares and circles) have almost superimposed trends, in terms of both slope and normalized integral values: this suggests that the respective protons belong to the same species. Indeed, they are compatible with the CHO and *para*-proton of benzaldehyde, respectively, thus featuring a 1:1 integration relationship. The other two signals (*i.e.*, gray squares and circles) feature a 2:1 ratio, and they are compatible with the iminic CHN proton (gray circles) and iminic *ortho*-protons (gray squares). Furthermore, the latter trace has a slope which is similar to those observed for the aldehydic signal, thus suggesting that the two generation processes may be interconnected. In addition, the fact that benzaldehyde signals appear before the phenylmethanimine ones suggests that the involved hydrolytic mechanism proceeds with the initial formation of benzaldehyde, followed by the formation of phenylmethanimine starting from different transient species.

In Figure 10.5, a putative hydrolytic pathway is proposed; it is compatible with the aforementioned observations and with previously proposed mechanisms.[276–278] As noted above, the additional pathway leading to the formation of benzonitrile in FMTW experiments is supposed to be a dehydrogenation.

The following key points need to be analyzed to explain the proposed mechanism: i. the role of water; and ii. the role of the metallic surface in the COBRA-FTMW spectrometer with respect to the glass surfaces of the FM-mmW experiment. Within the former aspect, water is expected to play a crucial role, and the hydrolytic pathways is supposed to proceed as long as water is available. Indeed, the amount of water detected by means of the

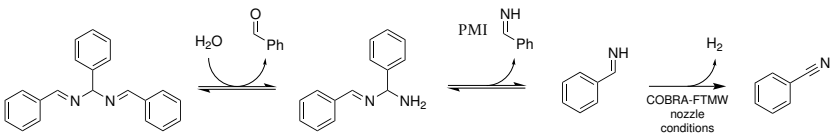


Figure 10.5: Proposed mechanism for HBA hydrolysis.

FTMW experiment as well as the residual water of the deuterated solvent is believed to be sufficient to observe the formation of the main hydrolytic products, *i.e.*, benzaldehyde at room temperature, and phenylmethanimine at higher temperatures.

The FM-mmW experiment can provide clues for a deeper understanding of the hydrolytic mechanism. At first, a tentative generation of phenylmethanimine in the gas phase has been carried out by thermolysis of hydrobenzamide. The solid has been placed in a glass tube, and heated up to 100 °C, also ensuring a uniform heating along the path to the absorption cell. While heating, a portion of the spectrum around 85.5 GHz has been scanned in the attempt of detecting two strong transitions of (*E*)-phenylmethanimine, as predicted by cm-wave measurements. No signal attributable to phenylmethanimine (nor benzonitrile) has been found. Typically, the whole glass apparatus of the FM-mmW is pumped continuously, thereby removing water, although its residual presence cannot be ruled out. However, no metallic surfaces are available in the instrument, thus formulating two hypotheses for the lack of formation of phenylmethanimine in the gas phase: i. water is not available in a sufficient amount to hydrolyze HBA; and/or ii. the metal catalysis is mandatory to obtain phenylmethanimine from HBA in the gas phase.

The reliability of spectral predictions is verified by the FM-mmW spectra of (*E*)-phenylmethanimine *via* FVP described above. In this experiment, benzonitrile has been found as a pyrolysis co-product, as proven by recording its rotational transitions. Furthermore, while the signals of phenylmethanimine reached their maximum intensity by setting the furnace temperature to 890 °C, the intensity of benzonitrile transitions kept increasing up to 1 200 °C, thus confirming the prevalence of benzonitrile at higher temperatures (in agreement with the cm-wave experiment). However, as mentioned above, no signal ascribable to the *Z* isomer has been found, despite the several attempts; in conclusion, the FM-mmW experiment proved the reliability of the low-frequency measurements, but it left some unexplored areas concerning phenylmethanimine formation from HBA in the gas phase when no metallic surfaces are involved.

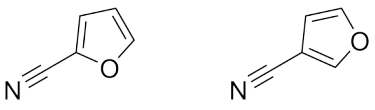
Chapter 11

2- and 3-Furonitrile

As explained in Section 1.3, 2- and 3-furonitrile (or cyanofuran, shown in Figure 11.1) represent valid candidates for the first detection of heterocycles derivatives in the interstellar medium. In order to provide the set of spectroscopic parameters mandatory for their astronomical search, an accurate computational and experimental characterization of both isomers has been carried out, and the results are collected in this Chapter. During the preparation of this manuscript, the author become aware of a recent publication by Esselman *et al.*, where the spectrum of 2-furonitrile has been exhaustively studied up to 750 GHz.[107] However, the measurements reported here have been conducted at the beginning of 2021, and are part of a wider project object of a future publication[279] without any conflict of interest with the results reported in Ref. 107.

11.1 Computational characterization

At first, the computational characterization of 2- and 3-furonitrile has been conducted at the revDSD level of theory, leading to the identification of the two structures corresponding to the minima sketched in Figure 11.1. Then, their equilibrium structure has been refined at the jun-ChS level, in order



2-furonitrile 3-furonitrile

Figure 11.1: Molecular structure of 2- and 3-furonitrile.

to obtain an accurate estimate of the equilibrium rotational constants. Both jun-ChS and revDSD geometries are collected in Table A.9. The vibrational corrections needed to derive the ground-state rotational constants have been evaluated at the revDSD level. The same level of theory has been used for the determination of quartic and sextic centrifugal distortion constants. The equilibrium values of the electric dipole moment components have been computed at the fc-CCSD(T)/jun-cc-pVTZ level of theory, while the nuclear quadrupole coupling constants at the equilibrium have been calculated at the jun-ChS level; both properties have then been vibrationally corrected at the revDSD level. The results are collected in Tables 11.1 and 11.2 for 2- and 3-furonitrile, respectively.

The jun-ChS model has also been used for the energetic characterization of the system, on top of jun-ChS optimized geometries: 3-furonitrile is found to lie 8.12 kJ mol^{-1} lower in energy with respect to 2-furonitrile, and 7.70 kJ mol^{-1} if the anharmonic ZPE is added (calculated at the revDSD level of theory).

11.2 Experimental study

From a spectroscopic point of view, 2- and 3-furonitrile (Figure 11.1) are nearly-prolate asymmetric-top rotors with $\kappa \approx -0.90$. Both molecules are planar and hence belong to the C_s symmetry point group. Given the values of the electric dipole moment components, reported in Tables 11.1 and 11.2 for the 2- and 3-substituted isomer, respectively, a strong *a*-type spectrum is expected for both isomers, together with a roughly 25-times and 90-times weaker *b*-type one. In particular, for these species, the values of the dipole moment component along the *a* axis are particularly large ($\mu_a = 4.48$ D for 2-furonitrile and $\mu_a = 3.78$ for the 3-substituted isomer), thus rendering these molecules ideal candidates for astronomical searches. However, both isomers possess a quadrupolar nucleus (the ^{14}N atom), and, therefore, the population of rotational levels is distributed among hyperfine states (*cf.* Equation 2.34). In the treatment of this interaction, the coupling scheme from Equation (2.61) has been used.

While no previous spectroscopic characterization is available for 3-furonitrile, the 2-substituted isomer has already been characterized in the centimeter-wave frequency range in Refs. 230 and 231. Large uncertainities have been reported on the experimental line positions in both works (100 kHz for Ref. 230 and 50 kHz for Ref. 231), and only a set of rotational constants and nuclear quadrupole coupling constants have been determined. The results reported in these two publications have been re-fitted

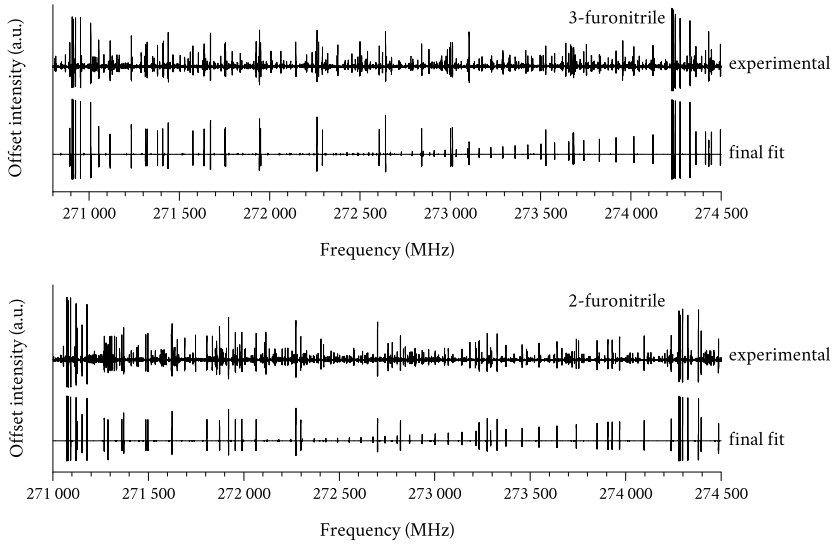


Figure 11.2: Comparison between the FM-mmW experimental spectrum and the simulation based on the fitted spectroscopic parameters for both 2- (bottom panel) and 3-furonitrile (top panel). A frequency range of *ca.* $B + C$ MHz has been considered, and the strongest patterns are constituted by transitions with values of J around 80 and K_a in the 0–5 range.

and collected in the "previous results" column of Table 11.1. As mentioned above, Esselman *et al.*[107] published a detailed characterization on the same species: these results are also included in Table 11.1.

At first, the experiment has been conducted by scholars at the School of Chemistry and Chemical Engineering of the Chongqing University (China), in the group led by Prof. Dr. Qian Gou, using a COBRA-type FTMW spectrometer. Thanks to the high resolution of this spectrometer, the experimental uncertainty on the frequency has been reduced by 25 to 50 times with respect to the previous works. Moreover, given the absence of strong interfering lines—the samples are commercially available with high purity, see Section 6.4—and the accuracy of the spectral predictions,*ca.* 170 and 200 lines have been recorded and assigned in the cm-wave range for the 2- and 3-substituted isomers, respectively.

Subsequently, the two regioisomers have been characterized at higher frequency up to the submillimeter-wave range, thereby exploiting the FM-mmW spectrometer described in Section 6.2 (laboratory conditions are described in Section 6.4). Using the spectral predictions based on the refined parameters from the cm-wave experiment, the extension in

Table 11.1: Ground-state spectroscopic parameters of 2-furonitrile (*S*-reduction, I^r representation). Values in parentheses denote one standard deviation and apply to the last digit of the constants.

		2-furonitrile			
		exp.[a]	theo.[b]	prev.[c]	from Ref. 107[d]
A	MHz	9 220.2506(1)	9 235.1	9 220.11(1)	9 220.2514(1)
B	MHz	2 029.27372(2)	2 030.9	2 029.271(3)	2 029.27414(2)
C	MHz	1 662.64314(2)	1 664.2	1 662.642(2)	1 662.64352(2)
D_J	Hz	59.598(5)	56.0		59.680(2)
D_{JK}	kHz	2.9094(1)	2.85		2.91036(2)
D_K	kHz	0.296(8)	0.272		0.2975(2)
d_1	Hz	−15.166(2)	−14.2		−15.1757(5)
d_2	Hz	−9.994(3)	−9.51		−9.9979(3)
H_J	µHz	−31.8(3)	−30.8		−28.01(8)
H_{JK}	mHz	4.32(1)	4.23		4.507(2)
H_{KJ}	mHz	−35.91(8)	−35.0		−35.91(2)
H_K	mHz		32.2		33.8(2)
h_1	µHz	−1.8(2)	−1.73		−0.85(2)
h_2	µHz	18.1(2)	17.3		18.83(3)
h_3	µHz	3.56(9)	3.37		3.72(1)
L_J	nHz				0.017(1)
L_{JJK}	nHz				−11.41(3)
L_{JK}	µHz	0.149(5)			0.1509(5)
L_{KKJ}	µHz	−1.51(2)			−1.529(4)
L_K	µHz				1.14(4)
l_{JK}	pHz				−31.4(7)
l_{KKJ}	pHz				−7.0(5)
l_K	pHz				−1.6(1)
½χ_{aa}	MHz	−6.440(1)	−6.43	−6.4(4)	
¼($\chi_{bb} - \chi_{cc}$)	MHz	0.2566(4)	0.269	0.25(9)	
$\lvert\mu_a\rvert$	D	Y[e]	4.48	4.62(2)	
$\lvert\mu_b\rvert$	D	Y	0.89	0.90(1)	
# lines		1 315		39	10 143
J^{max}, K_a^{max}		95, 45		8, 3	190, 59
rms error	kHz	23.7		108.0	40.0
σ		0.94		1.36	0.80

[a] Missing values have been kept fixed at the theoretical counterpart in the fit procedure. If no theoretical value is present, the term has not been included in the fit procedure. [b] Jun-ChS equilibrium rotational constants augmented by vibrational corrections at the revDSD level. Quartic and sextic centrifugal distortion constants have been evaluated at the revDSD level of theory. The equilibrium values of the dipole moment components have been determined at the fc-CCSD(T)/jun-cc-pVTZ level of theory, while the nuclear quadrupole coupling constants at the equilibrium have been calculated exploiting the jun-ChS scheme. The last two quantities have been augmented by anharmonic vibrational corrections at the revDSD level. [c] Re-fitted using the *S*-reduction from Refs. 230 and 231. Dipole moment components from Ref. 230. [d] The values of the off-diagonal octic terms l_1 and l_{JJK} have been kept fixed to zero in the fit procedure. [e] For each μ_g dipole moment component, "Y" refers to detected *g*-type transitions.

Table 11.2: Ground-state spectroscopic parameters of 3-furonitrile (*S*-reduction, I^r representation). Values in parentheses denote one standard deviation and apply to the last digit of the constants.

		3-furonitrile	
		exp.[a]	theo.[b]
A	MHz	9 296.5489(1)	9 306.3
B	MHz	1 940.26649(2)	1 942.2
C	MHz	1 604.63205(2)	1 606.2
D_J	Hz	54.754(4)	51.8
D_{JK}	kHz	2.9624(1)	2.91
D_K	kHz	0.41(1)	0.353
d_1	Hz	−13.366(2)	−12.5
d_2	Hz	−9.089(3)	−8.69
H_J	µHz	−29.4(3)	−29.2
H_{JK}	mHz	4.31(1)	4.31
H_{KJ}	mHz	−38.03(9)	−37.6
H_K	mHz		34.5
h_1	µHz	−2.0(2)	−2.00
h_2	µHz	16.2(2)	15.7
h_3	µHz	3.26(8)	3.08
L_{JK}	µHz	0.154(5)	
L_{KKJ}	µHz	−1.62(2)	
$\frac{3}{2}\chi_{aa}$	MHz	−6.338(1)	−6.59
$\frac{1}{4}(\chi_{bb}-\chi_{cc})$	MHz	0.1580(3)	0.159
$\lvert\mu_a\rvert$	D	Y[c]	3.78
$\lvert\mu_b\rvert$	D	Y	0.40
# lines		1 377	
J^{max}, K_a^{max}		98, 45	
rms error	kHz	15.0	
σ		1.05	

[a] Missing values have been kept fixed at the theoretical counterpart in the fit procedure. [b] Jun-ChS equilibrium rotational constants augmented by vibrational corrections at the revDSD level. Quartic and sextic centrifugal distortion constants have been evaluated at the revDSD level of theory. The equilibrium values of the dipole moment components have been determined at the fc-CCSD(T)/jun-cc-pVTZ level of theory, while the nuclear quadrupole coupling constants at the equilibrium have been calculated exploiting the jun-ChS scheme. The last two quantities have been augmented by anharmonic vibrational corrections at the revDSD level. [c] For each μ_g dipole moment component, "Y" refers to detected g-type transitions.

the 240–320 GHz range proceeded without any difficulty. A portion of *ca.* $B + C$ MHz of the experimental spectrum of both 2- and 3-furonitrile is reported in Figure 11.2, where it is compared with the simulation based on the fitted spectroscopic parameters. More than 1 100 lines have been recorded and analyzed for both isomers, with values of J and K_a as high as 95 and 45, respectively. Both *a*- and *b*-type transitions have been assigned for each regioisomer in both FTMW and FM-mmW experiments; the uncertainty on the frequency value has been set to 2 kHz for the former spectrometer, and to 15 kHz for the latter.

For both isomers, the total rotational Hamiltonian is that of Equation (10.1); the spectroscopic parameters have been determined and collected in Tables 11.1 and 11.2 for 2- and 3-furonitrile, respectively. The rotational and nitrogen quadrupole coupling constants as well as the entire set of centrifugal distortion constants up to the sixth order (with the exception of H_K) and two octic terms (namely, L_{JK} and L_{KKJ}) have been determined with high precision for both 2- and 3-furonitrile. This notwithstanding, given the wide spectral coverage and the impressive number of transitions analyzed by Esselman *et al.*, it is not significant to discuss the rotational and centrifugal distortion constants determined in this work for 2-furonitrile. It should however be noticed the agreement between the results of this work and the ones reported in Ref. 107; that being said, the accurate determination of the nitrogen quadrupole coupling costants (achieved thanks to the cm-wave measurements) represents the primary contribution of this work to 2-furonitrile spectroscopic characterization.

Moving to 3-furonitrile, this work represents the first determination of its spectroscopic parameters, which have been determined with an accuracy comparable to that of 2-furonitrile. The experimental accuracy has been reproduced: the rms error of the fit is *ca.* 15 kHz, and the a standard deviation is close to unity. For 3-furonitrile, a complete set of quartic and sextic centrifugal distortion constants as well as nuclear quadrupole coupling constants have been determined accurately, the only exception being H_K, which has been kept fixed at the theoretical value in the fitting procedure. In addition, two higher-order terms (the octics L_{JK} and L_{KKJ}) have been included and determined with a good precision. From an inspection of both Tables 11.1 and 11.2, a very good agreement is noticed between the theoretical and experimental ground-state rotational constants, with an average discrepancy below 0.1%. Moreover, computed quartic and sextic centrifugal distortion terms are also in good agreement with their experimental counterparts, showing an absolute discrepancy of 6–7% for the quartic terms and 2–4% for the sextic ones, respectively. Finally, an

excellent agreement is found on nuclear quadrupole coupling constants: on average, the absolute deviation is below 3%.

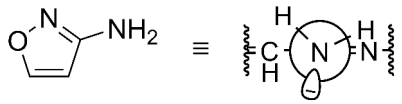
Chapter 12

3-Aminoisoxazole

As detailed in Section 1.3, 3-aminoisoxazole (see Figure 12.1) is a key reactive species in the scenario of the RNA-world hypothesis. In this Chapter,* the results of its computational and experimental characterization are provided and discussed.

12.1 Computational characterization

To locate the minimum energy structure of 3-aminoisoxazole, a preliminary scan of its PES has been carried out at the revDSD level of theory. In order to explore the possible orientations of the NH_2 group with respect to the isoxazole ring, two internal coordinates have been investigated, namely the dihedral angles defined by the N=C–N frame and each hydrogen atom of the amino moiety (therefore leading to two different N=C–N–H dihedral angles, see Table A.10). During the scan procedure, the other structural



3-aminoisoxazole

Figure 12.1: Molecular structure of 3-aminoisoxazole. In the right panel, the Newman representation along the nitrogen atom of the amino group and the adjacent carbon atom is shown. In order to highlight the orientation of the NH_2 moiety, only a portion of the isoxazole ring is represented, and the wavy lines are used to refer to the remaining part.

*The contents of this Chapter are taken from Ref. 280 after re-elaboration.

parameters have been optimized at each step. The minimum energy structure resulting from this analysis is shown in Figure 12.1. In particular, in the right panel of Figure 12.1, the Newman representation through the nitrogen atom of the amino group and the adjacent carbon atom of the ring is shown. For the sake of clarity, only a portion of the ring is represented, and wavy lines are used to refer to the remaining part. From this representation, is evident that the optimized geometry belongs to the C_1 symmetry point group: indeed, while the aminic nitrogen atom lies on the plane defined by the isoxazole ring, the two hydrogen atoms are out of that plane. This orientation is similar to that of several other molecules featuring a planar ring and an amino group, *e.g.*, 2-aminopyridine,[281,282] 3- and 4-aminopyridine,[283] 2-aminooxazole,[284] and aniline.[285–289] For these molecules, the angle between the bisector of the HNH angle and the extension of the C–NH_2 bond (which is a useful measure of the degree of non-planarity of the amino group) ranges from 27° to 42°, with the largest value being observed in 2-aminooxazole. For 3-aminoisoxazole, this angle is even larger: 45°.

In the rotational spectroscopy studies of all the molecules mentioned above, all rotational transitions are split into doublets due to the inversion motion of the NH_2 group; hence, the same behavior is expected for 3-aminoisoxazole. For details on molecular inversion states, the reader is referred to Section 2.5. In this case, the corresponding double-well potential energy surface has been evaluated at the revDSD level (see Figure 12.2), indeed confirming the presence of two equivalent structures and the planarity of the transition state.

Subsequently, to further improve the energetics, the jun-ChS composite scheme has been employed on top of the revDSD optimized geometries. An inversion barrier of 5.1 kJ mol^{-1} (2.4 kJ mol^{-1} if the harmonic ZPE contribution at the revDSD level is added) has been determined, in general agreement with the values found for 2-aminooxazole[284] (5.1 kJ mol^{-1}) and aniline[289] (6.4 kJ mol^{-1}). For both 3- and 4-aminopyrdine, values in the 3–5.4 kJ mol^{-1} range have also been estimated.[283] Therefore, considering the values of the computed inversion barrier, it is anticipated that the splitting in the rotational transitions should be also observed in 3-aminoisoxazole.

This being the first attempt to record the rotational spectrum of 3-aminoisoxazole, the accurate prediction of the spectroscopic parameters is required. To this aim, the revDSD geometries have been refined by resorting to the jun-ChS composite scheme. Both computed geometries are reported in Table A.10. Although the accuracy of the jun-ChS method is well documented (see Table 4.1), it is possible to compare the optimized

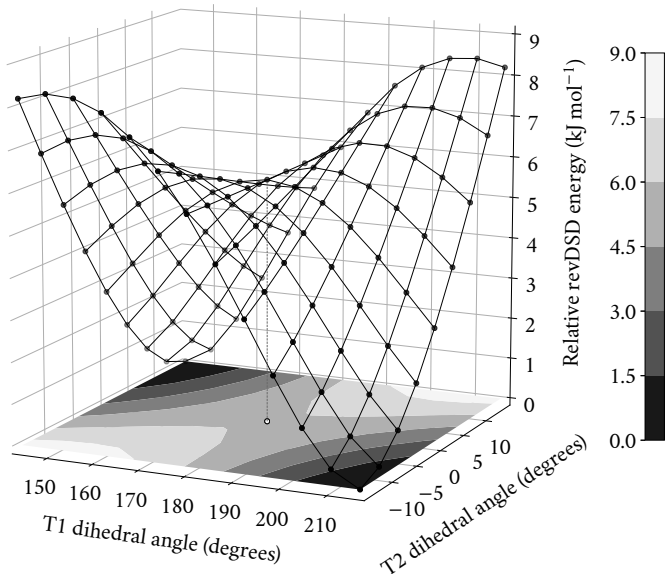


Figure 12.2: Representation of a portion of the PES of 3-aminoisoxazole, calculated *via* a relaxed scan at the revDSD level. The dihedral angles are defined according to the Z-matrix in Table A.10. On the grayscale coded plane, the white circle indicates the point where T1 and T2 dihedral angles have values of 180° and 0°, respectively. In this point, 3-aminoisoxazole has a totally planar structure, and this corresponds to the inversion transition state (the saddle point on the overlying surface).

geometry with the SE equilibrium structure of isoxazole, based on the assumption that the planar ring is essentially unchanged when moving from isoxazole to 3-aminoisoxazole. As a matter of fact, the observed average discrepancy is found to be 3 mÅ for bond lenghts and 0.1° for angles, thus confirming the accuracy and reliability of the jun-ChS equilibrium geometries.

Finally, the straightforwardly derived jun-ChS equilibrium rotational constants have been augmented by the vibrational corrections at the revDSD level. The electric dipole moment components as well as quartic and sextic centrifugal distortion constants have been calculated at the revDSD level. The equilibrium nitrogen quadrupole coupling constants have been determined exploiting the jun-ChS approach, and they have been vibrationally corrected at the revDSD level. These spectroscopic parameters are reported in Table 12.2.

12.2 Experimental study

From a spectroscopic point of view, 3-aminoisoxazole (Figure 12.1) is a nearly-prolate asymmetric top with $\kappa = -0.69$. A strong b-type spectrum combined with a less intense a-type one is expected to be registered, according to the values of the molecular electric dipole components ($\mu_b = 3.08$ D and $\mu_a = 1.19$ D, see Table 12.2). Moreover, the spectrum is expected to show a complicated hyperfine structure, due to the presence of two quadrupolar nuclei. In addition, as already discussed above and suggested by similar spectroscopic studies,[281–289] two inversion states are expected because of the large amplitude motion of the NH_2 moiety.

By relying on the computational prediction of the rotational spectrum, a few $K_a = 0, 1$ transitions have been successfully identified in the COBRA-FTMW experiment. Their assignment allowed to refine the rotational constants, and, with an iterative procedure, the analysis of the 6–24 GHz spectrum has been completed, thereby analyzing more than 200 hyperfine components of both a- and b-type transitions. An example of the hyperfine structure is reported in Figure 12.3 for the $3_{1,3} \leftarrow 2_{0,2}$ transition. Thanks to the excellent experimental resolution (*ca.* 2 kHz) achievable with the COBRA-FTMW spectrometer and its high sensitivity, 28 different hyperfine components of this transition have been assigned (26 of which are shown in Figure 12.3, see Table 12.1). Indeed, the uniqueness of the hyperfine structure

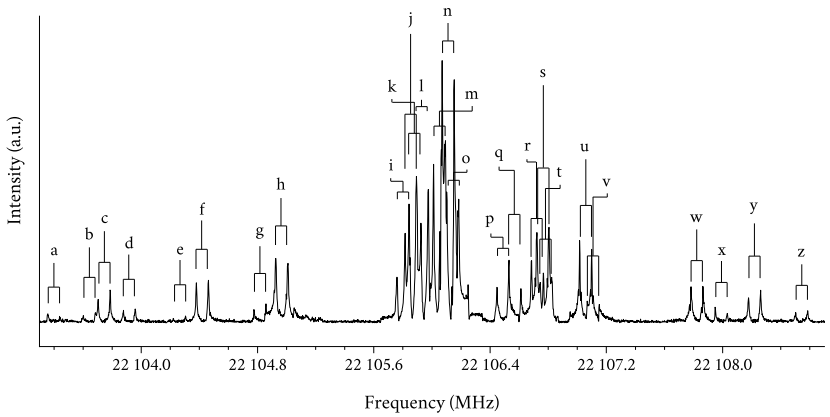


Figure 12.3: Hyperfine structure of the $3_{1,3} \leftarrow 2_{0,2}$ transition of 3-aminoisoxazole recorded using the COBRA-FTMW spectrometer. The 7 MHz scan has been registered with 10 000 shots every 100 kHz. All lines appear as doublets due to Doppler effect. The labels refer to the assigned hyperfine components, which are collected in Table 12.1.

Table 12.1: Hyperfine components of the $3_{1,3} \leftarrow 2_{0,2}$ transition of 3-aminoisoxazole. The quantum numbers are defined in Equation (12.2), while the letters refer to the labeling used in Figure 12.3. The observed frequencies are expressed in MHz.

F_1	F	F_1'	F'	obs. ν[a]		F_1	F	F_1'	F'	obs. ν[a]	
2	1	2	2	22 102.7616	–	4	5	3	4	22 106.1066	n
2	3	2	2	22 103.3964	a	3	4	3	3	22 106.1440	o
4	3	2	3	22 103.6426	b	4	3	1	2	22 106.4889	p
3	2	2	2	22 103.7442	c	3	3	2	3	22 106.5718	q
2	3	2	3	22 103.9182	d	2	1	1	0	22 106.7267	r
3	2	2	3	22 104.2622	e	2	3	1	2	22 106.7649	s
3	2	2	1	22 104.4204	f	3	2	1	1	22 106.7854	t
4	3	3	3	22 104.8174	g	4	4	3	3	22 107.0592	u
3	4	2	3	22 104.9667	h	3	2	1	2	22 107.1113	v
2	1	1	1	22 105.8016	i	4	4	3	4	22 107.8255	w
2	2	2	1	22 105.8548	j	2	2	3	2	22 107.9924	x
4	4	2	3	22 105.8839	l	2	2	1	1	22 108.2224	y
4	3	3	2	22 105.9353	k	2	2	1	2	22 108.5464	z
3	3	2	2	22 106.0530	m	3	3	3	2	22 108.8643	–

[a] The experimental uncertainty has been set to 2 kHz.

has been of great help in the identifications of the transitions belonging to 3-aminoisoxazole. However, no evidence of the second inversion state has been found in the FTMW experiment.

The measurements in the cm-wave region allowed to determine the ground state rotational, quartic centrifugal distortion, and nitrogen quadrupole coupling constants. From that, the rotational spectrum of 3-aminoisoxazole has been predicted in the millimeter-wave region, without, however, accounting for the hyperfine structure, as represented in the top panel of Figure 12.4. The choice of neglecting the contribution of quadrupole coupling is based on the expectation that the hyperfine structure is much smaller than the line width in that frequency region. In fact, taken the $15_{0,14} \leftarrow 14_{1,14}$ transition as example (which is the lowest in the frequency range considered in Figure 12.4), the splitting due to quadrupole coupling is determined to be *ca.* 50 kHz, to be compared with an observed linewidth HWHM of 150–200 kHz in the mm-wave region.

From the inspection of Figure 12.4, it is clear that the prediction of the mm-wave transitions based on FTMW data is affected by negligible errors, especially for the $K_a = 0, 1$ a- and b-type transitions, on which the impact of the uncertainties is small. Consequently, the spectral assignment has been rather straightforward; however, as shown in the bottom panel of Figure 12.4, a second set of lines sharing a remarkably similar pattern is

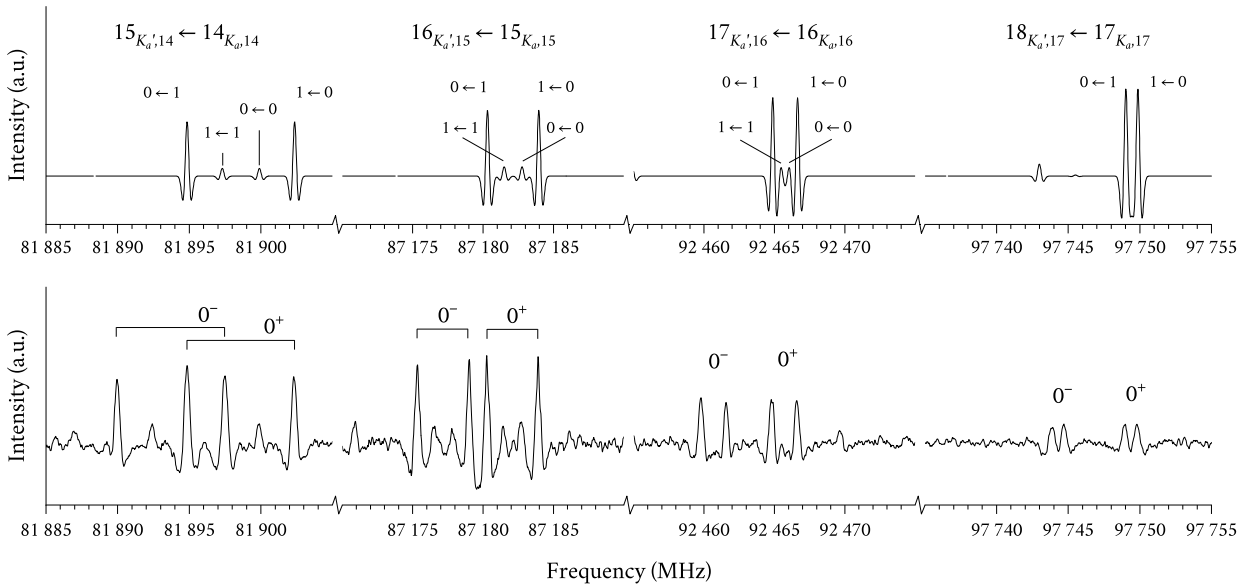


Figure 12.4: Simulated (based on FTMW results, top panel) and experimental (FM-mmW, bottom panel) typical pattern of the $K_a = 0, 1$ transitions of 3-aminoisoxazole in the millimeter-wave region. In the bottom panel, the two inversion states are easily distinguishable.

immediately evident. This set of transitions has been identified as belonging to the 0^- inversion state. After a first estimate of the spectroscopic parameters for the 0^- state and the inclusion of the measurements in the 80–115 GHz frequency range for the 0^+ state in the original cm-wave set, the characterization has been extended up to 320 GHz. No *c*-type transitions have been recorded (μ_c = 1.09 D), and the reason for this is explained later. In total, more than 1 100 and 900 lines have been recorded and assigned for the 0^+ and 0^- states, respectively, involving energy levels with J values up to 68 and K_a values up to 25.

At this point, the identification of the 0^- state in the mm-wave spectrum deserves a note. While the centimeter and millimeter/submillimeter spectra appear to have contradictory results, the profound different physical conditions need to be considered. In the FM-mmW experiment, the measurements are carried out at room temperature, thus populating both inversion states. On the contrary, in the FTMW experiment, a strong rotational cooling follows the pulsed supersonic expansion in the cavity, therefore depopulating the higher-energy states (*i.e.*, the 0^- state in the present case).

While the global rotational Hamiltonian of the 0^- state is reported in Equation (2.58), the one for the 0^+ state can be expressed as

$$\hat{\mathcal{H}}_{\text{rot}} = \hat{\mathcal{H}}^{S}_{\text{rot}} + \hat{\mathcal{H}}^{\text{r}}_{\text{hfs}} + \hat{\mathcal{H}}^{\text{a}}_{\text{hfs}}\,, \tag{12.1}$$

where $\hat{\mathcal{H}}^{S}_{\text{rot}}$ is that of Equation (2.58), while $\hat{\mathcal{H}}^{\text{r}}_{\text{hfs}}$ and $\hat{\mathcal{H}}^{\text{a}}_{\text{hfs}}$ are the hyperfine-structure Hamiltonians for the nitrogen atom of the isoxazole ring (referred

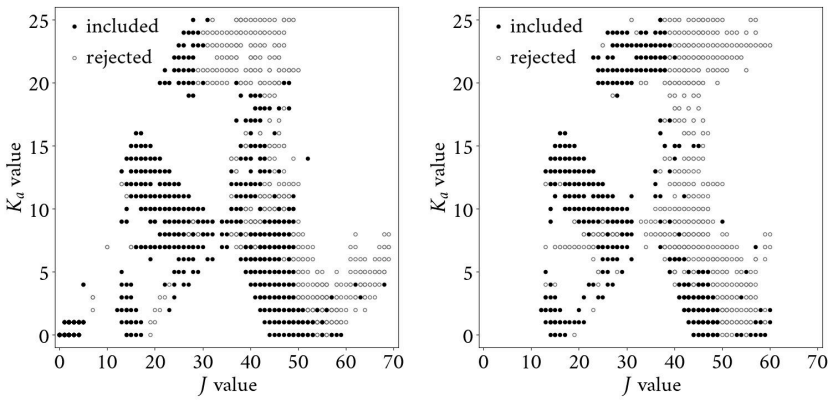


Figure 12.5: Graphical representation of the included and rejected transitions of the 0^+ (left panel) and 0^- (right panel) states of 3-aminoisoxazole in the fitting procedure.

Table 12.2: Ground-state spectroscopic parameters of the 0^+ and 0^- states of 3-aminoisoxazole (*S*-reduction, I^r representation). Values in parentheses denote one standard deviation and apply to the last digits of the constants.

		3-aminoisoxazole		
		exp. 0^+	exp. 0^-	theo.[a]
A	MHz	9 356.3832(1)	9 353.3440(3)	9 365.2
B	MHz	3 669.78547(7)	3 667.5030(1)	3 673.9
C	MHz	2 642.30281(5)	2 642.2541(1)	2 645.9
D_J	kHz	0.23386(1)	0.22837(4)	0.218
D_K	kHz	1.6012(2)	1.5558(4)	1.49
D_{JK}	kHz	1.5376(3)	1.5724(3)	1.53
d_1	Hz	−77.064(8)	−76.72(1)	−72.2
d_2	Hz	−21.705(4)	−24.063(7)	−21.5
$\frac{3}{2}\chi^{r\ b}_{aa}$	MHz	8.640(2)		8.90
$\frac{1}{4}(\chi^r_{bb} - \chi^r_{cc})$	MHz	−0.7448(7)		−0.761
$\frac{3}{2}\chi^a_{aa}$	MHz	3.803(3)		3.63
$\frac{1}{4}(\chi^a_{bb} - \chi^a_{cc})$	MHz	1.3939(6)		1.41
$\lvert\mu_a\rvert$	D	Y[c]	Y	1.19
$\lvert\mu_b\rvert$	D	Y	Y	3.08
$\lvert\mu_c\rvert$	D	N	N	1.09
# lines		623	529	
J^{max}, K_a^{max}		68, 25	60, 24	
rms error	kHz	40.4	44.8	
σ		1.47	1.50	

[a] Jun-ChS equilibrium rotational constants augmented by vibrational corrections at the revDSD level. Quartic centrifugal distortion and dipole moment components at the revDSD level. Equilibrium nuclear quadrupole coupling constants at the jun-ChS level augmented by vibrational corrections at the revDSD level. [b] The superscript "r" refers to the nitrogen atom in the isoxazole ring, while "a" to the one in the amino group. [c] For each μ_g dipole moment component, "Y" and "N" refer to detected and non-detected g-type transitions, respectively.

to as "r") and that of the amino group (indicated with the superscript "a"), respectively (according to Equation 2.65). The interactions of their electric quadrupole moments with the electric field gradient at the corresponding nuclei lead to the definition of two additional quantum numbers, with the coupling scheme

$$\begin{aligned} \mathbf{F}_1 &= \mathbf{J} + \mathbf{I}_{\mathrm{N}}^{\mathrm{r}} ; \\ \mathbf{F} &= \mathbf{F}_1 + \mathbf{I}_{\mathrm{N}}^{\mathrm{a}} . \end{aligned} \tag{12.2}$$

The fitting procedure was rather challenging, and a graphical representation of the encountered difficulties is depicted in Figure 12.5, which shows the included (black dots) and rejected (white dots) lines in terms of their J and K_a values. All recorded lines could be assigned without any doubt, thanks to both the accuracy of the spectroscopic parameters (*cf.* Table 12.2) and the absence of strong interfering lines in the spectrum. Indeed, the first vibrationally excited state (NH_2 wagging) lies at about 266 cm^{-1} (harmonic value calculated at the revDSD level of theory) and its population at room temperature should be around 28% of the ground state one, according to Boltzmann distribution. However, while *ca.* 1 300 lines could be well reproduced in the fitting procedure, more than 600 transitions (all in the 80–320 GHz range) showed a discrepancy between the observed and calculated frequencies which was more than four times the experimental accuracy (set to 2 kHz for the FTMW experiment and 30 kHz for the FM-mmW one). After several attempts to include them, all these lines have been excluded from the fit.

This behavior is likely due to a Coriolis interaction between the two inversion states: due to the orientation of the electric dipole components, *a*- and *b*-type transitions occur within each state, while the *c*-type ones connect the 0^+ and 0^- states. Indeed, the *c* axis is perpendicular to the isoxazole ring (*cf.* Figure 12.1), and its sign changes upon inversion. Therefore, these transitions are a great tool for the determination of the energy difference between the two states, which would be of great help in the description of this interaction. However, the energy difference may be rather high: as a matter of fact, no *c*-type transitions have been observed in the entire mm-wave range. For instance, for vinylamine,[290] the Coriolis interaction between the inversion states could be described thanks to the recording of the $0^- \leftarrow 0^+$ band, and the energy difference was determined to be 45.4871(74) cm^{-1}. Instead, for 2-, 3-, and 4-aminopyridine, aniline, and 2-aminooxazole, the energy separation was only estimated to be lower than 100 cm^{-1}, thus leading to *c*-type transitions taking place the far-infrared region.

In the present case, the two states had to be fitted separately with a

model that cannot include the energy separation nor the Coriolis interaction. As a consequence, a significant number of transitions could not be incorporated in the analysis. Taking a closer look at Figure 12.5, it appears that the 0^+ state behave globally better with respect to the 0^- state. In fact, for the latter, even low-K_a transitions above $J = 50$ have been discarded, together with those with high values of both J and K_a. Instead, for both states, no strong deviations were found for transitions with J values below 40 (except 0^- lines with $K_a = 7, 8$, which have been almost completely excluded).

The spectroscopic parameters determined by the fitting procedure are collected in Table 12.2 together with the computed values. The agreement between theoretical and experimental parameters is evident: the average relative difference on the rotational constants is found to be *ca.* 0.1%. The discrepancies for quartic centrifugal distortion and nitrogen quadrupole coupling constants are larger, but on average are below 5%. Owing to the inclusion in the fit of both *a*- and *b*-type transitions, all rotational constants have been accurately determined, with the resulting relative error being about 2 parts per 10^7, while a precision of 2 parts per 10^4 has been obtained for the quartic centrifugal distortion constants. The inclusion of the sextic centrifugal distortion constants in the fit parameters did not provide any significant improvement to the analysis, and therefore none of them have been determined, and their values have been kept fixed to zero.

However, the rms error of the fit reflects the lack in the global Hamiltonian of terms describing the interaction between the two inversion states. Indeed, since the standard deviation of the fit is close to 1.5, the experimental accuracy of the retrieved frequencies has not been well reproduced. To conclude, a comparison of these results with respect to those of similar molecules studied in the literature is deserved. For all the species previously mentioned (*i.e.*, 2-aminopyridine,[281,282] 3- and 4-aminopyridine,[283] 2-aminooxazole,[284] and aniline[285–289]), the measurements have been carried out below 110 GHz; indeed, the difficulties in our analysis arose above this frequency value. A comprehensive ro-vibrational work in the far-infrared region on all the species bearing the NH_2 moiety in combination with a planar ring is thus desirable to fill the lack of information in their spectroscopic characterization. In particular, for 3-aminoisoxazole, this would allow to overcome the fitting problems related to the neglect of the Coriolis interaction, which led to the scarcely-conditioned description of the system.

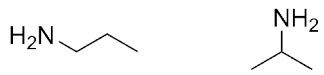
Chapter 13

iso- and n-Propylamine

In addition to the prebiotic interest of primary amines (*cf.* Section 1.3), thanks to the rich structural variety of propylamine isomers (C_3H_9N, see Figures 13.1–13.3), these species can be considered a valid test case for understanding how substituents affect the stability of isolated primary amines and, consequently, their chemical properties. In this Chapter,* the results of a joint spectroscopic and computational investigation of iso-propylamine (iPA) and n-propylamine (nPA) conformers are reported, aiming to provide a detailed spectroscopic characterization for astrochemical purposes.

13.1 Conformational analysis

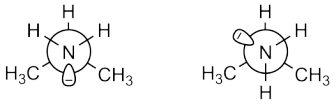
An accurate computational study on the two structural isomers of propylamine, namely iPA and nPA (shown in Figure 13.1), has been carried out to provide a reliable guide to the experiment. At first, a preliminary study of the PES has been carried out at the B3S level of theory; this pointed out the existence of two and five stable conformations for iPA and nPA, respectively. After identification of the minima, small portions of the



n-propylamine (nPA) iso-propylamine (iPA)

Figure 13.1: Molecular structure of propylamine isomers.

*The contents of this Chapter are taken from Ref. 291 after re-elaborations.

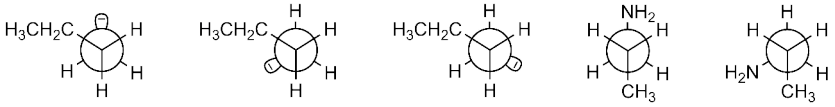


t-iPA *g*-iPA

Figure 13.2: Newman projections of iso-propylamine conformers. The letters *t* and *g* refer to *trans* and *gauche*, respectively. The structure of the conformational transition state is at the eclipsed configuration which converts one minimum into the other by rotation along the N–C bond.

PES have been investigated in order to determine the structures of the connecting transition states. Subsequently, all the stationary points geometries have been refined using a double-hybrid functional combined with a minimally-augmented triple-ζ basis set, namely the B2PLYP-D3(BJ)/maug-cc-pVTZ-dH level (from here on referred to as mB2-dH).

From a thorough analysis of the PES of the propylamine isomers, it results that the branched isomer, *i.e.*, iPA, can exist in the *trans* and *gauche* (doubly degenerated) conformational forms, hereafter labeled as *t*-iPA and *g*-iPA and represented in Figure 13.2. Instead, the linear isomer nPA (where the NH_2 moiety is at one end of the carbon chain) is significantly more flexible, thus it shows a greater number of stable conformers (*tt*, *tg*, *gg*, *gg′*, and *gt*, as demonstrated by low-resolution Raman and infrared investigations[292–294]), which are graphically displayed in Figure 13.3. In the labeling of the conformers, the first letter refers to the conformation of the NC–CC dihedral angle—with *t* and *g* referring to *trans* and *gauche*, respectively—while the second letter refers to the value of the ∶N–CC dihedral angle. For the latter, starting from the *trans* position, the clockwise and counterclockwise rotations of the NH_2 group by 120° are indicated by *g* and *g′*, respectively, as shown in Figure 13.3. All the optimized geometries



xg′-nPA *xg*-nPA *xt*-nPA *tx*-nPA *gx*-nPA

Figure 13.3: Newman projections of n-propylamine conformers. The letters *t* and *g* refer to *trans* and *gauche*, respectively, while *x* refer to any of them. The structures of each conformational transition state coincides with the eclipsed configuration thorough which a minimum is converted into the other by rotation along the N–C or NC–CC bonds.

have been collected in Tables A.11 for the minima and in Table A.12 for the conformational transition states. The structures of the latter have not been graphically reported for the sake of conciseness, but they can be retrieved at the eclipsed configurations which convert one minimum into the other.

Exploiting the mB2-dH geometries as reference, the energetic of the systems has been evaluated at the ChS level, while the mB2-dH level has been employed also for harmonic ZPE corrections. For the iPA conformers, a barrier of 10.7 kJ mol^{-1} (9.2 kJ mol^{-1} if harmonic ZPE at the mB2-dH level is included) is found, while the energy difference between the conformers is 2.0 kJ mol^{-1} (1.8 kJ mol^{-1} with harmonic mB2-dH ZPE). A high barrier and a small energy difference should favor the spectral observation of both isomers.

The energetic characterization of nPA is summarized in Figure 13.4, from this it is evident that the *tt* conformer is the most stable one, followed by *tg*, which is only 0.2 kJ mol^{-1} higher in energy. The *gx* conformers (with *x* referring to either *t*, *g* or *g'*) lie within 2 kJ mol^{-1} with respect to *tt*:

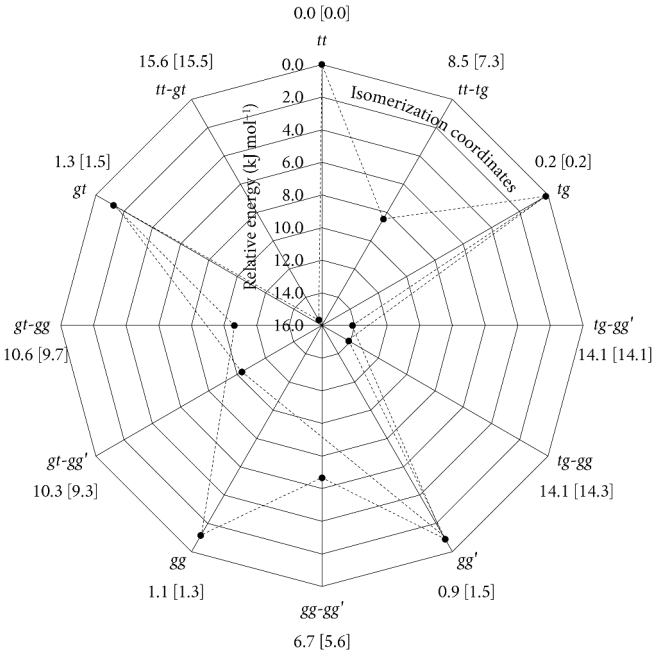


Figure 13.4: Schematic representation of the relative ChS energies of the n-propylamine conformers at the mB2-dH geometry, including the conformational transition states. The ChS energies augmented by the harmonic ZPE correction at the mB2-dH level are reported in parentheses. All values are in kJ mol^{-1}.

1.3 kJ mol^{-1} for *gt*, 1.1 kJ mol^{-1} for *gg*, and 0.9 kJ mol^{-1} for *gg′*. Consequently, the population at room temperature of nPA is spread over all conformers, which are therefore expected to be abundant enough to be observed in the spectrum.

Finally, concerning the spectroscopic characterization, the ChS model has been employed for obtaining accurate equilibrium geometries (see Table A.11) and, thus, the equilibrium rotational constants; the vibrational corrections have been evaluated at the B3S level. Quartic centrifugal distortion constants, dipole moments and nuclear quadrupole coupling constants have been evaluated at the mB2-dH level of theory, while the sextic centrifugal distortion terms have been determined at the B3S level as a by-product of the anharmonic calculations. The quantum-chemical spectroscopic parameters are collected in Table 13.1 for iPA, and in Tables 13.2 and 13.3 for the nPA conformers.

13.2 Experimental study

Rotational spectra have been recorded by scholars at the School of Chemistry and Chemical Engineering of the Chongqing University (China), in the group led by Prof. Dr. Qian Gou, between 2 and 22 GHz with a COBRA-type FTMW spectrometer and, then, in Bologna, using the FM-mmW spectrometer (described in Section 6.2) in the 80–400 GHz range. The laboratory conditions are described in Section 6.4.

iso-Propylamine, iPA

Both conformers of iPA (Figure 13.2) have been identified in a low-resolution study reporting the Raman spectra in gas, liquid and solid phases.[295,296] Instead, only the rotational spectrum of the most stable *t*-iPA has been previously observed.[227,228] In this work, its laboratory investigation has been extended, while attempts to detect the elusive *g*-iPA conformer have been carried out without success.

The *trans* isomer is a nearly oblate asymmetric-top rotor with $\kappa = 0.81$, belonging to the C_s point group. Its permanent electric dipole moment almost completely lies along the *c* axis, while μ_a is null by symmetry. As a consequence, the rotational spectrum of *t*-iPA is dominated by *c*-type transitions, whereas the *b*-type ones are expected to be roughly 150-times weaker. The rotational energy levels of *t*-iPA may be described using Watson's semi-rigid Hamiltonian either in the *S*- or *A*-reduction. The latter has been employed in quantum-chemical calculations as well as in

the spectral analysis to facilitate the comparison with previous data. The rotational Hamiltonian can be summarized as the sum of two terms

$$\hat{\mathcal{H}}_{\mathrm{rot}} = \hat{\mathcal{H}}^{A}_{\mathrm{rot}} + \hat{\mathcal{H}}_{\mathrm{hfs}} \,, \tag{13.1}$$

where $\hat{\mathcal{H}}^{A}_{\mathrm{rot}}$ is given in Equation (2.57) and $\hat{\mathcal{H}}_{\mathrm{hfs}}$ is the hyperfine-structure Hamiltonian (see Equation 2.65) due to the presence of a nitrogen atom, whose nuclear spin is $I_{\mathrm{N}} = 1$. The coupling scheme of Equation (2.61) has been used.

Concerning *t*-iPA, some rotational lines below 50 GHz have already been measured using a Stark modulation spectrometer[228] and a waveguide FTMW spectrometer,[227] with the hyperfine splittings resolved for many transitions in both experiments. However, due to the limited frequency and quantum number ranges, the centrifugal analysis was not well constrained (see the "previous results" column in Table 13.1).

In this work, the spectrum of *t*-iPA has been extended up to *ca.* 400 GHz, recording transitions between rotational energy levels with J and K_c maximum values as high as 30 and 20, respectively. A portion of the experimental spectrum is shown in Figure 13.5. While in the cm-wave experiment the hyperfine structure has been resolved, at higher frequency it has been partially solved only for ten high-K_c lines: all remaining transitions (more than 150) have been observed as blended into a single line. In addition, in the FTMW experiment, some *b*-type transitions have been measured despite their intrinsic low intensity. Indeed, the conditions of a supersonically expanded jet allowed to assign confidently five *b*-type lines of *t*-iPA. Conversely, their weakness made them impossible to be unequivocally identified in the FM-mmW experiment, even if predicted with good accuracy. The previous and the newly observed transitions have been fitted to the Hamiltonian of Equation (13.1). The accuracy is estimated to be 10–20 kHz for the millimeter- and submillimeter-wave regions and 3 kHz in the cm-wave region.

Thanks to the inclusion in the fit of both *b*- and *c*-type transitions, most of the spectroscopic parameters have been obtained with satisfactory accuracy. The results are collected in Table 13.1 together with the computed values. A rms error of 14 kHz together with a standard deviation σ below unity demonstrate that the fit was able to reproduce the experimental accuracy. The comparison points out a good agreement between theoretical and experimental data, with discrepancies on the rotational constants and centrifugal distortion constants within 0.1% and 5%, respectively. The uncertainties on the experimental rotational constants are smaller than 1 kHz, while the quartic centrifugal distortion terms have been evaluated

Table 13.1: Ground-state spectroscopic parameters of the *trans* and *gauche* conformers of iso-propylamine (*A*-reduction, III^l representation). Values in parentheses denote one standard deviation and apply to the last digits of the constants.

		t-iPA			*g*-iPA
		exp.[a]	theo.[b]	prev.[c]	theo.[b]
A	MHz	8 331.9030(2)	8 321.7	8 331.905(2)	8 173.6
B	MHz	7 977.3355(2)	7 982.7	7 977.334(2)	7 958.4
C	MHz	4 656.9166(6)	4 652.2	4 656.989(8)	4 658.1
Δ_J	kHz	7.1659(5)	7.26	7.2(3)	7.02
Δ_{JK}	kHz	−11.865(2)	−12.1	−12.17(5)	−10.9
Δ_K	kHz	5.6(1)	5.81	0.0[d]	4.92
δ_J	kHz	−0.14925(9)	−0.142	−0.1480(6)	−0.120
δ_K	kHz	9.227(6)	9.40	8.38(8)	7.48
Φ_J	mHz	11.3(5)	11.3		−36.2
Φ_{JK}	Hz	−0.61(3)	−0.596		−0.406
Φ_{KJ}	Hz	1.9(1)	1.88		1.50
Φ_K	kHz		−1.30		−1.06
ϕ_J	mHz		−2.21		−1.94
ϕ_{JK}	Hz	0.077(7)	0.0950		0.0324
ϕ_K	Hz	−6.7(4)	−6.73		−8.32
χ_{aa}	MHz	1.789(2)	2.06	1.788(4)	−3.39
χ_{bb}	MHz	2.569(2)	2.68	2.567(4)	2.24
$\lvert\mu_a\rvert$	D				1.15
$\lvert\mu_b\rvert$	D	Y[e]	0.07	0.10(4)	0.36
$\lvert\mu_c\rvert$	D	Y	1.30	1.19(3)	0.40
# lines		235		49	
J^{max}, K_c^{max}		30, 20		30, 7	
rms error	kHz	14.0		9.1	
σ		0.95		1.66	

[a] Missing values have been kept fixed at the theoretical counterpart in the fit procedure. [b] ChS equilibrium rotational constants augmented by vibrational corrections at the B3S level. Quartic centrifugal distortion and nuclear quadrupole coupling constants as well as dipole moment components at the mB2-dH level. Sextic centrifugal distortion constants at the B3S level. [c] Re-fitted from Refs. 227 and 228, which include only *c*-type transitions. Dipole moment components from Ref. 228. [d] Even if included in the original analysis by Keussen and Dreizler in Ref. 227 with a relative error of *ca.* 50% of its value, in this re-fit it has been kept to zero. [e] For each μ_g dipole moment component, "Y" refers to detected *g*-type transitions.

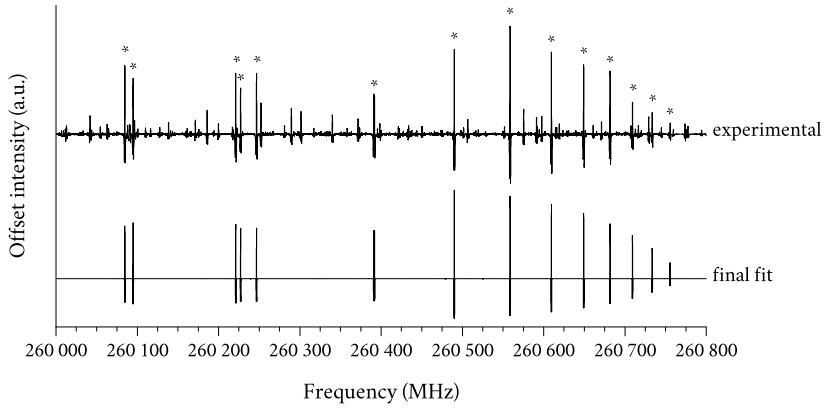


Figure 13.5: Comparison between the FM-mmW spectrum of *t*-iPA and the prediction based on the final spectroscopic parameters. The transitions marked with an asterisk have been included in the fit procedure.

with relative errors within 2%. In addition to a complete set of quartic centrifugal distortion constants, five out of seven sextic terms have been determined in the fit procedure.

Table 13.1 also includes the computed spectroscopic parameters for *g*-iPA (Figure 13.2), for which the rotational spectrum could not be assigned confidently, regardless of the apparently favorable energetic conditions. However, the detection of *g*-iPA isomers may be prevented by its being a nearly oblate asymmetric-top rotor with a medium-intensity *a*-type spectrum; in fact, this results in a sparse rotational spectrum lacking of characteristic patterns, therefore complicating the assignment of lines. Moreover, this behavior is also replicated by other small amines like aminoacetonitrile[297] and propargylamine:[298] both have been suggested to exist in *trans* and *gauche* forms by quantum-chemical studies or low-resolution experiments, but only their most stable *trans* forms have been revealed by high-resolution molecular spectroscopy.

n-Propylamine, nPA

The conformers of nPA (Figure 13.3) are nearly prolate asymmetric-top rotors ($\kappa = -0.98$ for *tx* conformers and $\kappa = -0.85$ for *gx* ones) with a total permanent dipole moment in the 1.2–2.0 D range. Only *tt*-nPA is of C_s symmetry, while all other conformers belong to the C_1 point group. The Hamiltonian used to describe the energy levels of nPA is analogous to that of Equation (13.1): given the absence of previous works for comparison, the

Table 13.2: Ground-state spectroscopic parameters of the *tx* conformers of n-propylamine (*S*-reduction, I^r representation). Values in parentheses denote one standard deviation and apply to the last digits of the constants.

		tt-nPA		*tg*-nPA
		exp.[a]	theo.[b]	theo.[b]
A	MHz	24 633.52(1)	24 616.8	25 043.5
B	MHz	3 687.5444(2)	3 690.5	3 732.8
C	MHz	3 473.8741(1)	3 477.2	3 491.1
D_J	kHz	0.81577(2)	0.815	0.821
D_{JK}	kHz	−2.1225(6)	−2.31	−2.56
D_K	kHz	47.8(7)	49.2	51.8
d_1	kHz	−0.06605(2)	−0.0664	−0.0779
d_2	Hz	0.665(7)	0.371	−1.87
H_J	mHz		0.176	0.173
H_{JK}	mHz		0.156	−1.12
H_{KJ}	Hz	−0.068(3)	−0.0744	−0.0873
H_K	Hz		0.279	0.273
h_1	mHz		0.0449	0.0428
h_2	µHz		0.133	3.29
h_3	µHz		−0.349	0.0694
χ_{aa}	MHz	−0.425(2)	−0.498	2.68
χ_{bb}	MHz	−1.367(3)	−1.58	0.130
$\lvert\mu_a\rvert$	D	Y[c]	0.98	0.05
$\lvert\mu_b\rvert$	D	Y	1.05	0.67
$\lvert\mu_c\rvert$	D			1.06
# lines		213		
J^{max}, K_a^{max}		47, 17		
rms error	kHz	10.4		
σ		0.83		

[a] Missing values have been kept fixed at the theoretical counterpart in the fit procedure. [b] ChS equilibrium rotational constants augmented by vibrational corrections at the B3S level. Quartic centrifugal distortion and nuclear quadrupole coupling constants as well as dipole moment components at the mB2-dH level. Sextic centrifugal distortion constants at the B3S level. [c] For each μ_g dipole moment component, "Y" refers to detected *g*-type transitions.

Table 13.3: Ground-state spectroscopic parameters of the *gx* conformers of n-propylamine (*S*-reduction, I^r representation). Values in parentheses denote one standard deviation and apply to the last digits of the constants.

		gt-nPA		*gg*-nPA		*gg′*-nPA	
		exp.[a]	theo.[b]	exp.[a]	theo.[b]	exp.[a]	theo.[b]
A	MHz	13 760.890(2)	13 784.0	13 901.575(1)	13 997.2	13 844.244(1)	13 894.4
B	MHz	4 873.886(1)	4 818.6	4 908.5029(7)	4 866.4	4 994.927(1)	4 975.7
C	MHz	4 149.318(2)	4 109.0	4 217.5655(6)	4 203.3	4 186.243(1)	4 179.3
D_J	kHz	4.25(7)	4.32	4.42(4)	4.29	4.46(4)	4.68
D_{JK}	kHz	−17.5(3)	−18.6	−16.8(1)	−17.4	−19.6(3)	−19.9
D_K	kHz		58.8		58.3		57.7
d_1	kHz	−1.55(6)	−1.29	−1.16(3)	−1.27	−1.60(4)	−1.48
d_2	kHz		−0.0947		−0.100		−0.114
χ_{aa}	MHz	−3.143(3)	−3.36	0.235(2)	0.269	1.996(3)	2.26
χ_{bb}	MHz	2.321(3)	2.46	−1.953(2)	−2.24	2.337(3)	2.51
$\lvert\mu_a\rvert$	D	Y[c]	1.27	Y	0.78	N	0.07
$\lvert\mu_b\rvert$	D	N	0.05	Y	1.01	Y	0.08
$\lvert\mu_c\rvert$	D	Y	0.61	Y	0.25	Y	1.23
# lines		35		49		26	
J^{max}, K_a^{max}		3, 1		3, 1		4, 2	
rms error	kHz	3.0		3.9		3.5	
σ		1.00		1.28		1.18	

[a] Missing values have been kept fixed at the theoretical counterpart in the fit procedure. [b] ChS equilibrium rotational constants augmented by vibrational corrections at the B3S level. Quartic centrifugal distortion and nuclear quadrupole coupling constants as well as dipole moment components at the mB2-dH level. Sextic centrifugal distortion constants at the B3S level. [c] For each μ_g dipole moment component, "Y" and "N" refer to detected and non-detected *g*-type transitions, respectively.

S-reduced form of Watson's Hamiltonian (see Equation 2.58) has been used, which usually requires a lower number of centrifugal distortion terms for a converged fit.

Analogously to iPA, the spectral assignment started from the FTMW experiment. The room temperature population of the various nPA conformers is retained inside the Fabry–Pérot cavity due to the high barriers that separate the minima (see Figure 13.4). Indeed, the rotational spectra of four conformers (*tt*, *gg*, *gg'*, and *gt*) could be assigned by relying on the computed rotational and nuclear quadrupole coupling constants. As an example, the $2_{1,2} \leftarrow 1_{1,1}$ transition of *tt*-nPA is shown in Figure 13.6, where the hyperfine structure is resolved. Unfortunately, the *tg* conformer was not detected, the reason being addressed later. Having refined the spectroscopic parameters, the study has been extended in the millimeter- and submillimeter-wave range using the FM-mmW spectrometer. In this experiment, only the spectrum of the *tt* conformer could be recorded and assigned up to 310 GHz. Both *a*- and *b*-type transitions have been measured with similar intensities, in accordance with the predicted dipole moment components, while no *c*-type transitions are observable by symmetry. The same uncertainties assigned to iPA transitions (3 kHz for FTMW data and 10–20 kHz for FM-mmW data) have been used.

The results of the fit procedure for *tt*-nPA are collected in Table 13.2, together with the computed parameters for the *tt* and *tg* conformers. Thanks to the assignment of more than 200 lines, the rotational constants, all quartic and one sextic (H_{JK}) centrifugal distortion terms, and the nuclear quadrupole coupling constants have been determined with high accuracy. An excellent agreement is found between theoretical and experimental outcomes, with the discrepancies on the rotational constants being smaller than 0.1%. A good agreement is noted also for the remaining terms, with the only exception being d_2, whose computed value is nearly half the experimental one. Excluding this outlier, the average absolute discrepancy is *ca.* 3%. The quality of this fit is demonstrated by the low standard deviation ($\sigma = 0.83$) and rms error (10 kHz): the chosen Hamiltonian describes appropriately the *tt*-nPA conformer.

The situation is different for the four remaining conformers. Although the *gx* conformers are easily identified in the FTMW spectrum, their rotational lines at higher frequency seem to deviate from the semi-rigid rotor approximation. The fact is not surprising, as it has already been observed in similar systems (*e.g.*, n-propanol from Ref. 299). In particular, the rotational spectrum of n-propanol has been measured in the 8–550 GHz range, and resonance effects due to Coriolis interactions between the conformers

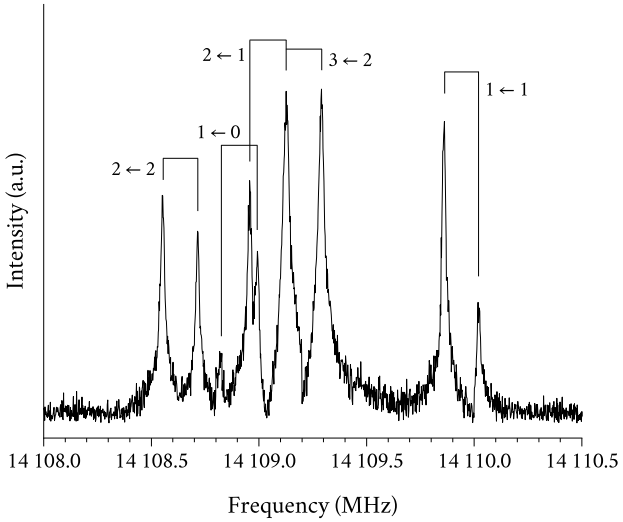


Figure 13.6: FTMW spectrum of the $2_{1,2} \leftarrow 1_{1,1}$ transition of *tt*-nPA showing the fully resolved hyperfine structure. $F \leftarrow F'$ quantum numbers are reported, and Doppler splittings due to the coaxial arrangement are shown.

were recognized.[299] Similarly to nPA, n-propanol can exist in five energetically close conformations—which are analogous to the nPA ones—whose rotational spectra can be coupled through Coriolis operators. In such a case, the Hamiltonian of Equation (13.1) cannot reproduce the rotational energy levels: a more appropriate Hamiltonian (developed for ethanol in Ref. 300) has been applied.[301] Nevertheless, it is beyond the scope of this thesis to perform a detailed analysis of the Coriolis interactions present in the ro-vibrational manifolds of nPA: therefore, for the *gx*-nPA family, the discussion is limited to the low-frequency spectra.

As anticipated, all the *gx*-nPA conformers (Figure 13.3) have been observed in the FTMW spectrum. Few tens of lines have been recorded and analyzed for each conformer, allowing the determination of the rotational constants, some quartic centrifugal distortion terms and the nuclear quadrupole coupling constants. The results are reported in Table 13.3 together with the computed counterparts. Again, a very good agreement between theoretical and experimental values is found: on average, the absolute deviation between the computed and experimental rotational constants is *ca.* 0.6%, and 6% for the centrifugal distortion terms.

The data collected in Table 13.3 provide a valid starting point for searching rotational transitions of the *gx*-nPA family at higher frequencies. To some extent, some spectral lines close to 80 GHz are easily assignable, but

their inclusion in the fit leads to anomalous centrifugal distortion terms, even if the discrepancy from the prediction is only a few MHz. These effects are evidence of the aforementioned perturbation effects.

Finally, the non-detection of the *tg*-nPA conformer appears to be puzzling. However, if the nPA system is assumed to behave similarly to n-propanol, the perturbation on the *tg* conformer is expected to be considerably larger than that in the *gx* series, and to occur even at the lowest J levels. As a matter of fact, this perturbation effect could also explain the anomalous value of d_2 determined for *tt*-nPA.

Chapter 14

Astrophysical Implications and Conclusions

This thesis collected the spectroscopic data of eight potential interstellar species, which are mandatory to guide their search in the interstellar medium. In addition to their characterization by means of rotational spectroscopy, an accurate computational study has been carried out for each species, in order to obtain a reliable support for the laboratory work and suitable constraints for the spectroscopic parameters that could not be experimentally determined. Moreover, this thesis included the definition, validation, and testing of the TM-SE(-LR) approach, which exploits datasets of combined experimental and theoretical data to provide accurate equilibrium geometries for medium- to large-sized systems.

Chapter 7 presented a benchmark study based on the availability of accurate semi-experimental equilibrium rotational constants for 21 molecular species to assess the reliability, robustness, and accuracy of the TM-SE(-LR) approach. Each of the selected molecules has been seen as formed by two fragments, whose accurate semi-experimental structures is available; the modifications taking place when moving from the isolated fragments to the molecular species of interest have been accunted for at the revDSD level of theory. This led to the definition of the TM-SE equilibrium structure. A further improvement has been obtained by correcting the inter-fragments parameters with the linear regression corrective terms available for the selected level of theory, resulting in the TM-SE-LR approach. For the TM-SE and the TM-SE-LR equilibrium rotational constants, the average relative deviation with respect to the accurate semi-experimental equilibrium rotational constants has been found to be *ca.* 0.3% and 0.2%, respectively. According to the detailed investigation of the equilibrium structures of

a subset of the original dataset (which included seven species whose accurate r_e^{SE} geometries are available), these discrepancies reflect structural differences of *ca.* 1 mÅ for bond lengths and 0.1° for angles.

Furthermore, the TM-SE(-LR) has been applied for the determination of accurate equilibrium rotational constants for six out of eight species investigated in this thesis, namely (*E*)- and (*Z*)-cyanovinylacetylene, allenylacetylene, (*E*)- and (*Z*)-phenylmethanimine, 2- and 3-furonitrile, and 3-aminoisoxazole. All species show a noticeable improvement when moving from the revDSD equilibrium rotational constants to the TM-SE counterparts when compared with semi-experimental equilibrium rotational constants. The agreement is further increased when the TM-SE-LR approach is exploited. In particular, for a subset of the investigated species (*i.e.*, the mono-substituted aromatic rings, which are characterized by a certaing degree of structural rigidity), the approach is found to perform superbly. Since the only quantum-chemical calculation required by the TM-SE(-LR) approach is the geometry optimization at the revDSD level of theory, this methodology can provide equilibrium structures with an accuracy close to that typical of sophisticated composite schemes based on the coupled cluster *ansatz* while retaining the cost-effectiveness of DFT methods. This represents an extremely promising tool for the accurate structural characterization of medium- to large-sized molecular systems, which are of current interest in rotational spectroscopy, Astrochemistry, and many other fields.

Chapters 8–13 concerned the computational and laboratory study of the eight species investigated in this thesis, namely allenylacetylene, cyanovinylacetylene, phenylmethanimine, 2- and 3-furonitrile, 3-aminoisoxazole, and iso- and n-propylamine. They have been selected by means of their relevance in Astrochemistry, as they are involved in different aspects of this broad research field, ranging from prebiotic chemistry to the formation of aromatic species. To date, most of these species have not been found in the interstellar medium, the only exceptions being (*E*)-cyanovinylacetylene and allenylacetylene, which have been identified thanks to measurements in the centimeter-wave range. Therefore, in order to enlighten the way towards their first detection, also exploiting the wide spectral coverage offered by ALMA (which opens the astronomical investigation at higher frequencies also for detected species), the rotational spectrum of each species has been recorded and analyzed up to the millimeter-/submillimeter-wave range. The results are presented and discussed in the corresponding Chapters. Overall, the obtained line catalogs satisfy the accuracy requirement needed for an astronomical search: to this purpose, a list of intense lines that

Table 14.1: List of intense rotational transitions suitable for an astronomical search for the investigated species, together with the upper state energy level. The calculated frequencies are expressed in MHz, while the energy in K.

Molecule	Transition[a]		calc. ν[b]	E_u
allenylacetylene	$28_{0,28} \leftarrow 27_{0,27}$		137 496.48(4)	96.6
(*E*)-cyanovinylacetylene	$38_{0,38} \leftarrow 37_{0,37}$		109 805.845(1)	102.9
(*Z*)-cyanovinylacetylene	$35_{0,35} \leftarrow 34_{0,34}$	$35 \leftarrow 34$	137 845.206(3)	120.7
	$35_{0,35} \leftarrow 34_{1,34}$			
	$35_{1,35} \leftarrow 34_{0,34}$			
	$35_{1,35} \leftarrow 34_{1,34}$			
(*E*)-phenylmethanimine	$38_{0,38} \leftarrow 37_{1,37}$	$38 \leftarrow 37$	92 741.868(4)	87.8
	$38_{1,38} \leftarrow 37_{0,37}$			
	$38_{0,38} \leftarrow 37_{1,37}$	$37 \leftarrow 36$	92 741.874(4)	
	$38_{1,38} \leftarrow 37_{0,37}$			
	$38_{0,38} \leftarrow 37_{1,37}$	$39 \leftarrow 38$		
	$38_{1,38} \leftarrow 37_{0,37}$			
(*Z*)-phenylmethanimine	$20_{0,20} \leftarrow 19_{0,19}$		48 986.75(3)	25.2
2-furonitrile	$35_{1,35} \leftarrow 34_{1,34}$		118 042.653(1)	103.3
	$35_{0,35} \leftarrow 34_{0,34}$		118 042.740(1)	
3-furonitrile	$35_{1,35} \leftarrow 34_{1,34}$		113 924.745(1)	99.7
	$35_{0,35} \leftarrow 34_{0,34}$		113 924.924(1)	
3-aminoisoxazole	$27_{0,27} \leftarrow 26_{0,26}$		145 301.750(2)	99.2
	$27_{0,27} \leftarrow 26_{1,26}$			
	$27_{1,27} \leftarrow 26_{0,26}$			
	$27_{1,27} \leftarrow 26_{1,26}$			
t-iso-propylamine	$21_{1,20} \leftarrow 20_{0,20}$	$21 \leftarrow 20$	342 213.85(1)	113.6
		$22 \leftarrow 21$	342 214.277(1)	
		$20 \leftarrow 19$		
tt-n-propylamine	$24_{0,24} \leftarrow 23_{1,23}$	$23 \leftarrow 22$	164 372.875(4)	102.1
		$25 \leftarrow 24$		
		$24 \leftarrow 23$	164 372.986(4)	
	$24_{0,24} \leftarrow 23_{0,23}$	$24 \leftarrow 23$	169 125.675(3)	
		$23 \leftarrow 22$	169 125.703(3)	
		$25 \leftarrow 24$		

[a] The column on the right denotes the $F \leftarrow F'$ quantum numbers. Blended hyperfine structures have not been reported for the sake of conciseness. [a] The values in parentheses denote the predicted uncertainties on the frequency. Missing values imply that the transition is blended with the precedent one.

are favorable candidates for detection have been reported in Table 14.1, together with the emission upper state energy level (E_u).

Moreover, the experimental spectroscopic parameters derived in Chapters 8–13 have been used to predict the rotational spectrum of each species up to 500 GHz. The simulations have been carried out at two different typical astronomical temperatures, namely $T = 30$ K and $T = 100$ K. The former temperature has been chosen to represent starless and dense core regions, such as TMC-1, while the latter is referred to the SFRs, the regions around protostars, and outflow shocks. The spectra have been collected in Figures 14.1 and 14.2 for all species except phenylmethanimine isomers, whose simulations have been reported in Figure 14.3. In fact, the latter predicted rotational spectra extend over a shorter frequency range (0–150 GHz for the *E* isomer and 0–50 GHz for the *Z* one) due to their limited experimental characterization (*cf.* Chapter 10). From the observation of Figures 14.1 and 14.2, it is noted that, for allenylacetylene, (*E*)-cyanovinylacetylene, and 2- and 3-furonitrile, at $T = 30$ K the spectra show the maximum intensity at *ca.* 100 GHz. Moving to $T = 100$ K, the maximum shifts at higher frequency (150–200 GHz). Analogously, for (*Z*)-cyanovinylacetylene, 3-aminoisoxazole, and *t*-iso-propylamine, the maximum of intensity is reported at *ca.* 200 GHz at $T = 30$ K and at 300–350 GHz when setting $T = 100$ K. Moreover, is it noted that (*Z*)-cyanovinylacetylene, 3-aminoisoxazole, and *tt*-n-propylamine show both an *a*- and *b*-type spectra of comparable intensity, while the simulations for the remaining species are dominated by either one of them. The only exception is *t*-iso-propylamine, whose *c*-type transitions are prevailing.

Chapter 8 proved the importance of the extension of the spectral investigation up to the millimeter-wave frequency domain to provide a reliable catalog of rest frequencies for allenylacetylene. In this case, centrifugal distortion effects have been found to be particularly prominent, as highlighted by the large value of the D_K constant (*ca.* 2 MHz, also confirmed by accurate *ab initio* calculations). This may indicate some degree of floppiness along the *a* principal axis, with the rigid allenyl and ethynyl moieties becoming more and more aligned as the centrifugal force increases. For this molecule, extrapolation to higher frequency from centimeter-wave measurements leads to completely unreliable predictions, thus pointing out the importance of the present work.

Chapter 9 presented the recording and analysis of almost 1 500 new rotational transtions for (*E*)- and (*Z*)-cyanovinylacetylene, which provide a robust base for the astronomical search for the *Z* isomer in the interstellar medium. Despite it being less stable than the *E* form (as proven by the

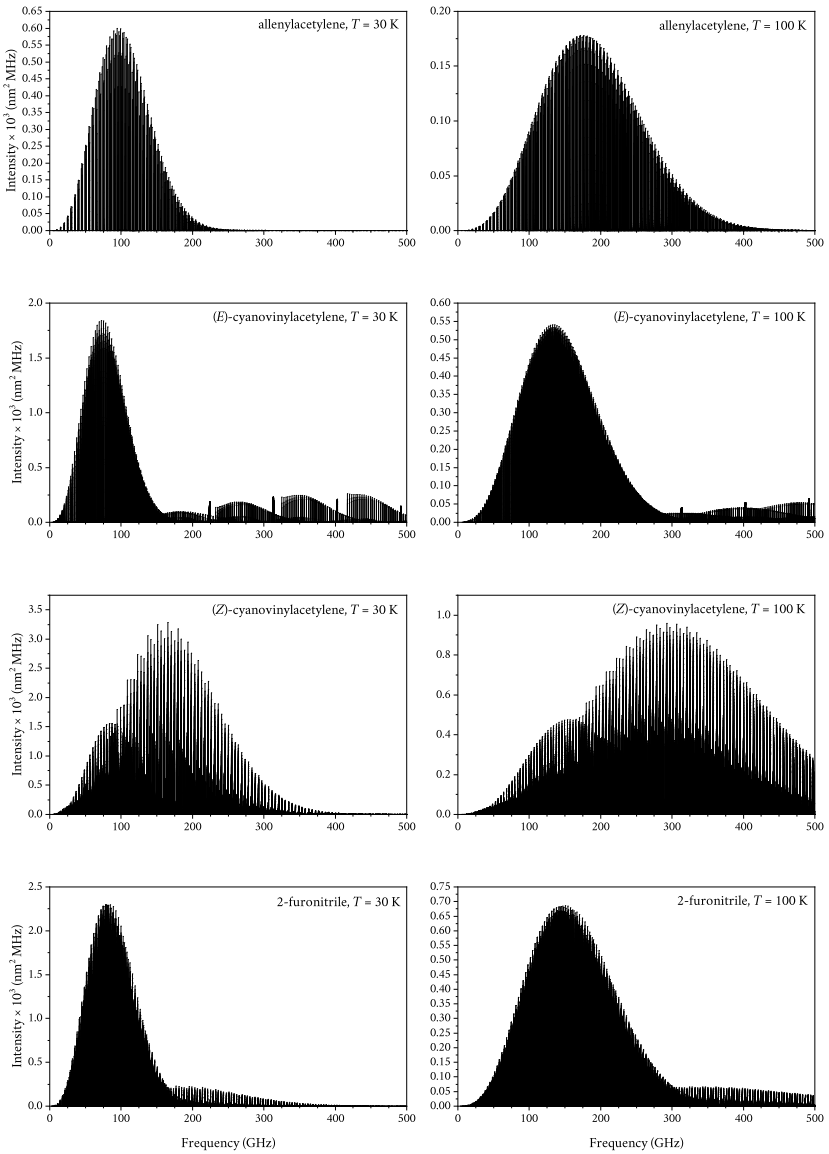


Figure 14.1: From top to bottom: simulated rotational spectrum of allenylacetylene, (*E*)-cyanovinylacetylene, (*Z*)-cyanovinylacetylene, and 2-furonitrile at 30 K (left column) and 100 K (right column).

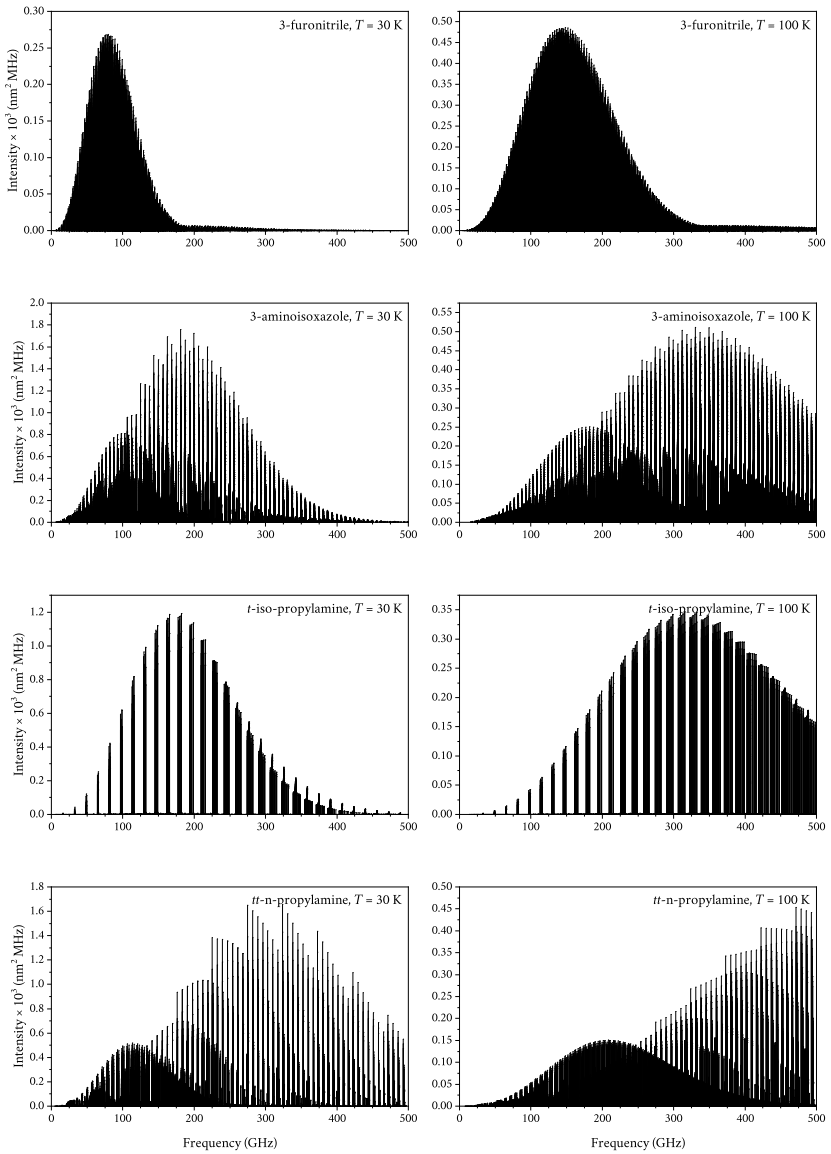


Figure 14.2: From top to bottom: simulated rotational spectrum of 3-furonitrile, 3-aminoisoxazole, *t*-iso-propylamine, and *tt*-n-propylamine at 30 K (left column) and 100 K (right column).

theoretical energetic characterization), its presence cannot be ruled out. Recent finding on imines[302] demonstrated that both geometrical isomers can coexist in cold interstellar gas, and pointed out the importance of quantum tunnelling in the *E*-*Z* isomerization process.

Subsequently, the detailed structural and energetic description of phenylmethanimine isomers provided in Chapter 10 led to their first laboratory identification. This work is a prerequisite for the possible radio astronomical detection of these species in the interstellar medium, relying on accurate rotational rest frequencies. In view of the strong chemical connection between benzonitrile and phenylmethanimine, an astronomical search in the TMC-1 region is suggested. In addition, a possible hydrolytic mechanism of phenylmethanimine formation by hydrobenzamide thermolysis has been studied in detail by means of rotational and nuclear magnetic resonance spectroscopy.

Chapter 11 concerned 2- and 3-furonitrile, which represent ideal candidates among the heterocycles family to be searched for in the interstellar medium. To this purpose, this work filled the lack of accurate spectral line catalogs for 3-furonitrile, a species whose rotational spectrum was com-

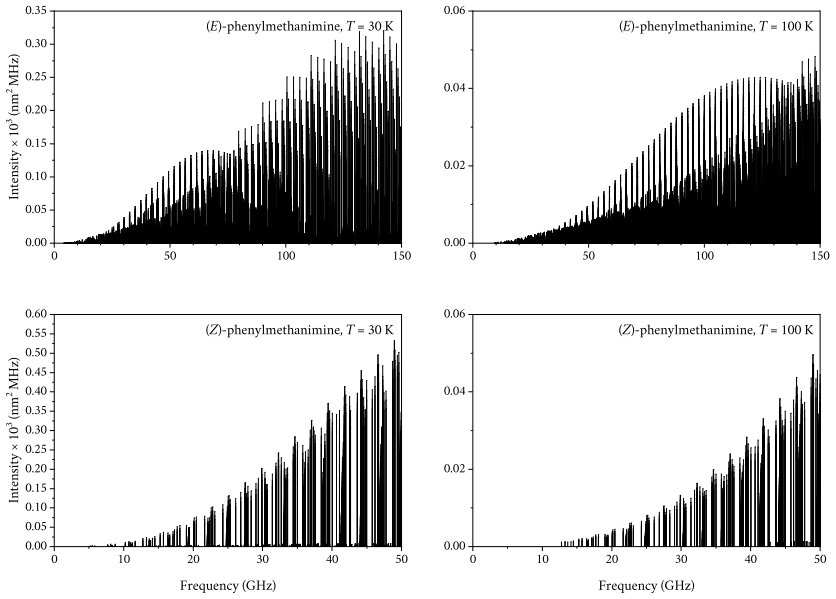


Figure 14.3: Simulated rotational spectrum of (*E*)-phenylmethanimine (top panels) and (*Z*)-phenylmethanimine (bottom panels) at 30 K (left column) and 100 K (right column).

pletely unknown. Moreover, the high-resolution measurements carried out in the centimeter-wave region allowed the determination of precise nitrogen quadrupole coupling constants for 2-furonitrile, thus adding an important piece of information towards an accurate and complete set of spectroscopic constants for this species. To date, the parameters reported in this thesis and in Ref. 107 enable the radio astronomical search for both furonitrile isomers in a wide frequency range and in different interstellar environments, such as cold and warm gases, turbulent and quiescent clouds, early-stage and more evolved objects. According to the energetic characterization and to the minimum energy principle,[263–265] 3-furonitrile is suggested to be the best candidate for detection.

Given its crucial prebiotic role, the first detection of 3-aminoisoxazole (treated in Chapter 12) in the interstellar medium would provide important insights into the unsolved question of the endogenous or exogenous origin of biomolecules on the early Earth. To guide the astronomical search, this work presented more than 500 accurate rest frequencies up to 320 GHz. In addition, the theoretical description of a portion of the potential energy surface close to the minimum energy structure has been carried out, pointing out the double-well shape due to the inversion motion of the amine moiety. A strong Coriolis coupling between the two inversion states prevented experimental accuracy to be achieved in the fitting procedure: no information on the energy difference between the two states could be derived. Based on the available literature for similar species, the inter-state transitions required for the characterization of the Coriols coupling take place in the far infrared region, thus preventing any conclusion to be drawn in this work. Despite the difficulties in the analysis, the line catalog of rotational transitions together with the spectroscopic parameters deliver the accuracy needed to guide the astronomical search of 3-aminoisoxazole in the interstellar medium.

Finally, the experimental and theoretical characterizations of iso- and n-propylamine and their conformers have been discussed in Chapter 13. In addition to a detailed conformational analysis, the rotational spectra of *t*-iso-propylamine and *tt*-n-propylamine (*i.e.*, the most stable conformers of the two isomeric forms) have been investigated up to the submillimeter-wave domain and satisfactorily reproduced with experimental accuracy. Some perturbation effects, likely due to Coriolis resonances, have been recognized in the spectra of the *gg*, *gg′*, and *gt* conformers of n-propylamine. However, the robust sets of spectroscopic constants determined in this work for *t*-iso-propylamine and *tt*-n-propylamine provide reliable spectral predictions for rotational transitions up to 500 GHz (see Figure 14.2), thus

enabling dedicated astronomical searches of propylamine isomers.

Appendices

Appendix A

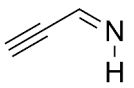
Supplementary Material

A.1 Determination of the r_e^{SE} structures

The SE method defined in Section 4.1 has been applied to the determination of the accurate equilibrium geometry of (*Z*)-propargylimine (see Figure A.1), using the data from Refs. 247 and 257. In these works, the rotational constants of the three ^{13}C isotopic species and two deuterated variants (–ND and –CCD), in addition to those of the parent species, for a total of six isotopologues, have been obtained. The vibrational corrections to equilibrium rotational constants ($\Delta B_{0,\mathrm{vib}}^{\mathrm{theo}}$, see Equation 4.2) have been determined at the B3 level of theory for each isotopologue. Given the planarity of (*Z*)-propargylimine, the zero-point inertial defect Δ_0 needs to be defined:[303]

$$\Delta_0 = I_0^c - I_0^a - I_0^b\,, \tag{A.1}$$

where I_0^g is the ground-state moment of inertia along the g-axis, with $g = a, b, c$. This is a measure of the planarity of the molecule: in fact, Δ_0 is null for a rigid planar molecule and takes negative values for rigid non-planar structures.[303] For (*Z*)-propargylimine, based on the ground-state experimental rotational constants from Ref. 247, Δ_0 is equal to 0.17 u Å^2, thus confirming that it belongs to the C_s symmetry point group. As a

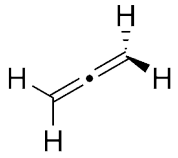


(*Z*)-propargylimine

Figure A.1: Molecular structure of (*Z*)-propargylimine.

consequence, the use of all three rotational constants A, B, C in the LSF procedure causes redundancies; only B and C have been employed and equally weighted. This led to the determination of a reliable partial SE equilibrium structure, whose is reported in Table A.2. The guess geometry (*i.e.*, the starting point of the LSF procedure) and the fixed parameters (whose choice has been driven by the availability of the isotopic substution at the involved atoms) have been taken from the TM-SE level.

In addition, the SE method has been employed for the determination of the equilibrium geometry of allene, represented in Figure A.2, using the data collected in Ref. 259 from the high-resolution infrared spectroscopy experiments on C_3H_4 and its fully-deuterated isotopologue C_3D_4 (see Refs. 304–308). Allene is a prolate symmetric-top rotor, therefore only A and B rotational constants could be determined from the experiment, while its molecular structure is entirely described by only three parameters: therefore, the determination of its equilibrium SE geometry presents a constrain on the number of the degrees of freedom. To address this issue and provide a reliable estimate of the structural parameters, a high-level computational characterization of the equilibrium geometry of allene has been carried out. Indeed, its symmetry (allene belongs to the D_{2d} symmetry point group) allows to exploit computationally demanding composite schemes such as the CCSD(T)/CBS+CV+fT+fQ and the CCSD(T)/CBS+CV models (*cf.* Section 3.7). The computed geometries as well as the equilibrium SE structures of allene are reported in Table A.3. For the latter, the vibrational corrections to equilibrium rotational constants have been determined at the revDSD level; the same level of theory has been chosen for the guess geometry. From the inspection of Table A.3, it is noted the excellent agreement between the SE equilibrium values and their computed counterparts. To assess the impact of the constrain on the degrees of freedom, the SE equilibrium structure has also been determined with the A1 angle (defined in Table A.3) fixed at the CCSD(T)/CBS+CV+fT+fQ value. In this case, the results are again in agreement with the theoretical data,



allene

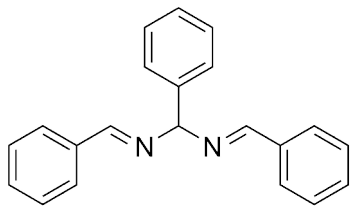
Figure A.2: Molecular structure of allene.

with a maximum discrepancy of 0.3 mÅ (the equilibrium SE values for the C–H and C=C bond lenghts are 1.3067(8) Å and 1.08123(4) Å, respectively).

A.2 Synthesis and characterization of hydrobenzamide

The synthesis of 1-phenyl-*N*,*N'*-bis(phenylmethylene)-methanediamine (hydrobenzamide or HBA, Figure A.3) has been carried out as described in literature.[222–226] Ammonium hydroxide solution (28–30% NH_3 basis, 510 mmol, 33.2 mL) has been added to benzaldehyde (39.4 mmol, 4 mL) and the mixture was stirred at room temperature for three hours. A white precipitate appeared, which was recovered by filtration through fritted funnel and washed with small portions of cold diethyl ether (10 mL, three times). HBA was obtained as white solid (2.747 g, 9.2 mmol, 70%) and used in the spectroscopic experiments without any further purification (purity $\geq$ 99% by ^{1}H-NMR).

A Varian Mercury 400 MHz spectrometer has been used for the ^{1}H-NMR spectrum of HBA, while the thermal analysis has been conducted using a Varian Inova 600 MHz equipped with an automated triple-resonance broadband probe, operating at a field of 14.4 T. All chemical shifts (δ) are reported in ppm from tetramethylsilane with the residual solvent resonance as the internal standard ($CHCl_3$: δ = 7.27 ppm, $C_2H_2Cl_4$: δ = 6.0 ppm). The signals that arise in the HBA spectrum after the addition of a few drops of benzonitrile are: i. a doublet at 7.68 ppm, overlapping with a benzaldehyde triplet; ii. a doublet at 7.63 ppm; and iii. a triplet at 7.50 ppm, overlapping with an HBA multiplet. The ATR-FTIR spectrum has been obtained with a Bruker Alpha System spectrometer, while the GS-MS spectrum has been taken by EI ionization at 70 eV on a Hp5971 with GC



hydrobenzamide
(HBA)

Figure A.3: Molecular structure of hydrobenzamide.

injection. An interesting feature arising from the GC-MS analysis is that the base peak (m/z: 194) represents the loss of phenylmethanimine from the molecular ion, whose peak is absent.

Characterization of HBA

^{1}H-NMR ($CDCl_3$, 400 MHz, 25 °C, ppm) δ: 8.60 (2H, s, PhCH=N), 7.90–7.84 (4H, m, ArH), 7.55–7.50 (2H, m, ArH), 7.47–7.41 (6H, m, ArH), 7.41–7.35 (2H, m, ArH), 7.33–7.27 (1H, m, ArH), 5.99 (1H, s, CHN_2).

^{13}C-NMR ($CDCl_3$, 100 MHz, 25 °C, ppm) δ: 160.8 (2C), 141.9, 136.2 (2C), 131.2 (2C), 128.8 (4C), 128.70 (4C), 128.68 (2C), 128.0, 127.4 (2C), 92.8.

ATR-FTIR (cm^{-1}) $\tilde{\nu}$: 3059, 3031, 2856, 1638, 1492, 1450, 1578, 1357, 1314, 1296, 1216, 1168, 1075, 1030, 914, 755, 700, 632, 565, 483.

MS (EI, m/z): 194, 165, 152, 116, 104, 89, 77.

A.3 Optimized geometries

Table A.1: Optimized revDSD geometries and SE equilibrium structures of the fragments used in the TM-SE approach validation. The SE structures are taken from Refs. 169–172 if not stated otherwise. Distances are in Å, angles in degrees.

ZMAT		revDSD	SE
methanimine			
H	R1	1.0211	1.0195(3)[a]
N 1 R1	R2	1.2721	1.2709(1)
C 2 R2 1 A1	R3	1.0932	1.092(3)
H 3 R3 2 A2 1 0.	R4	1.0887	1.0840(3)
H 3 R4 2 A3 1 180.	A1	110.37	110.35(4)
	A2	124.47	123.7(5)
	A3	118.72	119.3(5)
formaldehyde			
X	R1	1.2072	1.2047(1)
C 1 1.5	R2	1.1041	1.1003(1)
O 2 R1 1 90.	A1	121.78	121.65(1)
H 2 R2 3 A1 1 -90.			
H 2 R2 3 A1 1 90.			
hydrogen cyanide			
H	R1	1.0674	1.0651(1)
C 1 R1	R2	1.1575	1.1533(1)
X 2 1.5 1 90.			
N 2 R2 3 90. 1 180.			
acetylene			
C	R1	1.2062	1.2036(1)
C 1 R1	R2	1.0639	1.0611(1)
X 2 1.5 1 90.			
X 3 1.5 2 90. 1 180.			
X 1 1.5 2 90. 3 180.			
X 5 1.5 1 90. 2 180.			
H 2 R2 3 90. 4 0.			
H 1 R2 5 90. 6 0.			
ethene			
X	R1	1.3317	1.3311(1)
C 1 1.5	R2	1.0835	1.0807(1)
C 2 R1 1 90.	A1	121.50	121.42(1)
H 3 R2 2 A1 1 90.			
H 3 R2 2 A1 1 -90.			
H 2 R2 3 A1 1 -90.			
H 2 R2 3 A1 1 90.			
acetonitrile			
N	R1	1.1598	1.1554(3)[b]
C 1 R1	R2	1.4614	1.4586(3)
X 2 1.5 1 90.	R3	1.0897	1.0865(1)
X 3 1.5 2 90. 1 0.	A1	109.90	109.85(1)

(continues)

(Table A.1)

ZMAT		revDSD	SE
C 2 R2 3 90. 4 180.			
H 5 R3 2 A1 3 0.			
H 5 R3 2 A1 3 -120.			
H 5 R3 2 A1 3 120.			
benzene			
X	R1	1.3939	1.3916(1)
X 1 1.5	R2	1.0837	1.0799(1)
C 2 R1 1 90.			
C 2 R1 1 90. 3 60.			
C 2 R1 1 90. 4 60.			
C 2 R1 1 90. 5 60.			
C 2 R1 1 90. 6 60.			
C 2 R1 1 90. 7 60.			
H 3 R2 8 120. 2 180.			
H 4 R2 3 120. 2 180.			
H 5 R2 4 120. 2 180.			
H 6 R2 5 120. 2 180.			
H 7 R2 6 120. 2 180.			
H 8 R2 7 120. 2 180.			
furane			
X	R1	1.3624	1.3598(4)
O 1 1.5	R2	1.3588	1.3542(4)
C 2 R1 1 A1	R3	1.0764	1.0739(3)
C 2 R1 1 A1 3 180.	R4	1.0776	1.0743(3)
C 3 R2 2 A2 1 180.	A1	126.66	126.75(3)
C 4 R2 2 A2 1 180.	A2	110.62	110.79(4)
H 3 R3 2 A3 1 0.	A3	115.86	115.82(3)
H 4 R3 2 A3 1 0.	A4	126.33	126.4(1)
H 5 R4 3 A4 2 180.			
H 6 R4 4 A4 2 180.			
isoxazole			
C	R1	1.4213	1.4210[c]
C 1 R1	R2	1.3577	1.3515
C 2 R2 1 A1	R3	1.3435	1.3414
O 3 R3 2 A2 1 0.	R4	1.0793	1.0759
H 1 R4 2 A3 3 180.	R5	1.0759	1.0731
H 2 R5 1 A4 3 180.	R6	1.0777	1.0756
H 3 R6 2 A5 1 180.	R7	1.3121	1.3071
N 1 R7 2 A6 3 0.	A1	103.03	102.99
	A2	110.51	110.75
	A3	129.09	129.34
	A4	128.77	128.44
	A5	133.65	133.52
	A6	112.20	112.16
ammonia			
X	R1	1.0137	1.0111(2)
N 1 1.5	A1	112.12	
H 2 R1 1 A1	∠HNH	106.70	106.87(3)

(continues)

(Table A.1)

ZMAT	revDSD	SE
H 2 R1 1 A1 3 -120. H 2 R1 1 A1 3 120.		

[a] Ref. 185. [b] Ref. 238. [c] Ref. 187. No standard deviation on the parameters has been reported in the original work.

Table A.2: Computed and SE geometries for the molecules of the r_e^{SE} subset. The reference list for the SE structures can be found in Table 7.1. Distances are in Å, angles in degrees.

ZMAT		revDSD	SE	TM-SE	TM-SE-LR[+c][a]
vinylacetylene					
H	R1	1.0637	1.0617(4)	1.0609	
C 1 R1	R2	1.2111	1.2072(4)	1.2085	
X 2 1.5 1 90.	R3	1.4285	1.4267(4)	1.4285	1.4258
X 3 1.5 2 90. 1 0.	R4	1.0852	1.0819(4)	1.0824	
C 2 R2 3 90. 4 T1	R5	1.3387	1.3381(4)	1.3381	
X 5 1.5 2 90. 3 180.	R6	1.0823	1.0798(4)	1.0795	
X 6 1.5 5 90. 2 0.	R7	1.0828	1.0798(4)	1.0800	
C 5 R3 6 90. 7 T2	A1	116.37	116.60(2)	116.37	[116.57]
H 8 R4 5 A1 6 90.	A2	119.95	120.1(2)	119.95	
C 8 R5 9 A2 5 180.	A3	120.61	120.5(2)	120.61	
H 10 R6 8 A3 9 0.	A4	121.33	121.2(2)	121.33	
H 10 R7 8 A4 9 180.	T1	−179.43	−179.2(2)	−179.43	
	T2	−178.50	−178.5(2)	−178.50	
vinyl cyanide					
C	R1	1.3361	1.3351(4)	1.3355	
C 1 R1	R2	1.4336	1.4315(4)	1.4336	1.4309
C 2 R2 1 A1	R3	1.1627	1.1584(4)	1.1586	
X 3 1.5 2 90. 1 180.	R4	1.0823	1.0797(4)	1.0796	
X 4 1.5 3 90. 2 180.	R5	1.0819	1.0785(4)	1.0791	
N 3 R3 4 A2 5 0.	R6	1.0831	1.0800(4)	1.0804	
H 1 R4 2 A3 3 0.	A1	122.29	122.03(7)	122.29	[122.09]
H 1 R5 2 A4 3 180.	A2	88.97	89.14(7)	88.97	
H 2 R6 1 A5 7 180.	A3	121.44	121.31(7)	121.44	
	A4	120.43	120.48(7)	120.43	
	A5	121.54	121.66(7)	121.54	
trans-acrolein					
C	R1	1.0853	1.0827(2)	1.0825	
H 1 R1	R2	1.0824	1.0792(2)	1.0796	
H 1 R2 2 A1	R3	1.3369	1.3356(2)	1.3362	
C 1 R3 3 A2 2 180.	R4	1.0841	1.0815(1)	1.0813	
H 4 R4 1 A3 3 0.	R5	1.4732	1.4700(1)	1.4732	1.4702
C 4 R5 1 A4 3 180.	R6	1.1085	1.1049(2)	1.1047	
H 6 R6 4 A5 1 0.	R7	1.2139	1.2105(1)	1.2114	
O 6 R7 7 A6 4 180.	A1	117.29	117.48(2)	117.29	

(continues)

(Table A.2)

ZMAT		revDSD	SE	TM-SE	TM-SE-LR[+c][a]
	A2	122.14	122.09(2)	122.14	
	A3	122.57	122.76(2)	122.57	
	A4	120.38	120.22(1)	120.38	[120.18]
	A5	114.86	115.02(2)	114.86	
	A6	120.94	120.99(2)	120.94	
cis-acrolein					
C	R1	1.0838	1.0810(5)	1.0810	
H 1 R1	R2	1.0824	1.0791(3)	1.0796	
H 1 R2 2 A1	R3	1.3362	1.3359(4)	1.3356	
C 1 R3 3 A2 2 180.	R4	1.0848	1.0814(3)	1.0821	
H 4 R4 1 A3 3 0.	R5	1.4844	1.4818(4)	1.4844	1.4814
C 4 R5 1 A4 3 180.	R6	1.1060	1.1020(3)	1.1022	
H 6 R6 4 A5 1 180.	R7	1.2142	1.2101(3)	1.2117	
O 6 R7 4 A6 1 0.	A1	118.34	118.57(6)	118.34	
	A2	121.62	121.59(5)	121.62	
	A3	121.37	121.59(4)	121.37	
	A4	121.62	121.28(3)	121.62	[121.42]
	A5	115.49	115.82(4)	115.49	
	A6	124.31	123.88(3)	124.31	
trans-1,3-butadiene					
X	R1	0.7286	0.7274(1)	0.7286	0.7271
X 1 1.5	R2	1.3392	1.3380(2)	1.3386	
C 1 R1 2 90.	R3	1.0870	1.0842(1)	1.0842	
C 1 R1 2 90. 3 180.	R4	1.0824	1.0796(1)	1.0797	
C 3 R2 1 A1 2 180.	R5	1.0846	1.0817(1)	1.0819	
C 4 R2 1 A1 2 0.	A1	123.74	123.54(1)	123.74	[123.54]
H 3 R3 1 A2 2 0.	A2	116.65	116.65(2)	116.65	
H 4 R3 1 A2 2 180.	A3	121.54	121.45(1)	121.54	
H 5 R4 3 A3 1 180.	A4	121.08	120.96(1)	121.08	
H 6 R4 4 A3 1 180.					
H 5 R5 3 A4 1 0.					
H 6 R5 4 A4 1 0.					
trans-glyoxal					
X	R1	0.7602	0.7573(2)	0.7602	0.7585
X 1 1.5	R2	1.1040	1.1007(3)	1.1002	
C 1 R1 2 90.	R3	1.2086	1.2045(3)	1.2061	
C 1 R1 2 90. 3 180.	A1	115.17	115.25(2)	115.17	
H 3 R2 1 A1 2 0.	A2	121.36	121.28(3)	121.36	[121.16]
H 4 R2 1 A1 2 180.					
O 4 R3 1 A2 2 0.					
O 3 R3 1 A2 2 180.					
(*Z*)-propargylimine					
H	R1	1.0644	1.059(1)[b]	1.0616	
C 1 R1	R2	1.2106	1.207(3)	1.2080	
X 2 1.5 1 90.	R3	1.4393	1.436(2)	1.4393	1.4366
X 3 1.5 2 90. 1 180.	R4	1.0887	1.0840[c]	1.0840	
C 2 R2 3 A1 4 0.	R5	1.2786	1.272(3)	1.2774	
X 5 1.5 2 90. 3 180.	R6	1.0222	1.021(2)	1.0206	

(continues)

(Table A.2)

ZMAT		revDSD	SE	TM-SE	TM-SE-LR[+c][a]
X 6 1.5 5 90. 2 0.	A1	89.62	89.9(8)	89.62	
C 5 R3 6 A2 7 180.	A2	89.77	90.0(9)	89.77	
H 8 R4 5 A3 6 0.	A3	116.27	116.3(4)	116.27	[116.47]
N 8 R5 9 A4 6 180.	A4	118.03	118.03[c]	118.03	
H 10 R6 8 A5 5 0.	A5	109.99	109.99[c]	109.99	

[a] All parameters not reported are identical to the TM-SE values. The value in square brackets is used when the TM-SE-LR+c approach is considered. [b] The determination of the r_e^{SE} has been carried out in Section A.1. [c] Fixed at the TM-SE value.

Table A.3: Optimized geometries of allene at the revDSD, CCSD(T)/CBS+CV, and CCSD(T)/CBS+CV+fT+fQ levels and SE equilibrium structure. Distances are in Å, angles in degrees.

ZMAT		revDSD	CBS+CV	CBS+CV+fT+fQ	SE
allene					
X	R1	1.3077	1.30673	1.30702	1.307(3)
C 1 1.5	R2	1.0837	1.08100	1.08104	1.081(6)
C 2 R1 1 90.	A1	120.94	120.876	120.891	120.9(6)
C 2 R1 1 90. 3 180.					
H 4 R2 2 A1 1 90.					
H 3 R2 2 A1 1 180.					
H 3 R2 2 A1 1 0.					
H 4 R2 2 A1 1 -90.					

Table A.4: Optimized geometries of allenylacetylene at the jun-ChS, jB2, and TM-SE(-LR) levels. Two different quantum-chemical methods for the allene fragment have been used. Distances are in Å, angles in degrees.

ZMAT		jun-ChS	jB2	revDSD	TM-SE[-LR][a,b]	TM-SE[-LR][a,c]
allenylacetylene						
H	R1	1.0829	1.0843	1.0862	1.08351	1.08355
C 1 R1	R2	1.4275	1.4229	1.4286	[1.42595]	[1.42595]
C 2 R2 1 A1	R3	1.3147	1.3133	1.3161	1.31514	1.31543
C 2 R3 1 A2 3 180.	R4	1.3039	1.3002	1.3047	1.30376	1.30406
X 4 1.5 2 90. 1 0.	R5	1.0812	1.0820	1.0843	1.08162	1.08166
C 4 R4 5 A3 2 180.	R6	1.2064	1.2076	1.2105	1.20788	1.20788
H 6 R5 4 A4 5 T1	R7	1.0615	1.0604	1.0637	1.06086	1.06086
H 6 R5 4 A4 5 –T1	A1	117.70	123.63	117.65	117.65	117.65
X 3 1.5 2 90. 1 0.	A2	119.19	118.87	119.04	119.04	119.04
C 3 R6 9 A5 2 180.	A3	89.30	89.36	89.36	89.36	89.36
X 10 1.5 3 90. 9 0.	A4	120.87	121.11	120.95	120.95	120.95
H 10 R7 11 A6 9 180.	A5	88.06	88.19	88.16	88.16	88.16
	A6	90.69	90.60	90.65	90.65	90.65
	T1	–90.21	–90.19	–90.18	–90.18	–90.18

[a] When the LR correction is taken into account, the values in square brackets must be considered. Otherwise, these values are identical to the revDSD ones. [b] The TM-QC method used for the allene fragment is based on CCSD(T)/CBS+CV computations. [c] The TM-QC method used for the allene fragment is based on CCSD(T)/CBS+CV+fT+fQ computations.

Table A.5: Optimized geometries of the cyanovinylacetylene isomers at the jun-ChS, jB2, and TM-SE(-LR) levels. All three possible choices of fragments have been included. Distances are in Å, angles in degrees.

ZMAT		jun-ChS	jB2	revDSD	TM-SE[-LR][a,b]	TM-SE[-LR][a,c]	TM-SE[-LR][a,d]
(*E*)-cyanovinylacetylene							
N	R1	1.1580	1.1611	1.1635	1.1594	1.1591	1.1594
C 1 R1	R2	1.4278	1.4221	1.4285	[1.4258]	1.4264	[1.4258]
X 2 1.5 1 A1	R3	1.0797	1.0805	1.0829	1.0801	1.0798	1.0799
C 2 R2 3 90. 1 180.	R4	1.3434	1.3447	1.3455	1.3449	1.3445	1.3449
H 4 R3 2 A2 3 0.	R5	1.0817	1.0825	1.0849	1.0821	1.0822	1.0816
C 4 R4 2 A3 3 180.	R6	1.4210	1.4148	1.4217	[1.4190]	[1.4190]	1.4198
H 6 R5 4 A4 2 0.	R7	1.2065	1.2080	1.2107	1.2081	1.2081	1.2069
C 6 R6 4 A5 2 180.	R8	1.0620	1.0612	1.0644	1.0616	1.0616	1.0624
X 8 1.5 6 90. 4 180.	A1	89.30	88.88	88.94	88.94	88.94	88.94
C 8 R7 9 A6 6 180.	A2	117.28	117.02	117.13	117.13	117.13	117.13
X 10 1.5 8 90. 9 0.	A3	121.35	121.94	121.24	121.24	121.24	121.24
H 10 R8 11 A7 9 180.	A4	119.81	119.50	119.71	119.71	119.71	119.71
	A5	122.58	123.24	122.91	122.91	122.91	122.91
	A6	88.65	88.37	88.42	88.42	88.42	88.42
	A7	90.50	90.45	90.47	90.47	90.47	90.47
(*Z*)-cyanovinylacetylene							
N	R1	1.1576	1.1609	1.1633	1.1592	1.1590	1.1592
C 1 R1	R2	1.4278	1.4221	1.4285	[1.4258]	1.4264	[1.4258]
X 2 1.5 1 90.	R3	1.0791	1.0798	1.0823	1.0795	1.0791	1.0798
C 2 R2 3 A1 1 180.	R4	1.3440	1.3458	1.3465	1.3459	1.3455	1.3460
H 4 R3 2 A2 3 0.	R5	1.0811	1.0818	1.0843	1.0815	1.0809	1.0810
C 4 R4 2 A3 3 180.	R6	1.4198	1.4137	1.4206	[1.4180]	[1.4180]	1.4187

(continues)

(Table A.5)

ZMAT								jun-ChS	jB2	revDSD	TM-SE[-LR][a,b]	TM-SE[-LR][a,c]	TM-SE[-LR][a,d]
H	6	R5	4	A4	2	180.	R7	1.2064	1.2078	1.2105	1.2079	1.2079	1.2066
C	6	R6	4	A5	2	0.	R8	1.0621	1.0612	1.0644	1.0616	1.0616	1.0624
X	8	1.5	6	90.	4	0.	A1	88.76	88.56	88.63	88.63	88.63	88.63
C	8	R7	9	A6	6	180.	A2	116.87	116.71	116.84	116.84	116.84	116.84
X	10	1.5	8	90.	9	0.	A3	122.51	123.16	120.36	120.36	120.36	120.36
H	10	R8	11	A7	9	180.	A4	118.70	118.26	118.52	118.52	118.52	118.52
							A5	123.81	124.50	124.14	124.14	124.14	124.14
							A6	90.54	90.83	90.83	90.83	90.83	90.83
							A7	90.19	90.19	89.18	89.18	89.18	89.18

[a] When the LR correction is taken into account, the values in square brackets must be considered. Otherwise, these values are identical to the revDSD ones. [b] The fragments are hydrogen cyanide, ethene, and acetylene. [c] The fragments are vinylcyanide and acetylene. [d] The fragments are vinylacetylene and hydrogen cyanide.

Table A.6: Computed jB2 geometry of vinylcyanoacetylene in Cartesian coordinates.

	x	*y*	*z*
C	2.880293	0.549489	0.0
H	2.559133	1.580364	0.0
H	3.940704	0.349362	0.0
C	1.998810	−0.456650	0.0
H	2.340978	−1.483770	0.0
C	0.595890	−0.262441	0.0
C	−0.610364	−0.124712	0.0
C	−1.964837	0.041934	0.0
N	−3.121279	0.184182	0.0

Table A.7: Optimized geometries of the phenylmethanimine isomers at the ChS, mB2, revDSD, and TM-SE(-LR) levels. Distances are in Å, angles in degrees.

ZMAT		ChS	mB2	revDSD	TM-SE[-LR][a]
(*E*)-phenylmethanimine					
C	R1	1.3891	1.3904	1.3920	1.3897
C 1 R1	R2	1.3910	1.3909	1.3936	1.3914
C 2 R2 1 A1	R3	1.3931	1.3958	1.3967	1.3944
C 3 R3 2 A2 1 0.	R4	1.3970	1.3989	1.4005	1.3982
C 4 R4 3 A3 2 0.	R5	1.3860	1.3863	1.3884	1.3861
C 5 R5 4 A4 3 0.	R6	1.0808	1.0808	1.0836	1.0798
H 1 R6 2 A5 3 180.	R7	1.0805	1.0806	1.0834	1.0796
H 2 R7 1 A6 6 180.	R8	1.0826	1.0824	1.0852	1.0814
H 3 R8 2 A7 1 180.	R9	1.0804	1.0800	1.0829	1.0791
H 5 R9 4 A8 3 180.	R10	1.0807	1.0808	1.0836	1.0798
H 6 R10 5 A9 4 180.	R11	1.4716	1.4699	1.4734	[1.4705]
C 4 R11 3 A10 2 180.	R12	1.0930	1.0941	1.0962	1.0950
H 12 R12 4 A11 3 0.	R13	1.2749	1.2755	1.2777	1.2765
N 12 R13 13 A12 3 180.	R14	1.0158	1.0173	1.0191	1.0175
H 14 R14 12 A13 4 180.	A1	119.74	119.77	119.75	119.75
	A2	120.42	120.52	120.48	120.48
	A3	119.63	119.41	119.51	119.51
	A4	119.92	120.04	119.99	119.99
	A5	120.00	120.02	120.02	120.02
	A6	120.19	120.18	120.19	120.19
	A7	120.11	120.04	120.04	120.04
	A8	118.69	118.66	118.72	118.72
	A9	119.88	119.87	119.88	119.88
	A10	119.29	119.32	119.28	119.28
	A11	115.20	114.78	114.97	114.97
	A12	122.76	122.74	122.76	122.76
	A13	110.42	110.78	110.35	110.35

(continues)

(Table A.7)

ZMAT								ChS	mB2	revDSD	TM-SE[-LR][a]
(*Z*)-phenylmethanimine											
C								R1	1.3891	1.3903	1.3920	1.3898
C	1	R1						R2	1.3903	1.3902	1.3929	1.3906
C	2	R2	1	A1				R3	1.3944	1.3968	1.3978	1.3955
C	3	R3	2	A2	1	0.		R4	1.3964	1.3982	1.3998	1.3976
C	4	R4	3	A3	2	0.		R5	1.3867	1.3870	1.3891	1.3869
C	5	R5	4	A4	3	0.		R6	1.0805	1.0806	1.0834	1.0796
H	1	R6	2	A5	3	180.		R7	1.0805	1.0805	1.0833	1.0796
H	2	R7	1	A6	6	180.		R8	1.0823	1.0821	1.0849	1.0811
H	3	R8	2	A7	1	180.		R9	1.0819	1.0816	1.0844	1.0806
H	5	R9	4	A8	3	180.		R10	1.0806	1.0806	1.0834	1.0796
H	6	R10	5	A9	4	180.		R11	1.4791	1.4775	1.4808	[1.4777]
C	4	R11	3	A10	2	180.		R12	1.0894	1.0902	1.0924	1.0877
H	12	R12	4	A11	3	0.		R13	1.2734	1.2739	1.2763	1.2751
N	12	R13	4	A12	3	180.		R14	1.0194	1.0209	1.0223	1.0207
H	14	R14	12	A13	4	0.		A1	119.80	119.83	119.82	119.82
								A2	120.76	120.83	120.81	120.81
								A3	119.00	118.82	118.90	118.90
								A4	120.38	120.48	120.44	120.44
								A5	120.14	120.15	120.16	120.16
								A6	120.19	120.17	120.18	120.18
								A7	120.07	120.02	120.01	120.01
								A8	120.29	120.17	120.27	120.27
								A9	119.83	119.83	119.83	119.83
								A10	118.28	118.37	118.29	118.29
								A11	114.55	114.25	114.38	114.38
								A12	116.79	116.54	116.58	116.58
								A13	111.04	111.39	110.98	110.98

[a] When the LR correction is taken into account, the values in square brackets must be considered. Otherwise, these values are identical to the revDSD ones.

Table A.8: Computed mB2 geometry of phenylmethanimine isomerization transition state in Cartesian coordinates.

	x	*y*	*z*		*x*	*y*	*z*
C	−2.242689	0.268402	0.0	H	−0.027340	−2.297768	0.0
C	−1.772569	−1.040677	0.0	H	0.738288	1.901097	0.0
C	−0.402609	−1.282022	0.0	H	−1.709603	2.350066	0.0
C	0.502563	−0.221441	0.0	C	1.961600	−0.516417	0.0
C	0.023748	1.089686	0.0	H	2.168984	−1.602911	0.0
C	−1.341473	1.333726	0.0	N	2.876181	0.320178	0.0
H	−3.306296	0.460873	0.0	H	3.594875	0.997478	0.0
H	−2.469209	−1.866940	0.0				

Table A.9: Optimized geometries of 2- and 3-furonitrile at the jun-ChS, revDSD, and TM-SE(-LR) levels. Distances are in Å, angles in degrees.

ZMAT		jun-ChS	revDSD	TM-SE[-LR][a]
2-furonitrile				
C	R1	1.3643	1.3616	1.3570
C 1 R1	R2	1.4328	1.4275	1.4266
C 2 R2 1 A1	R3	1.3672	1.3656	1.3610
C 3 R3 2 A2 1 0.	R4	1.3676	1.3649	1.3623
O 4 R4 3 A3 2 0.	R5	1.0764	1.0765	1.0740
H 1 R5 2 A4 3 180.	R6	1.0772	1.0773	1.0740
H 2 R6 1 A5 5 180.	R7	1.0775	1.0775	1.0742
H 3 R7 2 A6 1 180.	R8	1.4239	1.4180	[1.4154]
C 4 R8 3 A7 2 180.	R9	1.1653	1.1633	1.1592
X 9 1.5 4 90. 3 0.	A1	106.05	106.14	106.14
N 9 R9 10 A8 4 180.	A2	105.77	105.87	105.87
	A3	110.96	110.70	110.70
	A4	133.38	133.44	133.44
	A5	126.45	126.41	126.41
	A6	128.29	128.24	128.24
	A7	131.69	131.80	131.80
	A8	88.58	88.60	88.60
3-furonitrile				
C	R1	1.3509	1.3555	1.3510
C 1 R1	R2	1.4363	1.4393	1.4384
C 2 R2 1 A1	R3	1.3602	1.3657	1.3611
C 3 R3 2 A2 1 0.	R4	1.3460	1.3507	1.3481
O 4 R4 3 A3 2 0.	R5	1.0732	1.0760	1.0735
H 1 R5 2 A4 3 180.	R6	1.0741	1.0770	1.0737
H 2 R6 1 A5 5 180.	R7	1.0737	1.0764	1.0739
H 4 R7 3 A6 2 180.	R8	1.4213	1.4219	[1.4193]
C 3 R8 2 A7 1 180.	R9	1.1570	1.1628	1.1587
X 9 1.5 3 90. 2 0.	A1	105.46	105.57	105.57
N 9 R9 10 A8 3 180.	A2	106.34	106.33	106.33
	A3	110.23	110.17	110.17
	A4	133.35	133.47	133.47
	A5	127.45	127.38	127.38
	A6	132.66	132.80	132.80
	A7	127.44	127.40	127.40
	A8	89.36	89.30	89.30

[a] When the LR correction is taken into account, the values in square brackets must be considered. Otherwise, these values are identical to the revDSD ones.

Table A.10: Optimized geometries of 3-aminoisoxazole at the jun-ChS, revDSD, and TM-SE(-LR) levels. Distances are in Å, angles in degrees.

ZMAT		jun-ChS	revDSD	TM-SE[-LR][a]
3-aminoisoxazole				
O	R1	1.3993	1.4018	1.4018
N 1 R1	R2	1.3098	1.3156	1.3106
C 2 R2 1 A1	R3	1.4256	1.4282	1.4279
C 3 R3 2 A2 1 0.	R4	1.3359	1.3404	1.3384
C 1 R4 2 A3 3 0.	R5	1.3822	1.3851	[1.3828]
N 3 R5 2 A4 1 180.	R6	1.0738	1.0765	1.0737
H 4 R6 3 A5 2 180.	R7	1.0755	1.0782	1.0761
H 5 R7 1 A6 2 180.	R8	1.0048	1.0077	1.0051
H 6 R8 3 A7 2 T1	R9	1.0067	1.0097	1.0070
H 6 R9 3 A8 2 T2	A1	105.43	105.44	105.46
	A2	111.98	111.87	111.86
	A3	108.55	108.62	108.60
	A4	121.03	121.14	121.15
	A5	128.57	128.57	128.56
	A6	115.95	115.94	115.93
	A7	113.83	113.92	113.92
	A8	112.74	112.94	112.93
	T1	143.82	144.34	144.23
	T2	14.90	15.04	14.98

[a] When the LR correction is taken into account, the values in square brackets must be considered. Otherwise, these values are identical to the revDSD ones.

Table A.11: Optimized geometries of the propylamine isomers and their conformers at the ChS and mB2-dH levels. Distances are in Å, angles in degrees.

ZMAT		ChS	mB2-dH		ChS	mB2-dH
trans-iso-propylamine						
H	R1	1.0984	1.0985	T1	−119.35	−119.24
C 1 R1	R2	1.4640	1.4680	T2	−59.74	−59.34
N 2 R2 1 A1	R3	1.5182	1.5219	T3	57.47	58.14
C 2 R3 3 A2 1 T1	R4	1.0118	1.0127	T4	177.73	178.54
C 2 R3 3 A2 1 –T1	R5	1.0899	1.0893	T5	−62.51	−61.78
H 3 R4 2 A3 1 T2	R6	1.0903	1.0895			
H 3 R4 2 A3 1 –T2	R7	1.0910	1.0906			
H 4 R5 2 A4 1 T3	A1	111.65	111.92			
H 5 R5 2 A4 1 –T3	A2	108.72	108.71			
H 4 R6 2 A5 1 T4	A3	111.16	110.78			
H 5 R6 2 A5 1 –T4	A4	110.60	110.86			
H 4 R7 2 A6 1 T5	A5	109.90	110.06			
H 5 R7 2 A6 1 –T5	A6	111.05	111.10			

(continues)

(Table A.11)

ZMAT		ChS	mB2-dH		ChS	mB2-dH
gauche-iso-propylamine						
C	R1	1.0896	1.0890	A6	111.06	111.39
H 1 R1	R2	1.0883	1.0877	A7	110.91	110.99
H 1 R2 2 A1	R3	1.0928	1.0922	A8	110.17	110.38
H 1 R3 3 A2 2 T1	R4	1.5188	1.5224	A9	108.54	108.65
C 1 R4 3 A3 4 T2	R5	1.0932	1.0925	A10	110.26	110.00
H 5 R5 1 A4 3 T3	R6	1.5254	1.5294	A11	111.01	110.64
C 5 R6 1 A5 3 T4	R7	1.0911	1.0908	T1	−117.07	−116.73
H 7 R7 5 A6 1 T5	R8	1.0905	1.0901	T2	−120.76	−120.92
H 7 R8 5 A7 1 T6	R9	1.0922	1.0916	T3	−58.00	−59.22
H 7 R9 5 A8 1 T7	R10	1.4633	1.4677	T4	−177.15	−178.09
N 5 R10 1 A9 7 T8	R11	1.0136	1.0144	T5	60.09	59.61
H 11 R11 5 A10 1 T9	R12	1.0120	1.0130	T6	−179.54	−179.95
H 11 R12 5 A11 1 T10	A1	108.98	108.86	T7	−59.77	−60.35
	A2	108.28	108.11	T8	−125.64	−126.13
	A3	110.68	110.65	T9	60.97	61.56
	A4	108.48	108.27	T10	179.24	179.13
	A5	111.02	111.22			
tt-n-propylamine						
H	R1	1.0891	1.0888	A6	109.33	109.29
C 1 R1	R2	1.5227	1.5255	A7	110.86	110.55
C 2 R2 1 A1	R3	1.5237	1.5282	T1	59.77	59.81
C 3 R3 2 A2 1 180.	R4	1.4587	1.4622	T2	−121.96	−122.13
N 4 R4 3 A3 2 180.	R5	1.0907	1.0903	T3	58.08	57.91
H 2 R5 3 A4 4 T1	R6	1.0934	1.0929	T4	−59.33	−58.98
H 2 R5 3 A4 4 –T1	R7	1.0922	1.0917			
H 3 R6 4 A5 2 T2	R8	1.0115	1.0126			
H 3 R6 4 A5 2 –T2	A1	111.53	111.58			
H 4 R7 3 A6 2 T3	A2	111.96	112.51			
H 4 R7 3 A6 2 –T3	A3	115.77	116.13			
H 5 R8 4 A7 3 T4	A4	110.84	111.02			
H 5 R8 4 A7 3 –T4	A5	108.98	108.96			
tg-n-propylamine						
C	R1	1.0940	1.0934	A6	110.77	110.42
H 1 R1	R2	1.0911	1.0904	A7	111.44	110.99
H 1 R2 2 A1	R3	1.5175	1.5212	A8	112.00	112.52
C 1 R3 3 A2 2 T1	R4	1.0919	1.0913	A9	111.09	111.26
H 4 R4 1 A3 3 T2	R5	1.0975	1.0979	A10	111.30	111.35
H 4 R5 1 A4 3 T3	R6	1.4616	1.4649	A11	110.83	111.02
N 4 R6 1 A5 3 T4	R7	1.0117	1.0127	T1	−117.35	−117.17
H 7 R7 4 A6 1 T5	R8	1.0106	1.0117	T2	−61.20	−62.25
H 7 R8 4 A7 1 T6	R9	1.5221	1.5247	T3	−177.86	−178.48
C 1 R9 4 A8 7 T7	R10	1.0907	1.0903	T4	57.21	56.13
H 10 R10 1 A9 4 T8	R11	1.0890	1.0888	T5	63.27	64.03
H 10 R11 1 A10 4 T9	R12	1.0907	1.0903	T6	−177.57	−177.63
H 10 R12 1 A11 4 T10	A1	106.71	106.50	T7	179.61	178.75
	A2	108.65	108.55	T8	59.80	59.72
	A3	109.28	109.20	T9	180.01	179.89

(continues)

(Table A.11)

ZMAT								ChS	mB2-dH		ChS	mB2-dH
								A4	109.01	108.93	T10	−59.98	−60.13
								A5	110.52	110.69			
gt-n-propylamine													
C								R1	1.0937	1.0933	A6	111.32	111.14
H	1	R1						R2	1.0932	1.0927	A7	110.89	110.69
H	1	R2	2	A1				R3	1.5265	1.5310	A8	113.25	113.57
C	1	R3	3	A2	2	T1		R4	1.0924	1.0919	A9	110.69	110.89
H	4	R4	1	A3	3	T2		R5	1.0908	1.0905	A10	111.23	111.38
H	4	R5	1	A4	3	T3		R6	1.4595	1.4626	A11	111.59	111.60
N	4	R6	1	A5	3	T4		R7	1.0109	1.0114	T1	117.18	117.05
H	7	R7	4	A6	1	T5		R8	1.0117	1.0126	T2	64.33	63.58
H	7	R8	4	A7	1	T6		R9	1.5225	1.5251	T3	−52.11	−52.43
C	1	R9	4	A8	7	T7		R10	1.0913	1.0909	T4	−173.92	−174.41
H	10	R10	1	A9	4	T8		R11	1.0892	1.0889	T5	−60.23	−61.02
H	10	R11	1	A10	4	T9		R12	1.0906	1.0899	T6	58.96	57.76
H	10	R12	1	A11	4	T10		A1	106.65	106.42	T7	64.47	63.74
								A2	108.34	108.41	T8	57.82	58.74
								A3	109.23	109.23	T9	177.77	178.72
								A4	109.40	109.22	T10	−62.40	−61.42
								A5	116.13	116.54			
gg-n-propylamine													
C								R1	1.0944	1.0938	A6	110.51	110.32
H	1	R1						R2	1.0923	1.0917	A7	111.23	110.97
H	1	R2	2	A1				R3	1.5210	1.5247	A8	113.13	113.38
C	1	R3	3	A2	2	T1		R4	1.0921	1.0915	A9	110.48	110.67
H	4	R4	1	A3	3	T2		R5	1.0959	1.0965	A10	111.09	111.23
H	4	R5	1	A4	3	T3		R6	1.4633	1.4663	A11	110.90	110.82
N	4	R6	1	A5	3	T4		R7	1.0119	1.0128	T1	116.78	116.68
H	7	R7	4	A6	1	T5		R8	1.0104	1.0114	T2	69.67	68.31
H	7	R8	4	A7	1	T6		R9	1.5222	1.5249	T3	−47.31	−48.19
C	1	R9	4	A8	7	T7		R10	1.0911	1.0907	T4	−172.07	−173.41
H	10	R10	1	A9	4	T8		R11	1.0896	1.0893	T5	67.12	67.81
H	10	R11	1	A10	4	T9		R12	1.0881	1.0872	T6	−174.15	−173.91
H	10	R12	1	A11	4	T10		A1	106.40	106.17	T7	65.58	64.10
								A2	108.50	108.45	T8	61.22	61.74
								A3	109.29	109.19	T9	−179.05	−178.49
								A4	109.23	109.07	T10	−58.35	−57.72
								A5	110.61	110.86			
gg′-n-propylamine													
C								R1	1.0915	1.0908	A6	111.04	110.66
H	1	R1						R2	1.0922	1.0918	A7	111.20	110.88
H	1	R2	2	A1				R3	1.5195	1.5234	A8	112.89	113.28
C	1	R3	2	A2	3	T1		R4	1.0979	1.0984	A9	110.86	111.04
H	4	R4	1	A3	2	T2		R5	1.0905	1.0900	A10	111.09	111.20
H	4	R5	1	A4	2	T3		R6	1.4630	1.4663	A11	111.52	111.48
N	4	R6	1	A5	2	T4		R7	1.0105	1.0115	T1	−117.02	−116.89
H	7	R7	4	A6	1	T5		R8	1.0110	1.0117	T2	177.00	176.22
H	7	R8	4	A7	1	T6		R9	1.5235	1.5261	T3	60.12	59.86

(continues)

(Table A.11)

ZMAT		ChS	mB2-dH		ChS	mB2-dH
C 1 R9 4 A8 7 T7	R10	1.0917	1.0913	T4	−58.21	−58.43
H 10 R10 1 A9 4 T8	R11	1.0893	1.0889	T5	173.66	171.94
H 10 R11 1 A10 4 T9	R12	1.0903	1.0895	T6	−67.46	−69.93
H 10 R12 1 A11 4 T10	A1	107.41	107.24	T7	64.05	63.81
	A2	108.74	108.60	T8	56.30	56.62
	A3	108.87	108.87	T9	176.19	176.51
	A4	109.39	109.14	T10	−64.30	−63.97
	A5	111.00	111.23			

Table A.12: Optimized geometries in Cartesian coordinates at the mB2-dH level of propylamine conformational transition states.

	x	*y*	*z*		*x*	*y*	*z*
gg'-tg-n-propylamine				*gg-tg*-n-propylamine			
C	−0.623609	0.204952	0.553331	C	−0.620700	0.543995	0.225379
H	−0.348765	−0.548220	1.292366	H	−0.347368	1.302716	−0.508807
H	−0.893006	1.095379	1.121321	H	−0.881449	1.074057	1.142172
C	0.612253	0.490681	−0.318529	C	0.606146	−0.352788	0.464663
H	0.342024	0.359029	−1.373432	H	0.358816	−1.383907	0.210555
H	0.924519	1.528046	−0.205124	H	0.867527	−0.347833	1.528854
N	1.746743	−0.348067	0.088260	N	1.722325	0.046578	−0.401693
H	2.560687	−0.150152	−0.480426	H	2.033982	0.980780	−0.164673
H	1.526011	−1.327182	−0.050305	H	2.516521	−0.567375	−0.269868
C	−1.832294	−0.270600	−0.248033	C	−1.836489	−0.226680	−0.280513
H	−2.150580	0.484756	−0.966672	H	−2.146076	−0.985742	0.438497
H	−2.679739	−0.488627	0.400745	H	−2.685879	0.433346	−0.452628
H	−1.598980	−1.177648	−0.807237	H	−1.606594	−0.730830	−1.219018
tt-tg-n-propylamine				*gg'-gg*-n-propylamine			
C	0.573964	0.513709	−0.031456	C	−0.715243	0.627335	−0.288033
H	0.457494	1.187199	0.821621	H	−0.617052	0.670191	−1.374357
H	0.475303	1.126204	−0.930769	H	−1.207808	1.551412	0.024831
C	−0.553051	−0.510112	0.000460	C	0.683494	0.587209	0.322452
H	−0.412720	−1.156294	0.870963	H	0.594105	0.541298	1.410609
H	−0.475794	−1.154152	−0.878591	H	1.202043	1.520157	0.094886
N	−1.885892	0.106813	0.068284	N	1.461126	−0.560981	−0.171085
H	−1.822385	1.115244	0.051711	H	2.294082	−0.266083	−0.657836
H	−2.468136	−0.169398	−0.708152	H	1.750648	−1.173117	0.576573
C	1.954493	−0.132101	0.010777	C	−1.562795	−0.575939	0.106247
H	2.104074	−0.793205	−0.843132	H	−1.710292	−0.610865	1.187129
H	2.744314	0.617056	−0.007187	H	−2.545843	−0.542042	−0.361309
H	2.080280	−0.726670	0.915536	H	−1.075323	−1.500239	−0.198296
gg'-gt-n-propylamine				*gg-gt*-n-propylamine			
C	−0.693647	0.600901	−0.317824	C	−0.695600	0.614438	−0.321234
H	−0.585030	0.576492	−1.405514	H	−0.605135	0.607190	−1.410452
H	−1.174448	1.551587	−0.081711	H	−1.181041	1.555185	−0.054484

(continues)

(Table A.12)

	x	*y*	*z*		*x*	*y*	*z*
C	0.696505	0.594450	0.309766	C	0.698599	0.601020	0.296377
H	0.594354	0.560921	1.397630	H	0.595972	0.645270	1.384395
H	1.192885	1.537015	0.070798	H	1.229329	1.506057	−0.003538
N	1.536727	−0.517418	−0.161902	N	1.489085	−0.588018	−0.060494
H	1.888742	−1.076096	0.600977	H	0.983134	−1.193948	−0.691178
H	1.021176	−1.136512	−0.771202	H	2.351312	−0.334080	−0.518947
C	−1.585085	−0.547975	0.145290	C	−1.570897	−0.549288	0.138471
H	−1.738671	−0.506789	1.224069	H	−1.685707	−0.540015	1.222422
H	−2.562590	−0.506494	−0.332784	H	−2.563747	−0.493233	−0.305300
H	−1.148936	−1.519122	−0.086735	H	−1.145288	−1.514269	−0.135184
tt-gt-n-propylamine				*t-g*-iso-propylamine			
C	0.637255	0.556522	0.217143	C	−1.446089	−0.211207	−0.114202
H	0.381343	0.622466	1.276103	H	−2.107694	0.560519	0.281444
H	0.903064	1.568175	−0.093922	H	−1.816329	−1.183763	0.203837
C	−0.610486	0.073341	−0.560712	H	−1.487812	−0.179064	−1.203152
H	−0.357766	−0.828394	−1.120174	C	−0.009811	−0.006280	0.359865
H	−0.898836	0.817407	−1.302549	H	−0.007443	−0.031742	1.452464
N	−1.804740	−0.232615	0.228472	C	0.524626	1.350494	−0.085721
H	−1.605124	−0.952649	0.911903	H	−0.124946	2.159847	0.245043
H	−2.109522	0.583512	0.744176	H	1.523683	1.530194	0.310525
C	1.845848	−0.359540	0.043242	H	0.576171	1.392263	−1.175643
H	2.140919	−0.421394	−1.004406	N	0.840371	−1.105710	−0.137171
H	2.705777	−0.004156	0.609633	H	1.581515	−0.760137	−0.730966
H	1.618902	−1.371613	0.380712	H	1.274958	−1.615509	0.617945

Appendix B

Additional Publications

In this Appendix, two additional works are presented. These studies concern molecular complexes and, hence, they are only marginally inherent with the topics treated in this thesis. Indeed, the characterization of non-covalent adducts is an interesting branch of physical chemistry and spectroscopy, but it is a brief digression from astrochemical subjects.

B.1 Imidazole–water molecular complexes

The study reported in this Section (Ref. 309) addresses the characterization of the imidazole–water molecular complex, whose rotational spectrum has been reported for a number of isotopologues in Ref. 310. In this complex, water plays both the roles of hydrogen-bond donor and acceptor. This work exploited the TM-SE approach for the accurate determination of the equilibrium structure of the two isomeric forms, and its comparison with accurate quantum-chemical approaches, in order to assess the reliability of the method also for non-covalent complexes. In addition, the structural and energetic characteristics of the molecular complexes have been examined in detail and compared with those of analogous heterocycle–water adducts, also exploiting energy decomposition analyses.

THE JOURNAL OF PHYSICAL CHEMISTRY A
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCA

Article

Unveiling Bifunctional Hydrogen Bonding with the Help of Quantum Chemistry: The Imidazole-Water Adduct as Test Case

Published as part of The Journal of Physical Chemistry virtual special issue "Cheuk-Yiu Ng Festschrift".

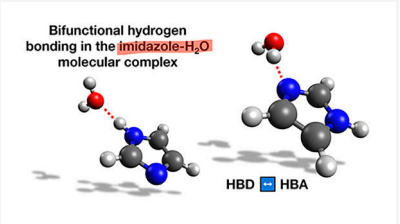
Alessio Melli, Vincenzo Barone,* and Cristina Puzzarini*

Cite This: *J. Phys. Chem. A* 2021, 125, 2989−2998

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: The ubiquitous role of water and its amphiprotic nature call for a deeper insight into the physical−chemical properties of hydrogen-bonded complexes formed with building blocks of biomolecules. In this work, the semiexperimental (SE) approach combined with the template model (TM) protocol allowed the accurate determination of the equilibrium structure of two isomeric forms of the imidazole-water complex. In this procedure, the integration of experiment (thanks to a recent rotational spectroscopy investigation) and theory is exploited, also providing the means of assessing the reliability and accuracy of different quantum-chemical approaches. Overall, this study demonstrated the robustness of the combined SE-TM approach, which can provide accurate results using affordable quantum-chemical methods. Finally, the structural and energetic characteristics of these complexes have been examined in detail and compared with those of analogous heterocycle−water adducts, also exploiting energy decomposition analyses.



INTRODUCTION

In recent years, the number of joint experimental−computational studies on noncovalent complexes has significantly increased, with the aim of obtaining a deeper knowledge of the underlying interactions.[1−9] The interest is often focused on spectroscopic and structural properties as well as on energetic characterizations, the latter often being coupled with their quantitative interpretation in terms of chemically meaningful concepts (e.g., electrostatics, induction, dispersion, etc.).[2,3,6,7,9]

Among different experimental techniques, rotational spectroscopy is the most suitable to derive structural information owing to the direct connection between rotational constants and molecular structure.[10] Rotational spectroscopy is indeed a powerful tool for molecular structure determinations, which are the mandatory prerequisites for investigating structure−activity relationships as well as deriving chemical and physical properties.[11−15] Furthermore, rotational spectroscopy investigations are performed in the gas phase, thus allowing one to focus on intrinsic effects without the perturbations introduced by the environmental effects operative in condensed phases.

Accurate structural determinations can be performed by exploiting the semiexperimental (SE) approach, which relies on extracting from the experimental outcomes the equilibrium structure details using quantum-chemical (QC) computations for providing the missing information.[16] Going more in detail, SE equilibrium structures can be obtained from a least-squares fit (LSF) of the SE equilibrium rotational constants, in turn derived from the experimental ground-state counterparts by subtracting the computed vibrational corrections.[16] Whenever experimental information is available for a sufficient number of isotopologues, a complete structural determination is possible. For small- and medium-sized organic and biological molecules, several studies have shown that vibrational corrections computed using hybrid or, even better, double-hybrid density functionals in conjunction with suitable basis sets have the required accuracy to obtain reliable results.[17,18] The application of such an approach has led to the development of a database of SE equilibrium structures containing information for an increasing number of species, and which already incorporates the most relevant prebiological building blocks.[17,18]

Received: February 24, 2021
Revised: March 22, 2021
Published: April 5, 2021

https://doi.org/10.1021/acs.jpca.1c01679
J. Phys. Chem. A 2021, 125, 2989−2998

Moving to molecular complexes, it becomes particularly challenging to obtain a set of experimental data sufficient for a complete structural determination. In particular, the most difficult information to retrieve is the position of hydrogen atoms because a full set of deuterated species is rarely experimentally available. In all cases where there are missing data, the usual approximation is to fix the intramolecular parameters at those of the isolated fragments and to fit only a limited number of intermolecular parameters. Unfortunately, this approximation is not free from difficulties, especially when flexible fragments are involved, and more advanced approaches are needed. To overcome this issue, we propose an extension of the template model approach (TMA),[17] which consists of correcting the structural parameters of a complex system, obtained at a suitable QC level, by using the corresponding SE values of suitable fragments, referred to as template models (TMs).[19] In the case of noncovalent molecular complexes, the TMs are the monomers, whose SE equilibrium structures are likely available in the above-mentioned database. The intramolecular parameters are greatly improved by the exploitation of the TMA, and they can thus be confidently kept fixed. Then, the number of experimental data becomes sufficient to optimize the intermolecular parameters.

Whenever some SE equilibrium intermolecular parameters cannot be determined, they can be obtained from partial geometry optimizations at an accurate QC level. On the basis of recent works,[20−22] the so-called jun-"cheap" composite scheme (hereafter jun-ChS) provides, for noncovalent complexes, interaction energies on par with the most sophisticated composite methods at a significantly reduced computational cost. It is, therefore, natural to investigate if this computational approach is able to also deliver accurate geometrical parameters to be employed in the interpretation and prediction of equilibrium rotational constants. In the perspective of studying even larger systems, the double-hybrid DSD-PBEP86 functional (in the recent reparametrization by Martin and co-workers,[23] rev-DSD-PBEP86) in conjunction with the jun-cc-pVTZ basis set[24,25] appears particularly promising and will be compared to jun-ChS results.

Once the methodology is defined and the computational level is selected, the next step is the choice of a challenging, yet representative, test case to assess the methodology. Among different noncovalent interactions, hydrogen bonds play a central role, especially for biological systems. Therefore, the interaction between building blocks of biomolecules and water represents a natural starting point for our analysis, also in view of the ubiquitous presence and the amphiprotic character of water. Focusing on the potential partner of the molecular complex, heterocyclic compounds appear interesting in view of the prebiotic character of the species containing the C−N moiety as well as their presence in important biological molecules. In this framework, imidazole is particularly appealing because of the concomitant presence of donor and acceptor sites.

In conclusion, this study will address the characterization (by means of the methodologies mentioned above) of the water−imidazole complex, whose rotational spectrum has recently been reported for a number of different isotopologues.[26] Interestingly, two different adducts have been experimentally observed (see Figure 1) in which water plays the role of either hydrogen bond (HB) donor (HBD) or acceptor (HBA).[26] On the contrary, in a previous experimental investigation with supersonic jet FTIR spectroscopy, only the HBD complex was detected, thus suggesting that it is significantly more stable than the HBA counterpart.[27] However, the study of ref 26 was not able to experimentally determine the relative stability of the two adducts. For this reason, an accurate energetic characterization is also deserved.

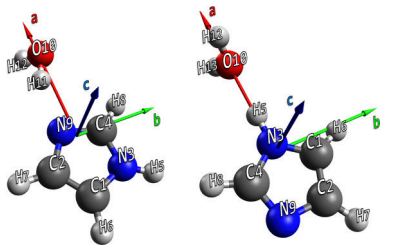


Figure 1. H_2O−Imid (left) and Imid−H_2O (right) structures (rev-DSD level) together with the atom numbering.

This manuscript is organized as follows. In the next section, the computational methodology is presented in all details. The following section will report and discuss the results of structural and energetic investigations. Finally, concluding remarks will be provided.

COMPUTATIONAL METHODOLOGY

The equilibrium structure of both isomers of the imidazole−water molecular complex (see Figure 1) has been evaluated by exploiting QC composite schemes as well as the SE approach. Both of them are described with some details in the following.

As briefly mentioned in the Introduction, within the SE methodology, the equilibrium structural parameters of the investigated system are determined from a LSF of the equilibrium moments of inertia of different isotopologues.[16] However, laboratory measurements give access to vibrational ground-state rotational constants (B_0^i, where i refers to the inertial axes a, b, c), thus requiring the vibrational corrections (ΔB_{vib}^i) to be subtracted in order to obtain the corresponding SE equilibrium rotational constants (B_e^i).[16] From the latter, the SE equilibrium moments of inertia are straightforwardly derived, these being inversely proportional to rotational constants.[10]

To better clarify the statements above, it has to be noted that, in the framework of vibrational perturbation theory to the second order (VPT2),[28] the equilibrium rotational constants can be expressed as

$$B_e^i = B_0^i - \Delta B_{vib}^i = B_0^i + \frac{1}{2}\sum_r \alpha_r^i \quad (1)$$

where the α_r's are the vibration−rotation interaction constants and the sum runs over all r vibrational modes. Noted is that the evaluation of the α_r's implies anharmonic force field calculations (for details, see, e.g., refs 29−32).

To recap, in the SE procedure, the second term on the right-hand side of eq 1 (i.e., ΔB_{vib}^i) is derived from QC calculations, while the first term is obtained from the experiment. ΔB_{vib}^i being significantly smaller than B_e^i, the former term can be determined at an affordable level of theory (such as methods rooted in the density functional theory, DFT) without

https://doi.org/10.1021/acs.jpca.1c01679
J. Phys. Chem. A 2021, 125, 2989−2998

Table 1. Equilibrium Geometries of H_2O−Imid and Imid−H_2O at Different Levels of Theory (Bonds in Å; Angles in deg)

H_2O−Imid	rev-DSD	ChS	ChS-mAUG	jun-ChS	Imid−H_2O	rev-DSD	ChS	ChS-mAUG	jun-ChS
C1−C2	1.3693	1.3652	1.3647	1.3647		1.3720	1.3679	1.3673	1.3673
N3−C1	1.3780	1.3755	1.3747	1.3749		1.3755	1.3731	1.3721	1.3723
C4−N3	1.3589	1.3553	1.3544	1.3547		1.3604	1.3569	1.3560	1.3562
N3−H5	1.0053	1.0031	1.0021	1.0022		1.0122	1.0086	1.0084	1.0085
C1−H6	1.0762	1.0741	1.0733	1.0735		1.0767	1.0747	1.0738	1.0740
C2−H7	1.0774	1.0753	1.0746	1.0747		1.0780	1.0758	1.0751	1.0753
C4−H8	1.0784	1.0769	1.0759	1.0761		1.0787	1.0764	1.0758	1.0760
C4−N9	1.3168	1.3122	1.3120	1.3122		1.3176	1.3132	1.3125	1.3125
O10−N9	2.8714	2.8470	2.8612	2.8594	O10−N3	2.9696	2.9773	2.9581	2.9644
O10−H11	0.9749	0.9733	0.9689	0.9684	O10−H12/13	0.9613	0.9582	0.9565	0.9565
O10−H12	0.9600	0.9569	0.9554	0.9554					
N3−C1−C2	105.26	105.28	105.26	105.27		105.44	105.45	105.46	105.46
C4−N3−C1	107.57	107.57	107.51	107.49		106.99	107.09	106.95	106.93
H5−N3−C4	126.10	126.06	126.12	126.12		126.39	126.39	126.43	126.41
H6−C1−C2	132.44	132.49	132.45	132.44		132.41	132.43	132.41	132.41
H7−C2−C1	128.20	128.28	128.17	128.17		128.03	128.15	128.03	128.02
H8−C4−C2	163.61	163.82	163.51	163.48		164.14	164.20	164.17	164.18
N9−C4−C2	38.17	38.16	38.23	38.25		38.53	38.46	38.57	38.59
O10−N9−C4	99.74	103.48	95.31	92.94	O10−N3−C1	126.92	126.72	126.87	127.42
H11−O10−C4	35.25	32.46	38.13	40.35	X11−O10−N3	163.51	187.85	172.45	174.07
H12−O10−H11	104.71	106.32	104.91	104.95	H12/13−O10−X11	52.64	52.95	52.74	52.71
O10−N9−C4−C2	174.76	171.56	174.38	176.24	H13−O10−H12	105.28	105.90	105.48	105.42
H11−O10−C4−C2	−4.86	−11.93	−4.59	−3.67	H12−O10−X11−C1	90.0	90.0	90.0	90.0
H12−O10−H11−C4	128.89	154.54	127.37	138.34	H13−O10−X11−C1	−90.0	−90.0	−90.0	−90.0

significantly affecting the accuracy of the resulting SE equilibrium rotational constants.[11,18,29,33] The calculation of the α_r's from DFT anharmonic force fields led to the collection of accurate SE equilibrium geometries for several systems, ranging from isolated molecules to clusters.[3,6,9,17,18,34−38] In this work, the global-hybrid B3LYP functional,[39,40] also incorporating the Grimme's DFT-D3 scheme[41] for the treatment of dispersion effects in conjunction with the Becke−Johnson (BJ) damping function,[42] has been used for the calculation of the vibrational contributions. The partially augmented double-ζ jun-cc-pVDZ basis set[24,25] has been employed in combination with B3LYP-D3(BJ). Hereafter, this level of theory is shortly referred to as B3.

The starting point of the LSF procedure is a guess geometry obtained by means of the TMA.[18] In the present context, the TM species are the monomers, imidazole and water, and their structural parameters within the complex have been adjusted using the SE equilibrium geometry of the isolated fragments.[a] For a generic intramolecular equilibrium structural parameter r, which has been optimized at the x level of theory, the TMA parameter (r_{TMA}) is calculated as

$$r_{TMA} = r_x + \Delta TM \tag{2}$$

where

$$\Delta TM = r_{SE}^{TM} - r_x^{TM} \tag{3}$$

In eqs 2 and 3, TM is the isolated fragment. The rev-DSD-PBEP86-D3(BJ) functional[23] in conjunction with the jun-cc-pVTZ basis set[24,25] (hereafter denoted as rev-DSD) has been considered as the x level of theory in the equations above because it is proven to offer a very good description of noncovalent complexes.[43]

In addition to the SE approach, the equilibrium structure of the molecular complex has been evaluated using the so-called "cheap" composite scheme (hereafter ChS).[44] The original version of this accurate yet cost-effective approach for medium-sized systems[45,46] is the following:[44]

$$r_{ChS} = r(CCSD(T)/VTZ) + \Delta r(MP2/CBS) + \Delta r(MP2/CV) + \Delta r(MP2/AUG) \tag{4}$$

The terms on the right-hand side are presented here below in order of appearance:

i. The starting point is the coupled-cluster ansatz including single and double excitations with a perturbative treatment of triples, CCSD(T),[47] within the frozen-core (fc-) approximation and in conjunction with the cc-pVTZ basis set.[24]

ii. The correction due to the extrapolation to the complete basis set (CBS) limit is evaluated using the n^{-3} formula[48] and applied to the geometrical parameters optimized using Møller−Plesset perturbation theory to second order[49] (fc-MP2). The cc-pVnZ (with n = T,Q) basis sets are employed.

iii. The core−valence (CV) correlation contribution is calculated as the difference between all-MP2/cc-pCVTZ[50] and fc-MP2/cc-pCVTZ optimized parameters, where "all-" denotes the correlation of all electrons.

iv. The effect of the inclusion of diffuse functions in the basis set is estimated as the difference between fc-MP2/aug-cc-pVTZ[24,51] and fc-MP2/cc-pVTZ optimized parameters.

The jun-ChS variant[20] is obtained by employing the partially augmented jun-cc-pVnZ basis sets[24,25] (in the place of cc-pVnZ). Since these sets already incorporate diffuse functions, the iv term is not considered in conjunction with them. Instead, the iii term is the same in both approaches. If the contribution of diffuse functions (iv) is neglected in the original scheme, the "ChS-mAUG" model is obtained.

https://doi.org/10.1021/acs.jpca.1c01679
J. Phys. Chem. A 2021, 125, 2989−2998

All DFT and MP2 calculations have been carried out using the Gaussian16 suite,[52] while the CCSD(T) computations have been performed with the CFOUR package.[53]

RESULTS AND DISCUSSION

In agreement with the IUPAC definition of hydrogen bond,[54] contrary to what used in ref 26, we denote as Imid−H_2O the isomer where the imidazole ring acts as HBD, and as H_2O−Imid the isomer where it acts as HBA.

Computational Results. A preliminary scan of the potential energy surface (PES) of the molecular imidazole−water complex has been carried out at the B3 level and confirmed the DFT results reported in ref 26. Next, the geometries of Imid−H_2O and H_2O−Imid have been optimized at the rev-DSD, ChS, ChS-mAUG, and jun-ChS levels of theory. Their rev-DSD structures are sketched in Figure 1, while the geometrical parameters obtained at the different computational levels are collected in Table 1. From an inspection of the results collected in this table, we note that the four different approaches provide very similar results for the intramolecular parameters, with rev-DSD showing deviations within a few mÅ for distances and well within 0.1° for angles. Larger deviations, even within the different ChS variants, are noted for intermolecular parameters and, in particular, for the N···O distance. From Table 1, the cost-effectiveness of the rev-DSD level is evident: while providing results in nearly quantitative agreement with those issuing from composite schemes, its cost is comparable with the cheapest step of the latter approaches (i.e., MP2 in conjunction with a triple-ζ basis set).

Unlike H_2O−Imid, the structure of Imid−H_2O is characterized by a symmetry plane containing the imidazole ring and the oxygen atom of the water molecule. Thus, Imid−H_2O belongs to the C_s symmetry point group. In order to ease the interpretation of the geometrical parameters, we have used a dummy atom (X) placed on the HOH angle bisector to locate the water molecule with respect to the ring (see Table 1). In addition to the symmetry issue, the HB geometry presents some slight differences between the two isomers. On average, the computed N···O distance is about 0.11 Å longer in Imid−H_2O than in H_2O−Imid. Furthermore, the O−H bond in H_2O−Imid and the N−H distance in Imid−H_2O, which are involved in the HB donation, show stretches of about 13 and 6 mÅ, respectively, with respect to the isolated fragments. Concerning Imid−H_2O, we note that, for the X11−O10−N3 angle, the original formulation of the ChS scheme leads to a value ∼15° larger than that obtained with jun-ChS and ChS-mAUG. This implies that ChS orients the water molecule differently. As expected, no significant differences are noted for the other geometrical parameters, the only exception being the HOH angle of water, which is overestimated at the ChS level in both H_2O−Imid and Imid−H_2O complexes. Noted is that the dihedral angles not explicitly reported in Table 1 are either 0° or 180°, as a consequence of the planarity of the imidazole ring.

For comparison purposes, we have also computed the structure of several complexes formed by water with N-containing heterocycles (Het) and reported the O···N distance at the rev-DSD level of theory (see Table 2). From the analysis of the results obtained, a systematic increase of the HB length when the heterocycle acts as HBD is noted, with a difference—by averaging on the two families—of 0.1 Å.

Table 2. Comparison between the O···N Distances in Several Het−H_2O and H_2O−Het Complexes at the rev-DSD Level (Bonds in Å)

heterocycle	r(O···N)
Het−H_2O	
indole	2.99
pyrrole	3.00
imidazole	2.97
H_2O−Het	
pyrazine	2.89
pyridazine	2.86
pyridine	2.89
pyrimidine	2.88
imidazole	2.87

Table 3 collects the rotational constants. While the B_e's have been straightforwardly derived from the equilibrium structures,[10] the B_0's have been obtained by augmenting the former with the B3 vibrational corrections. Then, the B_0's can be directly compared with the experimental counterparts. For H_2O−Imid, the ChS results are in remarkable agreement with the experiment,[26] indeed showing an average relative error of 0.16%. The origin of the worse performance of both ChS-mAUG and jun-ChS models can be traced back to the slightly different water−imidazole relative position. In the H_2O−Imid complex, the water molecule is characterized by a high mobility. As a consequence, different levels of theory lead to rather different values of the dihedral angles describing the position of water with respect to the imidazole ring. Deviations from the seemingly correct ChS value are noted when either neglecting diffuse functions (ChS-mAUG) or replacing the additive approximation by their systematic inclusion (jun-ChS), thus leading to larger errors on the B and C rotational constants. Moving to Imid−H_2O, a remarkable agreement is found for all ChS variants, with the jun-ChS approach showing the best performance with an average relative error below 0.1%. Similarly to H_2O−Imid, at different levels of theory, the N···O···X and O···N−C1 angles (which describe the water position) show differences when compared with each other, thus affecting the rotational constants determination. This is evident for the A constant of Imid−H_2O: at the ChS level it

Table 3. Rotational Constants for H_2O−Imid and Imid−H_2O (All Parameters in MHz)

	rev-DSD	ChS	ChS-mAUG	jun-ChS	exp[26]
			H_2O−Imid		
A_e	9481.5	9506.2	9519.3	9533.6	
B_e	1830.2	1835.5	1866.8	1879.0	
C_e	1539.9	1544.0	1567.0	1573.7	
A_0	9439.7	9464.5	9477.5	9491.8	9502.79(43)
B_0	1821.2	1826.6	1857.9	1870.1	1826.3929(13)
C_0	1526.8	1531.0	1554.0	1560.7	1531.9303(10)
			Imid−H_2O		
A_e	9504.5	9538.6	9564.3	9561.4	
B_e	1673.9	1672.4	1684.5	1680.2	
C_e	1432.7	1432.4	1441.8	1438.5	
A_0	9462.8	9496.9	9522.6	9519.8	9520.99(26)[a]
B_0	1657.8	1656.3	1668.4	1664.1	1662.91298(73)
C_0	1417.2	1416.9	1426.3	1423.0	1420.74502(72)

[a]The 0^+ state is reported.

https://doi.org/10.1021/acs.jpca.1c01679
J. Phys. Chem. A 2021, 125, 2989−2998

Table 4. Nuclear Quadrupole Coupling Constants of H_2O−Imid and Imid−H_2O Computed at Different Levels of Thoery (All Parameters in MHz)

		theory[26,a]	rev-DSD	CC[b]	jun-ChS	jun-ChS+vib[c]	exp[26]
				H_2O−Imid			
N3	χ_{aa}	1.192	1.102	1.136	1.173	1.120	1.143(19)
	$(\chi_{bb} - \chi_{cc})$	3.858	3.740	3.828	4.015	3.906	3.712(56)
N9	χ_{aa}	−3.783	−3.099	−2.892	−2.880	−2.825	−2.889(21)
	$(\chi_{bb} - \chi_{cc})$	−0.490	−0.847	−0.895	−1.128	−1.147	−1.07(11)
				Imid−H_2O			
N3	χ_{aa}	0.815	0.767	0.821	0.890	0.872	0.916(45)[d]
	$(\chi_{bb} - \chi_{cc})$	3.451	3.252	3.398	3.529	3.489	4.07(23)
N9	χ_{aa}	−2.267	−2.052	−1.970	−2.069	−2.013	−1.859(44)
	$(\chi_{bb} - \chi_{cc})$	−2.680	−2.052	−2.354	−2.487	−2.475	−2.47(25)

[a]B3LYP-D3(BJ)/aug-cc-pVTZ level of theory. [b]CCSD(T)/jun-cc-pVTZ level of theory. [c]The jun-ChS equilibrium values have been augmented by the vibrational corrections at the B3 level. [d]The 0^+ state is reported.

Table 5. Equilibrium Structures of Imidazole and Water (Bonds in Å; Angles in deg)

	SE*	rev-DSD	ChS	ChS-mAUG	jun-ChS
			Imid		
C1−C2	1.3624	1.3704	1.3665	1.3658	1.3658
N3−C1	1.3741	1.3775	1.3750	1.3741	1.3743
C4−N3	1.3613	1.3629	1.3594	1.3586	1.3588
N3−H5	1.0016	1.0051	1.0029	1.0019	1.0020
C1−H6	1.0765	1.0765	1.0744	1.0735	1.0737
C2−H7	1.0755	1.0778	1.0755	1.0749	1.0750
C4−H8	1.0772	1.0785	1.0763	1.0756	1.0758
C4−N9	1.3101	1.3148	1.3106	1.3098	1.3098
N3−C1−C2	105.42	105.15	105.19	105.17	105.17
C4−N3−C1	107.01	107.30	107.34	107.25	107.22
H5−N3−C4	126.20	126.36	126.29	126.38	126.38
H6−C1−C2	132.65	132.57	132.60	132.58	132.59
H7−C2−C1	127.89	127.94	128.07	127.95	127.95
H8−C4−C2	164.36	164.34	164.34	164.34	164.35
N9−C4−C2	38.61	38.45	38.39	38.49	38.51
			H_2O		
O−H	0.9573	0.9610	0.9586	0.9563	0.9563
H−O−H	104.53	104.46	105.15	104.56	104.48

[a]Asterisk refers to the "Additional notes" section.

deviates by 0.25%, while the relative discrepancy decreases by 1 order of magnitude moving to ChS-mAUG (0.02%) and jun-ChS (0.03%). This can be associated with the larger value, mentioned above, of ∠(X11−O10−N3), in the case of ChS.

Concerning rev-DSD, which is—as already mentioned—a level of theory affordable also for larger systems, the calculated B_0^i's are in good agreement with the experimental ones, indeed showing a relative error of about 0.4% for both H_2O−Imid and Imid−H_2O complexes, which improves to 0.3% if one does not consider the A constant. In this respect, if we compare our rev-DSD/B3 results (with B3 referring to the vibrational corrections) with the current practice of directly using the B_e^i's, evaluated from global-hybrid DFT calculations, for guiding spectral recording and analysis (as done, for example, in ref 26), a reduction of the discrepancies ranging from a factor of 2 to more than 1 order of magnitude is evident.

Finally, the imidazole ring is characterized by the presence of two nitrogen atoms, which are quadrupolar nuclei. Having significantly different local environments, the corresponding quadrupole coupling interactions lead to distinctive spectroscopic features.[26] Therefore, we decided to compare the experimental nuclear quadrupole coupling constants (χ_{ii}'s) with the calculated ones. The results, reported in Table 4, show an overall good agreement with experiment at any level of theory considered, with the largest discrepancies being observed for the experimental data affected by large uncertainty. A remarkable improvement is noted when going from B3LYP[26] to rev-DSD and CCSD(T)/jun-cc-pVTZ. Application of the jun-ChS approach only marginally changes the constants with respect to the latter level of theory.

Semiexperimental Structure. As mentioned in the Computational Methodology section, the determination of the SE equilibrium structure for the two molecular complexes starts from a guess geometry obtained with the TMA. For the exploitation of the SE approach, the experimental rotational constants of four and three isotopic species (reported in ref 26) for H_2O−Imid and Imid−H_2O, respectively, have been employed. The number of experimental data not being sufficient for a complete geometry evaluation, the LSF procedure has been applied to a reduced number of intermolecular structural parameters, while keeping the intramolecular and the other intermolecular parameters fixed at the guess values. Therefore, first of all, the accuracy of the

https://doi.org/10.1021/acs.jpca.1c01679
J. Phys. Chem. A 2021, 125, 2989−2998

equilibrium structures of the monomers, which are at the basis of the TMA, needs to be discussed.

Focusing on the isolated fragments (see Table 5), it is noted that all ChS schemes show a very good performance in terms of deviation from the SE equilibrium values. An inspection of Table 5 confirms the accuracy expected on the basis of the literature on this topic: 0.001−0.002 Å for bond lengths and 0.1−0.2° for angles.[44,45,55] A slight overestimation of the water angle is observed at the ChS level. A good agreement is also found between the rev-DSD and SE values, with an average error of 3 mÅ and 0.15° for bond lengths and angles, respectively.

The LSF procedure has been carried out using the molecular structure refinement (MSR) software.[56] For H_2O−Imid, the fit has been carried out using all rotational constants equally weighted, while the A's have been excluded from the fit of Imid−H_2O. Different sets of determinable geometrical parameters have been tested in the LSF procedure in order to obtain the most robust fit. Among these tests, inclusion of ∠(H−O−H) leads to precise (in terms of the statistical measures), but inaccurate, results for this angle. For instance, in the case of the Imid−H_2O isomer, the use of the TMA-rev-DSD geometry in the LSF leads to r(N···O) = 2.96855(7) Å and ∠(H−O−H) = 98.8(1)°, whereas the TMA-jun-ChS structure leads to r(N···O) = 2.96912(4) Å and ∠(H−O−H) = 99.42(5)°. In order to better analyze these trends, a two-dimensional scan has been carried out at the rev-DSD level, thereby exploring the portion of the PES ruled by the ∠(N···O···X) (143.5−183.5°) and ∠(H−O−H) (from 98.5° to more than 115°) angles. As clearly evident in the bottom panel of Figure 2, those structures with ∠(H−O−H) angles close to 105.5° are significantly more stable. To confirm the position of the minimum, a more focused second scan of the two-dimensional PES has been performed. The portion with ∠(H−O−H) = 104°−106.5° has been probed, using the previous range for the other coordinate (see Figure 2). These

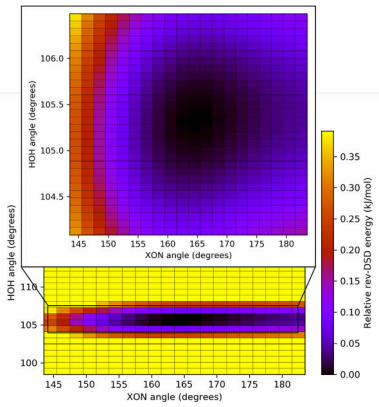


Figure 2. Two-dimensional scan of the PES near to the Imid−H_2O minimum.

investigations locate the minimum in the same position found in all geometry optimizations, therefore leading to the exclusion of the outcomes of this fitting procedure.

The source of the anomalous results addressed above may be traced back to the water mobility. Indeed, the presence of large amplitude motions has a huge impact on the geometry and questions the foundations of the SE approach, which relies on second-order vibrational perturbation theory. Experimental information on isotopic substitution at the hydrogen−water atoms would have helped in correctly deriving their positions within the molecular adduct. This cannot be accomplished by fitting instead intramolecular parameters such as the ∠(H−O−H) angle, the overall conclusion being the exclusion of any structural parameter involving water hydrogens from our LSF procedure. A further evidence of this problem has been met in the fit of Imid−H_2O. For this isomer, the value of the moment of inertia along the a axis (and, therefore, the rotational constant A) is strongly affected by the position of the hydrogen atoms of the water molecule, which, however, experience large amplitude motions. Within the VPT2 framework, the treatment of the latter represents a well-known issue,[57] which has been confirmed, and partially solved, for this adduct by the exclusion of the A's from the LSF procedure.

The results obtained by employing the TMA (fit 1) or theoretical (fit 2) intramolecular parameters are reported in Table 6 and confirm that the O···N distance is shorter in H_2O−Imid than in Imid−H_2O by at least 0.1 Å. The already mentioned difficulties in positioning the water molecule with respect to imidazole are at the origin of the slight differences in the SE equilibrium structure of H_2O−Imid derived from the different fits. However, by comparing the results of fit 1 and fit 2 and taking into consideration the confidence interval of one standard deviation, we note that the fitted values are comparable. For Imid−H_2O, the O···N distance as well as the ∠(N···O···X) and ∠(O···N−C1) angles are in remarkable agreement. Larger deviations are noted for the ChS results for both isomers, probably due to a greater difference in the "guess" values for the water molecule. Interestingly, the comparison of the results of fit 1 points out the strength of the TMA, which is able to provide reliable geometries even when the reference structure is obtained at the DFT level (instead of a computational expensive CC-based composite scheme).

Energetics. The energetic characterization of the two isomers has been carried out at both the jun-ChS and rev-DSD levels, with the former model for interaction energies being fully described in ref 20. Actually, the different "cheap" expressions can be derived from eq 4 by replacing the geometrical parameter r with the total energy. In particular, the jun-ChS model chemistry is proven to provide a very good compromise between accuracy and computational cost in the description of noncovalent interactions.[20] Furthermore, two different reference geometries (rev-DSD and jun-ChS) have been used in order to determine the influence of the structure on the energetics. The results are reported in Table 7. According to them, H_2O−Imid is about 6.2 kJ/mol more stable than the Imid−H_2O isomer, in agreement with both a previous FTIR experiment[27] and the findings of ref 26. With respect to the relative stability, the effect of the different geometry on the energy is negligible (i.e., less than 0.1 kJ/mol).

The jun-ChS and rev-DSD levels have also been employed to evaluate the interaction energy (int en), possibly accounting

https://doi.org/10.1021/acs.jpca.1c01679
J. Phys. Chem. A 2021, 125, 2989−2998

Table 6. Semiexperimental Equilibrium H_2O−Imid and Imid−H_2O Intermolecular Parameters (Bonds in Å; Angles in deg)[a]

	method	guess value	fit 1	fit 2[b]
		H_2O−Imid		
r(O···N)	rev-DSD	2.8714	2.86(2)	2.771(8)
	ChS	2.8470	2.80(2)	2.79(1)
	ChS-mAUG	2.8612	2.85(2)	2.87(2)
	jun-ChS	2.8594	2.82(4)	2.83(4)
∠(O···N−C4)	rev-DSD	99.74	101(3)	121(4)
	ChS	103.48	112(5)	117(5)
	ChS-mAUG	95.31	102(3)	100(3)
	jun-ChS	92.94	106(7)	104(7)
		Imid−H_2O		
r(N···O)	rev-DSD	2.9696	2.987(2)	2.9637(4)
	ChS	2.9773	2.981(1)	2.981(2)
	ChS-mAUG	2.9581	2.987(2)	2.995(2)
	jun-ChS	2.9644	2.986(2)	2.992(2)
∠(O···N−C1)	rev-DSD	126.92	137.0(6)	125(1)
	ChS	126.72	134.6(6)	133.6(7)
	ChS-mAUG	126.87	137.1(7)	138.4(7)
	jun-ChS	127.42	136.8(7)	137.9(7)
∠(N···O···X)	rev-DSD	163.51	155.3(6)	157.7(7)
	ChS	187.85	161.4(6)	162.0(7)
	ChS-mAUG	172.45	155.5(6)	155.4(7)
	jun-ChS	174.07	155.6(6)	155.4(6)

[a]The error within brackets is one standard deviation. [b]No TMA has been used to correct the intramolecular parameters.

Table 7. Energetics of the Two Imidazole−Water Complexes (All Results in kJ/mol)

			int en (jun-ChS)		int en (rev-DSD)	
	ref geom	rel en (jun-ChS)	CP	NCP	CP	NCP
H_2O−Imid	rev-DSD	0.00	−32.54	−32.76	−31.34	−32.91
	jun-ChS	0.00	−32.24	−32.45	−30.83	−32.38
Imid−H_2O	rev-DSD	6.20	−25.89	−25.97	−24.61	−25.97
	jun-ChS	6.27	−25.77	−25.86	−24.45	−25.80

for the basis set superposition error (BSSE) by means of the counterpoise (CP) correction.[58] The results (Table 7) show a good agreement between the two levels of theory (within 0.1 kJ/mol) when CP corrections are not incorporated (denoted as NCP). As expected, NCP and CP results are quite similar to one another at the jun-ChS level (within 0.2 kJ/mol) because of the extrapolation to the CBS limit, whereas the CP correction significantly worsens (by about 1.5 kJ/mol) the rev-DSD results (with respect to jun-ChS). It is thus confirmed that jun-ChS computations (for both energies and geometries) can be safely performed without any CP correction. The same applies to rev-DSD results, possibly due to to a fortuitous error compensation and/or the conditions employed in the parametrization of the functional.

From a close inspection of the results of Table 7, a conclusion on the deformation experienced by the monomers (from isolated to part of the adduct) can be drawn. The difference between the jun-ChS-CP interaction energies of the two isomers (6.65 and 6.46 kJ/mol at the rev-DSD and jun-ChS geometries, respectively) is mainly ascribable to the relative jun-ChS electronic energy (6.20 and 6.27 kJ/mol at the two reference geometries). The remaining contribution (0.45 and 0.20 kJ/mol), due to the deformation of the monomers, slightly favors the Imid−H_2O isomer.

While the interaction energy clearly points out a stronger interaction in H_2O−Imid, a deeper insight on the origin of different HB patterns can be gained by the natural energy decomposition analysis (NEDA).[59] This has been carried out at the B3 level using the NBO 7.0[60] suite of programs interfaced to the latest revision of the Gaussian16 program. Furthermore, in order to obtain general information, such an analysis has been extended to different N-heterocycle−water adducts. The results, collected in Table 8, show a close agreement between B3 (last column) and rev-DSD (last column of Table 7) total energies for the case of imidazole and point out a clear trend of the two hydrogen bond patterns

Table 8. Energy Decomposition Analysis[a] for Het−H_2O and H_2O−Het Complexes (Values in kJ/mol)

heterocycle	El[b]	CT[c]	core	total
		Het−H_2O		
indole	−45.0	−44.2	65.8	−23.4
pyrrole	−42.1	−40.4	60.7	−21.8
imidazole	−47.2	−44.7	66.7	−25.2
		H_2O−Het		
pyrazine	−55.8	−49.4	78.2	−27.0
pyridazine	−61.7	−57.2	87.5	−31.4
pyridine	−60.8	−60.0	89.3	−31.5
pyrimidine	−58.2	−50.8	80.6	−28.4
imidazole	−63.7	−58.9	89.2	−33.4

[a]Calculated at the B3LYP-D3(BJ)/jun-cc-pVTZ level of theory, using rev-DSD optimized geometries. [b]Electrostatic plus polarization interaction. [c]Charge transfer.

https://doi.org/10.1021/acs.jpca.1c01679
J. Phys. Chem. A 2021, 125, 2989−2998

(Het−H_2O and H_2O−Het). When the N-heterocycle molecule acts as HBD, the interaction is weaker than that experienced in those adducts where water acts as HBD, this outcome being in agreement with the basicity of the heterocycles and the O···N distances reported in Table 2. Within the two groups, no significant difference in the various contributions to the total interaction energy can be noted. Focusing on the total energy, the results within the two group are in line with the acidity and basicity of imidazole, which always shows the largest interaction energy. Indeed, the pK_A value of imidazole is 14.5, thus being more acidic than pyrrole. On the other hand, the pK_A value of the conjugated acid is ∼7, which means that imidazole is about 60 times more basic than pyridine.

The two attractive terms, the electrical interaction (which includes the electrostatic and polarization contributions) and the charge transfer, are similar, with the maximum discrepancy being smaller than 7.5 kJ/mol. Furthermore, similar values are also found for the repulsive core interaction, with average values of 64.4 and 85.0 kJ/mol for Het−H_2O and H_2O−Het, respectively. In both the investigated groups, imidazole is found to have the strongest interaction with the water molecule.

CONCLUSIONS

A computational characterization of H_2O−Imid and Imid−H_2O, two isomers of the imidazole−water complex, has been carried out using different levels of theory for the geometry optimization, also aiming to test the reliability of the jun-ChS composite scheme and the DSD-PBEP86 double-hybrid density functional in its most recent reparametrization. To check the performance of these two levels of theory, we relied on the results obtained in a recent experimental work based on rotational spectroscopy.[26] Indeed, rotational constants intrinsically contain structural information. We proceeded in two different ways: we derived the rotational constants from our structures to be compared with the experimental ones and we used the latter, available for different isotopic species, in the semiexperimental approach for deriving accurate structures (to be compared with the computed ones). Both strategies required the computation of the vibrational corrections to rotational constants (either to be added to the computed equilibrium rotational constants or to be subtracted from the experimental ground-state rotational constants), which have been determined from B3LYP-D3(BJ)/jun-cc-pVDZ anharmonic calculations within the VPT2 model. While both strategies pointed out the accuracy of the jun-ChS model applied to geometries and the reliability of the rev-DSD level of theory, some discrepancies have been found that could be ascribed to the high flexibility of the water molecule position within the adduct. Further developments in the treatment of large amplitude motions will be addressed in a future work.

The semiexperimental equilibrium structure of the two isomers has been determined by means of a least-squares fit of the SE equilibrium moments of inertia for several isotopologues, with the intramolecular parameters being kept fixed. The values used for the latter have been derived from the application of the template model approach. One important finding is that the use of this approach leads to robust and reliable results, which are independent of the level of theory employed. The overall conclusion is that the rev-DSD-PBEP86-D3(BJ)/jun-cc-pVTZ level of theory in conjunction with the TMA is a valuable choice for the determination of the equilibrium structure of medium-sized molecular complexes at a reasonable computational cost.

Finally, an accurate energetic characterization of the imidazole−water complexes has been carried out. In addition, their energy decomposition analysis (EDA) is reported also for several N-heterocycle complexes with water for comparison purposes. The EDA pointed out that the different contributions are essentially unchanged within the series of adducts considered.

In our opinion, together with the interest of the studied system, the proposed computational strategy paves the route toward accurate structural and energetic characterizations for large noncovalent complexes of current technological and biological interest.

AUTHOR INFORMATION

Corresponding Authors

Vincenzo Barone − *Scuola Normale Superiore, 56126 Pisa, Italy*; orcid.org/0000-0001-6420-4107; Email: vincenzo.barone@sns.it

Cristina Puzzarini − *Dipartimento di Chimica "Giacomo Ciamician", Università di Bologna, 40126 Bologna, Italy*; orcid.org/0000-0002-2395-8532; Email: cristina.puzzarini@unibo.it

Author

Alessio Melli − *Scuola Normale Superiore, 56126 Pisa, Italy; Dipartimento di Chimica "Giacomo Ciamician", Università di Bologna, 40126 Bologna, Italy*; orcid.org/0000-0002-8469-1624

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.1c01679

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work has been supported by MIUR (Grant Number 2017A4XRCA) and by the University of Bologna (RFO funds). The SMART@SNS Laboratory (http://smart.sns.it) is acknowledged for providing high-performance computing facilities.

ADDITIONAL NOTE

[a]The SE equilibrium structures of imidazole and water are available for download from smart.sns.it.

REFERENCES

(1) Hobza, P.; Muller-Dethlefs, K. *Non-Covalent Interactions; Theoretical and Computational Chemistry Series*; The Royal Society of Chemistry, 2009.

(2) Li, W.; Spada, L.; Tasinato, N.; Rampino, S.; Evangelisti, L.; Gualandi, A.; Cozzi, P. G.; Melandri, S.; Barone, V.; Puzzarini, C. Theory meets experiment for noncovalent complexes: The puzzling case of pnicogen interactions. *Angew. Chem., Int. Ed.* **2018**, *57*, 13853−13857.

(3) Obenchain, D. A.; Spada, L.; Alessandrini, S.; Rampino, S.; Herbers, S.; Tasinato, N.; Mendolicchio, M.; Kraus, P.; Gauss, J.; Puzzarini, C.; et al. Unveiling the sulfur−sulfur bridge: Accurate structural and energetic characterization of a homochalcogen intermolecular bond. *Angew. Chem., Int. Ed.* **2018**, *57*, 15822−15826.

(4) Kraus, P.; Obenchain, A. D.; Frank, I. Benchmark-quality semiexperimental structural parameters of van der Waals complexes. *J. Phys. Chem. A* **2018**, *122*, 1077−1087.

(5) Kraus, P.; Frank, I. Density functional theory for microwave spectra of noncovalent complexes: A benchmark study. *J. Phys. Chem. A* **2018**, *122*, 4894−4901.
(6) Wang, J.; Spada, L.; Chen, J.; Gao, S.; Alessandrini, S.; Feng, G.; Puzzarini, C.; Gou, Q.; Grabow, J.-U.; Barone, V. The unexplored world of cycloalkene−water complexes: Primary and assisting interactions unraveled by experimental and computational spectroscopy. *Angew. Chem., Int. Ed.* **2019**, *58*, 13935−13941.
(7) Li, W.; Li, M.; Jin, Y.; Gou, Q.; Grabow, J.-U.; Feng, G. Molecular structure and non-covalent interaction of 2-thiophenecarboxaldehyde and its monohydrated complex. *J. Chem. Phys.* **2019**, *151*, 164307.
(8) Al-Hamdani, Y. S.; Tkatchenko, A. Understanding non-covalent interactions in larger molecular complexes from first principles. *J. Chem. Phys.* **2019**, *150*, 010901.
(9) Chen, J.; Zheng, Y.; Melli, A.; Spada, L.; Lu, T.; Feng, G.; Gou, Q.; Barone, V.; Puzzarini, C. Theory meets experiment for elucidating the structure and stability of non-covalent complexes: Water−amine interaction as a proof of concept. *Phys. Chem. Chem. Phys.* **2020**, *22*, 5024−5032.
(10) Gordy, W.; Cook, R. L. *Microwave Molecular Spectra*; Wiley, 1984.
(11) Puzzarini, C.; Heckert, M.; Gauss, J. The accuracy of rotational constants predicted by high-level quantum-chemical calculations. I. Molecules containing first-row atoms. *J. Chem. Phys.* **2008**, *128*, 194108.
(12) Caminati, W.; Grabow, J.-U. In *Frontiers of Molecular Spectroscopy*; Laane, J., Ed.; Elsevier: Amsterdam, 2009; pp 455−552.
(13) Puzzarini, C. Rotational spectroscopy meets theory. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6595−6607.
(14) Puzzarini, C. Accurate molecular structures of small-and medium-sized molecules. *Int. J. Quantum Chem.* **2016**, *116*, 1513−1519.
(15) Pate, B. H.; Evangelisti, L.; Caminati, W.; Xu, Y.; Thomas, J.; Patterson, D.; Perez, C.; Schnell, M. In *Chiral Analysis*, 2nd ed.; Polavarapu, P. L., Ed.; Elsevier, 2018; pp 679−729.
(16) Pulay, P.; Meyer, W.; Boggs, J. E. Cubic force constants and equilibrium geometry of methane from Hartree−Fock and correlated wavefunctions. *J. Chem. Phys.* **1978**, *68*, 5077−5085.
(17) Penocchio, E.; Piccardo, M.; Barone, V. Semiexperimental equilibrium structures for building blocks of organic and biological molecules: The B2PLYP route. *J. Chem. Theory Comput.* **2015**, *11*, 4689−4707.
(18) Piccardo, M.; Penocchio, E.; Puzzarini, C.; Biczysko, M.; Barone, V. Semi-experimental equilibrium structure determinations by employing B3LYP/SNSD anharmonic force fields: Validation and application to semirigid organic molecules. *J. Phys. Chem. A* **2015**, *119*, 2058−2082.
(19) Puzzarini, C.; Barone, V. Diving for accurate structures in the ocean of molecular systems with the help of spectroscopy and quantum chemistry. *Acc. Chem. Res.* **2018**, *51*, 548−556.
(20) Alessandrini, S.; Barone, V.; Puzzarini, C. Extension of the "cheap" composite approach to noncovalent interactions: The jun-ChS scheme. *J. Chem. Theory Comput.* **2020**, *16*, 988−1006.
(21) Lei, J.; Alessandrini, S.; Chen, J.; Zheng, Y.; Spada, L.; Gou, Q.; Puzzarini, C.; Barone, V. Rotational spectroscopy meets quantum chemistry for analyzing substituent effects on non-covalent Interactions: The case of the trifluoroacetophenone-water complex. *Molecules* **2020**, *25*, 4899.
(22) Lupi, J.; Puzzarini, C.; Cavallotti, C.; Barone, V. State-of-the-art quantum chemistry meets variable reaction coordinate transition state theory to solve the puzzling case of the H_2S + Cl system. *J. Chem. Theory Comput.* **2020**, *16*, 5090−5104.
(23) Santra, G.; Sylvetsky, N.; Martin, J. M. Minimally empirical double-hybrid functionals trained against the GMTKN55 database: revDSD-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. *J. Phys. Chem. A* **2019**, *123*, 5129−5143.
(24) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(25) Papajak, E.; Truhlar, D. G. Convergent partially augmented basis sets for post-Hartree-Fock calculations of molecular properties and reaction barrier heights. *J. Chem. Theory Comput.* **2011**, *7*, 10−18.
(26) Gougoula, E.; Cole, D. J.; Walker, N. R. Bifunctional hydrogen bonding of imidazole with water explored by rotational spectroscopy and DFT calculations. *J. Phys. Chem. A* **2020**, *124*, 2649−2659.
(27) Zischang, J.; Lee, J. J.; Suhm, M. A. Communication: Where does the first water molecule go in imidazole? *J. Chem. Phys.* **2011**, *135*, 061102.
(28) Mills, I. M. In *Molecular Spectroscopy: Modern Research*; Rao, K. N., Matthews, C. W., Eds.; Academic Press, 1972; *Vol. 1*, pp 115−140.
(29) Pawłowski, F.; Jørgensen, P.; Olsen, J.; Hegelund, F.; Helgaker, T.; Gauss, J.; Bak, K. L.; Stanton, J. F. Molecular equilibrium structures from experimental rotational constants and calculated vibration−rotation interaction constants. *J. Chem. Phys.* **2002**, *116*, 6482−6496.
(30) Puzzarini, C.; Stanton, J. F.; Gauss, J. Quantum-chemical calculation of spectroscopic parameters for rotational spectroscopy. *Int. Rev. Phys. Chem.* **2010**, *29*, 273−367.
(31) Liévin, J.; Demaison, J.; Herman, M.; Fayt, A.; Puzzarini, C. Comparison of the experimental, semi-experimental and ab initio equilibrium structures of acetylene: Influence of relativisitic effects and of the diagonal Born−Oppenheimer corrections. *J. Chem. Phys.* **2011**, *134*, 064119.
(32) Puzzarini, C.; Bloino, J.; Tasinato, N.; Barone, V. Accuracy and interpretability: The devil and the holy grail. New routes across old boundaries in computational spectroscopy. *Chem. Rev.* **2019**, *119*, 8131−8191.
(33) Barone, V.; Biczysko, M.; Puzzarini, C. Quantum chemistry meets spectroscopy for astrochemistry: Increasing complexity toward prebiotic molecules. *Acc. Chem. Res.* **2015**, *48*, 1413−1422.
(34) Penocchio, E.; Mendolicchio, M.; Tasinato, N.; Barone, V. Structural features of the carbon−sulfur chemical bond: A semi-experimental perspective. *Can. J. Chem.* **2016**, *94*, 1065−1076.
(35) Herbers, S.; Kraus, P.; Grabow, J.-U. Accurate equilibrium structures of methyl methacrylate and methacrylic acid by microwave spectroscopy and dispersion corrected calculations. *J. Chem. Phys.* **2019**, *150*, 144308.
(36) Xie, F.; Fusè, M.; Hazrah, S. A.; Jaeger, W.; Barone, V.; Xu, Y. Discovering the elusive global minimum in a ternary chiral cluster: Rotational spectra of propylene oxide trimer. *Angew. Chem., Int. Ed.* **2020**, *59*, 22427−22430.
(37) Demaison, J.; Vogt, N.; Ksebafontov, N. D. Accuracy of semiexperimental equilibrium structures: Sulfine as an example. *J. Mol. Struct.* **2020**, *1206*, 127676.
(38) Heim, Z. N.; Amberger, B. K.; Esselman, B. J.; Stanton, J. F.; Woods, R. C.; McMahon, R. J. Molecular structure determination: Equilibrium structure of pyrimidine (m-$C_4H_4N_2$) from rotational spectroscopy (r_e^{SE}) and high-level ab initio calculation (r_e) agree within the uncertainty of experimental measurement. *J. Chem. Phys.* **2020**, *152*, 104303.
(39) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098.
(40) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785.
(41) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
(42) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456−1465.

https://doi.org/10.1021/acs.jpca.1c01679
J. Phys. Chem. A 2021, 125, 2989−2998

(43) Barone, V.; Ceselin, G.; Fusè, M.; Tasinato, N. Accuracy meets interpretability for computational spectroscopy by means of hybrid and double-hybrid functionals. *Front. Chem.* **2020**, *8*, 584203.
(44) Puzzarini, C.; Barone, V. Extending the molecular size in accurate quantum-chemical calculations: The equilibrium structure and spectroscopic properties of uracil. *Phys. Chem. Chem. Phys.* **2011**, *13*, 7189−7197.
(45) Puzzarini, C.; Biczysko, M.; Barone, V.; Pena, I.; Cabezas, C.; Alonso, J. L. Accurate molecular structure and spectroscopic properties of nucleobases: A combined computational−microwave investigation of 2-thiouracil as a case study. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16965−16975.
(46) Puzzarini, C.; Biczysko, M. Microsolvation of 2-thiouracil: Molecular structure and spectroscopic parameters of the thiouracil−water complex. *J. Phys. Chem. A* **2015**, *119*, 5386−5395.
(47) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479−483.
(48) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.* **1997**, *106*, 9639−9646.
(49) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618.
(50) Woon, D. E.; Dunning, T. H., Jr Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon. *J. Chem. Phys.* **1995**, *103*, 4572−4585.
(51) Kendall, R. A.; Dunning, T. H., Jr; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796−6806.
(52) 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.; et al. *Gaussian16*, Revision C.01; Gaussian Inc.: Wallingford, CT, 2019; see http://gaussian.com.
(53) Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G. *CFOUR, Coupled-Cluster techiques for Computational Chemistry*, version 1.2; 2008; see http://cfour.de.
(54) Arunan, E.; Desiraju, G. R.; Klein, R. A.; Sadlej, J.; Scheiner, S.; Alkorta, I.; Clary, D. C.; Crabtree, R. H.; Dannenberg, J. J.; Hobza, P.; et al. Definition of the hydrogen bond (IUPAC Recommendations 2011). *Pure Appl. Chem.* **2011**, *83*, 1637−1641.
(55) Puzzarini, C. Extrapolation to the complete basis set limit of structural parameters: Comparison of different approaches. *J. Phys. Chem. A* **2009**, *113*, 14530−14535.
(56) Mendolicchio, M.; Penocchio, E.; Licari, D.; Tasinato, N.; Barone, V. Development and implementation of advanced fitting methods for the calculation of accurate molecular structures. *J. Chem. Theory Comput.* **2017**, *13*, 3060−3075.
(57) Biczysko, M.; Bloino, J.; Puzzarini, C. Computational challenges in Astrochemistry. *WIRES: Comp. Mol. Sci.* **2018**, *8*, No. e1349.
(58) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553−566.
(59) Glendening, E. D. Natural energy decomposition analysis: Extension to density functional methods and analysis of cooperative effects in water clusters. *J. Phys. Chem. A* **2005**, *109*, 11936−11940.
(60) Glendening, E. D.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Karafiloglou, P.; Landis, C. R.; Weinhold, F. *NBO 7.0, Natural Bond Orbital Analysis Programs*, version 7.0; Theoretical Chemistry Institute, University of Wisconsin: Madison, WI, 2018; see https://nbo7.chem.wisc.edu.

https://doi.org/10.1021/acs.jpca.1c01679
J. Phys. Chem. A 2021, 125, 2989−2998

B.2 Water–amine molecular complexes

In order to gain a better understanding of the nature of weak and non-covalent interactions, in Ref. 311, the n-propylamine and iso-propylamine adducts with water have been accurately characterized by means of a joint experimental-theoretical approach, integrating rotational spectroscopy with energy decomposition analyses. From the experimental rotational constants for different isotopic species (obtained in Ref. 311 from the recording and analysis of the corresponding rotational spectra) and the computation of the corresponding vibrational corrections, the determination of the SE equilibrium structures for the intermolecular parameters has been carried out. In addition, the effect of the length and ramification of the alkyl chain on the water–amine interactions has been addressed, together with the effect of the latter on the stability order of propylamine conformers with respect to the isolated systems described in Chapter 13.

PCCP

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2020, **22**, 5024

Theory meets experiment for elucidating the structure and stability of non-covalent complexes: water–amine interaction as a proof of concept†

Junhua Chen,‡[a] Yang Zheng,‡[a] Alessio Melli,‡[b] Lorenzo Spada,[bc] Tao Lu,[a] Gang Feng,[ad] Qian Gou,*[ad] Vincenzo Barone*[c] and Cristina Puzzarini*[b]

Received 16th December 2019,
Accepted 11th February 2020

DOI: 10.1039/c9cp06768j

rsc.li/pccp

Several gas-phase spectroscopic investigations have focused on a better understanding of the nature of weak, non-covalent interactions in model systems. However, their characterization and interpretation are still far from being satisfactory. A promising route to fill this gap is offered by strategies in which high-resolution rotational spectroscopy is deeply integrated with state-of-the-art quantum-chemical methodology to accurately determine intermolecular parameters and interaction energies, with the latter interpreted by means of powerful energy decomposition analyses (EDAs). As a proof of concept of this approach, we have selected the adducts formed by *n*-propylamine (PA) and iso-propylamine (IPA) with water. Among the stable structures computationally predicted, four (out of five) isomers of the PA–water complex and two isomers (*trans* and *gauche*) of the IPA–water adduct have been characterized with supersonic jet Fourier transform microwave spectroscopy. Starting from the experimental rotational constants for different isotopic species, computation of the corresponding vibrational corrections allowed a semi-experimental determination of the intermolecular parameters. Different EDAs point out that in all cases a strong O–H···N hydrogen bond is the primary interaction. Accurate computations indicate that the length and ramification of the alkyl chain do not significantly affect the water–amine interactions, which – on the contrary – modify the stability order of PA conformers with respect to the isolated systems.

Introduction

Non-covalent molecular complexes, in addition to their intrinsic interest, provide the unique opportunity to gain insights on the interactions of a given molecule with the environment and, in particular, in condensed phases. Indeed, the first step toward their understanding is a complete characterization of the basic interactions occurring between one target molecule and one "environmental" species, with water being that of greatest interest. As a matter of fact, gas-phase investigations of non-covalent molecular complexes provide structural and dynamical information that is seldom straightforward to be derived from measurements in condensed phases. To exploit such investigations at best, a powerful strategy relies on high-resolution rotational spectroscopy supported and complemented by state-of-the-art quantum-chemical (QC) computations (see, *e.g.*, ref. 1–4). Although a large number of spectroscopic studies of molecular complexes also reporting some QC calculations can be found in the literature (see, *e.g.*, ref. 1 and 5–9), there is still a lack of full integration between experiment and theory, thus preventing a full reconciliation of accuracy and interpretation. Illustrative in this respect is the fact that the usual practice for deriving structural information from rotational spectra of different isotopic species is to rely on questionable, "old-fashioned" approximations such as the so-called substitution structure. On the other hand, the semi-experimental approach,[10] well-tested for isolated molecules (see, *e.g.*, ref. 11–14), allows the determination of not only equilibrium intramolecular, but also intermolecular, parameters with great accuracy (0.001 Å for distances

[a] *Department of Chemistry, School of Chemistry and Chemical Engineering, Chongqing University, Daxuecheng South Rd. 55, 401331 Chongqing, China. E-mail: qian.gou@cqu.edu.cn*
[b] *Department of Chemistry "Giacomo Ciamician", University of Bologna, Via F. Selmi 2, 40126 Bologna, Italy. E-mail: cristina.puzzarini@unibo.it*
[c] *Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy. E-mail: vincenzo.barone@sns.it*
[d] *Chongqing Key Laboratory of Theoretical and Computational Chemistry, Chongqing University, Daxuecheng South Rd. 55, 401331 Chongqing, China*

† Electronic supplementary information (ESI) available: PDF file collecting all transitions recorded together with the corresponding spectroscopic parameters from the fit. Detailed results from the SAPT and NBO analyses are also reported. See DOI: 10.1039/c9cp06768j

‡ These authors contributed equally to this work.

and 0.1 degrees for angles) by combining experimental rotational constants with computed vibrational corrections. A further aspect of this strong interplay of experiment and theory is the derivation of interaction energies with uncertainties as low as a fraction of kJ mol^{-1} combined with their interpretation in terms of chemically meaningful concepts by means of energy decomposition analyses (EDAs).

As a proof of concept of this approach, we have chosen the interaction between the primary amino group ($-NH_2$) and water, which plays a key role in several fields ranging from biology to atmospheric chemistry and astrochemistry. Despite the fact that several spectroscopic studies of water adducts with ammonia and amines have been previously reported, the number of investigations involving primary amines is quite limited (see, *e.g.*, ref. 15–20). Hydrogen bonds (HBs) are ubiquitous in organic or biological molecules involving either oxygen or nitrogen, their topologies and strengths being significantly tuned by the functional groups bonded to the interacting atoms. In this respect, the tuning of amine–water interaction by the length and/or ramification of the alkyl chain has not been fully characterized yet and any generalization from small to large alkyl amines is completely missing.

Alkyl amines have been detected systematically in the atmosphere[21–25] and represent one important class of compounds involved in the complex organic fraction of ambient aerosol and play a pivotal role in aerosol growth processes (see, *e.g.*, ref. 21–23) and in determining the hygroscopic properties of particles.[24] Although intermolecular complexes are not of direct astrophysical interest, ices, hydrates, and clathrates are formed from ammonia and other amines. The coexistence of a number of hydrate–clathrate systems[26,27] might determine, in part, the volatile inventories of comets, icy satellites, and other objects.[28,29] Last but not least, the fine tuning of amine properties has significant implications for their increasing employment in nano-engineering.[30,31] These examples point out that a deeper understanding of how amines interact with water warrants a thorough investigation. As a matter of fact, the development of amine–water intermolecular parameters is an area of active development,[32,33] which – however – suffers from the lack of accurate structures and interaction energies for reference amine–water complexes.

The above considerations prompted us to undertake a comprehensive study of the molecular complexes formed by *n*-propylamine (PA) and iso-propylamine (IPA) with one water molecule (W) by means of an integrated computational–experimental approach, which has already provided remarkable results in some recent studies.[3,4,34] PA and IPA have been selected because they represent a good compromise: on the one hand, they are systems with an alkyl chain sufficiently long to plausibly lead to significant effects; on the other hand, the sizes of their water complexes are such that they are still amenable to high-level QC calculations. To complement such investigation, the molecular complexes involving smaller primary amines, namely methylamine (MA) and ethylamine (both *trans*, ET, and *gauche*, EG), with one water molecule have also been characterized from a computational point of view in order to unveil general trends related to the lengthening of the alkyl chain.

Computational methods

The conformational potential energy surfaces (PESs) of both the PA–W and IPA–W adducts have been investigated using the B2PLYP-D3(BJ) double-hybrid functional[35] (with D3(BJ) denoting the correction for dispersion effects according to the Grimme's DFT-D3[36] scheme employing the Becke–Johnson damping function[37]) in conjunction with the maug-cc-pVTZ-*d*H basis set,[38] where "-*d*H" denotes that *d* functions on hydrogen atoms have been removed.[39] Geometry optimizations have been performed by both including and neglecting the counterpoise correction (CP)[40] to recover the basis set superposition error (BSSE), thus leading to two levels of theory shortly denoted as B2-CP and B2, respectively. While CP corrections are mandatory in order to obtain reliable and accurate interaction energies, their inclusion in geometry optimizations might lead to a potential improvement of the results, as recently demonstrated in ref. 41.

All stationary points have been confirmed and characterized by evaluating the corresponding Hessian matrix (harmonic force field in a normal mode representation) at the B2 level by means of analytical second derivatives.[42] Subsequently, energetics has been improved using the so-called "cheap" composite scheme (hereafter, simply referred to as ChS).[43–45] In this approach, the energy obtained using the coupled-cluster singles and doubles approach with perturbative inclusion of triple excitations, CCSD(T),[46] in conjunction with the cc-pVTZ[47] basis set and within the frozen-core (fc) approximation, is corrected for the contributions due to the extrapolation to the complete basis set (CBS) limit and to the core-valence (CV) correlation effects, evaluated using second-order Møller–Plesset theory (MP2):[48]

$$E(\mathrm{ChS}) = E(\text{fc-CCSD(T)/cc-pVTZ}) + \Delta E(\mathrm{CBS}) + \Delta E(\mathrm{CV}) \quad (1)$$

(i) ΔE(CBS) accounts for the extrapolation to the CBS limit; this term is computed by means of the two-point n^{-3} extrapolation formula of ref. 49, with n = 3 (cc-pVTZ), 4 (cc-pVQZ). (ii) ΔE(CV) recovers the CV contribution. This term is calculated as the difference between all-electrons and fc-MP2 calculations, both with the same cc-pCVTZ basis set.[50]

Electronic energies need to be further corrected for the BSSE error using the CP method, thus leading to the final ChS-CP energies. Zero-point energy (ZPE) corrections have been incorporated within the harmonic approximation at the B2 level (ChS-CP-ZPE).

Analogously, the equilibrium structures of methylamine–water (MAW) and of the two isomers of ethylamine–water complexes (ETW and EGW) have been optimized at the B2 and B2-CP level, and their interaction energy at the B2-CP level.

To spectroscopically characterize all PA–/IPA–W isomers, anharmonic force field calculations have been performed at the B3LYP-D3(BJ)/SNSD level[51,52] (hereafter denoted as B3). Within vibrational perturbation theory to second order,[53,54] these allow for the derivation of the vibrational corrections to rotational constants ($\Delta B_{\mathrm{vib}}^{i}$), which have been used to correct

View Article Online
PCCP Paper

the equilibrium rotational constants (B_e^i) straightforwardly derived from the equilibrium structure evaluations:[53]

$$B_0^i = B_e^i(\text{B2-CP}) + \Delta B_{vib}^i(\text{B3}) = B_e^i - \frac{1}{2}\sum_r \alpha_r^i \qquad (2)$$

where the α_r^i's are the vibration–rotation interaction constants, i refers to the inertial axes (a, b and c), and the sum runs over the r vibrational normal modes. B_0^i's denote the vibrational ground-state rotational constants. At the same time, the B2 harmonic force field provides, as a byproduct, predictions for quartic centrifugal distortion constants.

Because of the presence of a nitrogen atom (nuclear spin $I = 1$), the corresponding nuclear quadrupole coupling constants represent additional important spectroscopic parameters. To evaluate the corresponding tensor, the components of the electric field gradient tensor at the nitrogen nucleus have been computed at the B2 level. From them, the components of the nitrogen quadrupole coupling constants have then been obtained by means of the following expression:

$$\chi_{ij}\,(\text{MHz}) = 234.9647 \times Q\,(\text{barn}) \times q_{ij}\,(\text{a.u.}) \qquad (3)$$

where the first term on the right-hand side is a conversion factor. Q (barn) is the ^{14}N quadrupole moment (20.44(3) barn[55]) and q_{ij}'s (a.u.) are the components of the electric field gradient tensor, with i, j denoting the inertial axes. Vibrational corrections, at the B3 level, have also been incorporated.

Dipole moment components along the inertial axes, which in turn determine the type and the intensity of transitions observable in the rotational spectra, have been evaluated at the B2 level.

All computations have been performed using the Gaussian 16[56] program package.

Experimental section

Samples of PA (≥99%) and IPA (≥99%), both purchased from Adamas, were used without any further purification. A gas mixture of 1% PA (or IPA) in helium at a stagnation pressure of 0.2 MPa was streamed over water (main isotopologue, with $H_2{}^{18}O$ and HOD also considered), and supersonically expanded through a solenoid valve (Parker-General Valve, Series 9, nozzle diameter 0.5 mm) into the Fabry–Pérot cavity of the spectrometer.

Rotational spectra of the PA–W and IPA–W adducts were recorded in the 2.0–20.0 GHz frequency range using a highly integrated supersonic-jet Fourier-Transform Microwave (FTMW) spectrometer,[57] with a coaxially oriented beam-resonator arrangement (COBRA-type[58]), which has been described in detail elsewhere.[59] The spectral line positions were determined after Fourier transformation of the time-domain signal with 8k data points, recorded with 100 ns sample intervals. Each transition appears split into a Doppler doublet due to the coaxial resonator arrangement of the supersonic jet, the line position being obtained as the arithmetic mean of the frequencies of the two Doppler components. The estimated accuracy for frequency measurements is better than 3 kHz and the resolution better than 6 kHz.

Results and discussion

All possible isomers of the PA-/IPA–W systems obtained from the computational search are depicted in Fig. 1. Two stable isomers have been identified for the IPA–W adduct, where IPA adopts *trans* (TW) or *gauche* (GW) conformation. As far as PA–W is concerned, five stable isomers have been located on the PES, which are labelled as TGW, GG′W, TTW, GTW and GGW, where the first two letters denote the PA conformer involved in the molecular complex with water. The notation used for the PA conformers is based on the τ(CCCN) dihedral angle of the skeletal frame and the orientation of the lone pair in amino group (τ(CCN-lp)). The relative stability of the minima together with that of the transition states ruling their interconversion are shown in Fig. 2 (left panel for PA–W, right panel for IPA–W). As already noted for *n*-propanol,[60–62] the water cluster formation modifies the relative stability of the different isomers with respect to the corresponding isolated amine. In the case of PA, the complexation with water leads – in relative terms – to a destabilization of the TT and GG conformers and to a stabilization of TG, GT and GG′ (see Table 1).

Computed spectroscopic parameters and relative energies of all PA–/IPA–W isomers are collected in Table 1, with a more extended version of this table reported in the ESI.† Spectral searches using simulations based on the spectroscopic parameters reported in Table 1 have been performed following the stability order. For PA–W, TGW was firstly observed and assigned. After its frequencies were removed from the spectrum, rotational transitions of the GG'W, TTW and GTW isomers have been then easily assigned thanks to their characteristic ^{14}N quadrupole hyperfine patterns (see Fig. 3). In agreement with its relative stability (and thus its low population), no lines belonging to GGW have been found. The experimental outcomes are in line with the computed stabilities and the barrier heights ruling the interconversion (Fig. 2). For the IPA–W adduct, the assignment of the rotational spectra of both isomers was straightforward thanks again to the characteristic hyperfine structure due to the nitrogen quadrupole coupling (see Fig. 3).

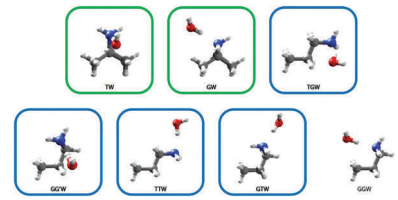


Fig. 1 Stable isomers of the IPA–W and PA–W adducts. The isomers within the frames (green for IPA–W, light blue for PA–W) have been experimentally observed.

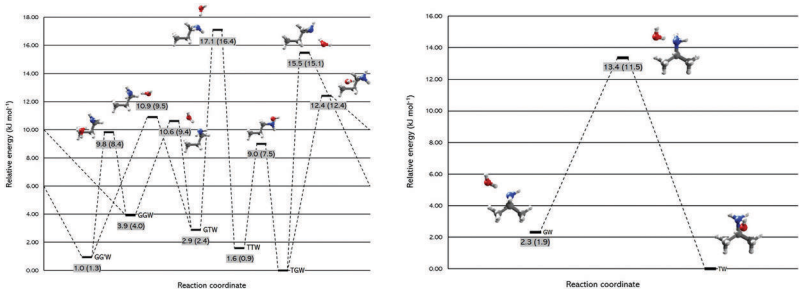


Fig. 2 Isomerization profile for the PA–W (left panel) and IPA–W (right panel) adducts: ChS-CP and ChS-CP-ZPE (within parentheses) relative energies in kJ mol^{-1}.

Table 1 Spectroscopic parameters and relative energies of the PA–W and IPA–W adducts[a]

		PA–W					IPA–W	
		TGW	GG'W	TTW	GTW	GGW	TW	GW
A_0/MHz	Theo.	5378.9	4963.8	9895.1	11290.1	4818.3	4805.5	7239.2
	Exp.	5466.2596(9)[b]	5003.9028(9)	9977.1499(8)	11477.9594(7)	—	4880.478(2)	7271.1043(1)
B_0/MHz	Theo.	2180.5	2450.3	1452.8	1471.6	2608.0	2520.2	2101.3
	Exp.	2172.6407(3)	2411.8257(3)	1467.3249(3)	1478.4918(2)	—	2496.0733(8)	2102.8262(3)
C_0/MHz	Theo.	1686.2	1995.4	1307.6	1418.1	1835.9	2425.5	1799.2
	Exp.	1688.9887(2)	1979.0333(3)	1319.2830(2)	1427.6159(2)	—	2401.5471(7)	1805.9597(3)
χ_{aa}/MHz	Theo.	−0.35	−1.61	−3.27	−4.10	−0.98	−2.88	−3.95
	Exp.	−0.400(3)	−1.626(2)	−3.195(5)	−3.963(5)	—	−2.730(3)	−3.825(8)
$(\chi_{bb} - \chi_{cc})$/MHz	Theo.	−3.52	−2.74	−0.31	0.60	−3.12	0.75	−0.10
	Exp.	−3.232(4)	−2.387(3)	−0.198(6)	0.627(6)	—	0.731(7)	−0.131(9)
ΔE/kJ mol^{-1}		0[c]	0.96	1.59	2.87	3.92	0	2.31
		(0.16)[c]	(0.88)	(0)	(1.26)	(1.15)	(0)	(2.02)
ΔE_0/kJ mol^{-1}		0	1.35	0.94	2.45	4.00	0	1.91

[a] Ground-state rotational constants (A_0, B_0, C_0) from B2-CP equilibrium rotational constants augmented by vibrational corrections at the B3 level (for details, see the ESI). Equilibrium nitrogen quadrupole coupling constants (χ_{ii}, with $i = a, b, c$) at the B2 level augmented by vibrational corrections at the B3 level (equilibrium values are provided in the ESI). Equilibrium (ΔE) and ZPE-corrected (ΔE_0) relative energies at the ChS-CP and ChS-CP-ZPE level, respectively. [b] Standard errors within parentheses are expressed in units of the last digit. [c] Values in parentheses are the relative energies of the isolated amine conformers at the ChS level.

For all molecular complexes, the measured transition frequencies have been fitted, using Pickett's SPFIT program,[63] to Watson's semirigid-rotor Hamiltonian[64] (I^r representation, S-reduction). Selected spectroscopic parameters are listed in Table 1 for the main isotopic species, while all measured transition frequencies as well as the complete list of the spectroscopic constants are available in the ESI.† In all cases, the standard deviation of the fit is small and on the order of the measurement accuracy (*i.e.* 2–4 kHz). In addition to the main isotopic species, water isotopologues were also investigated. As far as the PA–W isomers are concerned, the rotational spectra of the molecular complexes with ^{18}O-enriched water and DOH (with D involved in the HB) have been assigned. For the IPA–W isomers, transitions belonging to the IPA-$H_2{}^{18}$O/DOH/HOD isotopologoues could be successfully identified. The spectroscopic parameters for all isotopic species are summarized in the ESI.†

The comparison between experimental and computed spectroscopic parameters points out an overall good agreement. For rotational constants, the differences – in relative terms – range between 0.1% and 1.6%, which can be considered a very good result in view of the high flexibility of the systems under consideration. Analogously, the nuclear quadrupole coupling constants are well predicted. The good agreement noted for these two sets of spectroscopic parameters, which are strongly related to the molecular structure, suggests that the B2(-CP) level is suitable for describing flexible intermolecular complexes.

For each observed isomer, the investigation of different isotopologues has opened the way to the determination of a partial semi-experimental (SE) equilibrium structure (r_e^{SE}). This approach is based on a least-squares fit of the experimental ground-state rotational constants computationally corrected for vibrational contributions.[10]

View Article Online
PCCP Paper

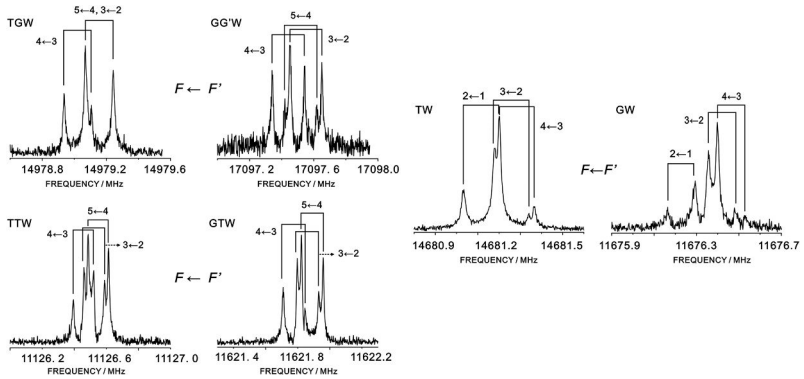


Fig. 3 Left panel: The $4_{0,4} \leftarrow 3_{0,3}$ rotational transition for the TGW, GG′W, TTW and GTW isomers. Right panel: The $3_{0,3} \leftarrow 2_{0,2}$ rotational transition of the TW and GW isomers. All spectra show the characteristic hyperfine structure and Doppler doubling. Hyperfine components are labelled using the F quantum number coming from the $F = I + J$ coupling scheme, with I being the nuclear spin of ^{14}N and J the rotational quantum number.

As shown in eqn (2), equilibrium rotational constants B_e^i's differ from the vibrational ground-state ones B_0^i's – derived from the spectral analysis – for a contribution that can be reliably computed $\left(\Delta B_{vib}^i = -1/2 \sum_r \alpha_r^i\right)$.[11,65] The resulting equilibrium rotational constants are thus defined as semi-experimental because of the mixed contributions. Although the number of isotopologues investigated in the present work is not sufficient for a complete structural characterization, they anyway allow the determination of the intermolecular parameters with an experimental-quality accuracy.[11,65,66]

The intramolecular bond lengths and angles have been kept fixed at their scaled values, which were obtained by means of the template approach.[11] While we refer the reader to ref. 11 for a detailed account, in the following the key aspects of this methodology are summarized. The structural parameters that cannot be determined in the fit are corrected by using high-level calculations performed for a smaller system, referred to as templating model (TM), which contains a similar structural frame:

$$r_e = r_e^{B2} + \Delta TM. \quad (4)$$

r_e^{B2} is the geometrical parameter obtained at the B2 level for the system under investigation, while ΔTM is the difference between the B2 equilibrium parameter and the corresponding reference value (in the present case, at the ChS + AUG level) for the TM molecule:

$$\Delta TM = r_e^{ChS+AUG}(TM) - r_e^{B2}(TM). \quad (5)$$

"ChS + AUG" denotes the ChS approach for geometries, which is analogous to the ChS model for energies, but it also includes a contribution for incorporating the effects of diffuse functions in the basis set (Δr(aug)):[43]

$$r(ChS + AUG) = r(CC/VTZ) + \Delta r(CBS) + \Delta r(core) + \Delta r(aug) \quad (6)$$

where Δr(aug) is calculated as the difference between MP2 calculations with the aug-cc-pVTZ[67] and cc-pVTZ basis sets, within the frozen-core approximation.

In eqn (4) and (5), TMs are the two monomers, the PA or IPA isolated isomer and the water molecule. The MSR software[68] has been used for the structural determinations above, with all rotational constants being equally weighted.

The SE intermolecular parameters obtained from the fit are collected in Table 2 and compared with their QC counterparts. It is noted that the two sets of values are in a reasonably good agreement, with the largest discrepancies noted for the ∠HON angle (with H being the hydrogen atom involved in the HB), which is – however – the worse determined parameter. It is noteworthy that exclusion of the ∠HON angle from the fit affects only negligibly the final values of the other two parameters.

Comparison of the computed intermolecular parameters of the PA–W and IPA–W isomers with those of MAW, ETW and EGW shows that the O···N intermolecular distance remains nearly unchanged along the series, this ranging from 2.87 Å to 2.89 Å at the B2 level and from 2.89 Å to 2.90 Å at the B2-CP level. This comparison suggests that the length of the alkyl chain and also its ramification do not have a relevant effect on the intermolecular interactions.

To shed light on the intermolecular interactions taking place in the molecular complexes investigated, different analyses have been carried out. The first one considered is the Johnson's non-covalent interaction (NCI) analysis,[69,70] which has been

Table 2 SE[a] (first row) and B2-CP (second row; B2 given in parentheses underneath) intermolecular parameters in the alkyl amine–water complexes

	MAW	ETW	EGW	TW	GW	TGW	GG′W	TTW	GTW
$R_{O\cdots N}$/Å	—	—	—	2.9135(4)	2.894(1)	2.88134(3)	2.895(1)	2.88586(6)	2.8830(6)
	2.890	2.893	2.890	2.903	2.887	2.888	2.897	2.893	2.892
	(2.877)	(2.879)	(2.876)	(2.889)	(2.872)	(2.873)	(2.882)	(2.879)	(2.879)
∠ONC/°	—	—	—	102.35(2)	98.61(6)	97.9754(2)	101.34(3)	96.6001(8)	97.44(4)
	96.06	96.69	97.96	101.29	98.66	97.58	102.06	96.90	96.69
	(95.76)	(96.37)	(97.33)	(100.75)	(98.05)	(96.95)	(100.40)	(96.53)	(96.29)
∠HON[b]/°	—	—	—	3.08(28)	5.78(83)	7.63(2)	4.48(93)	8.06(6)	7.39(13)
	9.22	8.96	9.31	8.54	9.18	9.42	7.87	8.80	8.76
	(9.33)	(9.08)	(9.61)	(8.74)	(9.47)	(9.74)	(8.33)	(8.91)	(8.88)
∠OHN[b]/°	—	—	—	175.36[c]	171.30[c]	168.50[c]	173.24[c]	167.86[c]	168.86[c]
	166.14	166.52	166.00	167.17	166.17	165.82	168.16	166.77	166.82
	(165.94)	(166.31)	(165.51)	(166.83)	(165.70)	(165.30)	(167.44)	(166.56)	(166.61)

[a] The uncertainty reported is one standard deviation of the fit. [b] The H atom is the one involved in the HB. [c] Derived parameter.

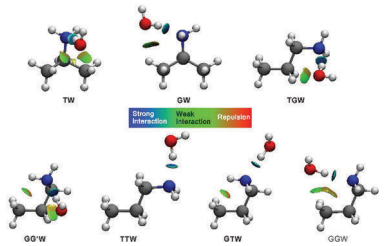


Fig. 4 Color-filled (color scale given in the inset) NCI isosurfaces of the reduced electron density gradient (s = 0.5). Isomers in bold have been experimentally observed.

performed using the Multiwfn program.[71] The results are graphically shown in Fig. 4, where the iso-surfaces of the reduced density gradient are shown together with a color legend box for the interaction strength. It is first of all noted that all molecular complexes present one strong O–H···N HB as primary interaction, which is – in the case of the TGW, GGW and GW isomers – associated with one weaker C–H···O HB. For the TW isomer, two weak C–H···O HBs are present together with the primary interaction, while in the case of the TTW isomer no additional interactions are found. The GTW complex shows, instead of an additional intermolecular HB, a weak intramolecular N–H···C interaction. Finally, the terminal methyl group of the GG′W isomer is able to form both intramolecular (N–H···C) and intermolecular (C–H···O) weak HBs.

To go deeper into the analysis of non-covalent interactions and to unravel their nature, the symmetry-adapted perturbation theory (SAPT)[72,73] EDA has been carried out at the SAPT2+3(CCD)/aug-cc-pVDZ-RI level using the PSI4 program.[74] The results are sketched in Fig. 5 and tabulated in the ESI† (Table S25), where the values for the water dimer as well as for MAW, ETW and EGW have been reported for comparison purposes. Focusing on PA–/IPA–W complexes, the interaction energies range between 29 and 32 kJ mol^{-1} for all isomers (between 31.2 and 33.5 kJ mol^{-1} and between 31.5 and 34.1 kJ mol^{-1} at the B2-CP and ChS level, respectively; see Table S25, ESI†) and also the individual contributions are similar. Inspection of the various contributions (see Table S25, ESI† and Fig. 5) makes it clear that the largest contribution is the repulsive exchange term, which is nearly compensated by the electrostatic contribution. The induction and dispersion terms, which are rather similar to one another, amount together to about two-thirds of the electrostatic contribution. For MAW as well as ETW and EGW, the situation is very similar both in terms of the single contributions and the overall interaction energy (ranging from ~29 to 31 kJ mol^{-1}). It is interesting to note that the interaction energies of all these complexes are about 50% larger than that of the water dimer, thus indicating that the primary interaction (O–H···N) occurring in the alkyl amine–water complexes is stronger than the O–H···O HB in the water dimer, as expected in view of the greater basicity of amines with respect to water. Overall, the data suggest that the lengthening of the alkyl chain and its ramification have a limited influence on the intermolecular interaction, its strength only marginally increasing along the series. This outcome confirms the overall conclusion drawn from the analysis of the intermolecular parameters.

The SAPT results point out that the interaction between PA and water is stronger in the case of the TG and GG′ conformers with respect to TT, GT and GG by about 2 kJ mol^{-1}. On the contrary, the strength of the interaction is nearly the same for the two IPA conformers. This explains the above-mentioned change of the stability order for the PA conformers when going from the isolate amine to the corresponding water complex, and also why the relative stability remains unchanged for the two IPA conformers.

The results of the natural bond order (NBO) analysis[75] (performed at the B3LYP-D3(BJ)/maug-cc-pVTZ-dH level of theory and detailed in the ESI†) point out that the strength of the O–H···N HB, *i.e.* the interaction between the nitrogen lone pair and the antibonding O–H orbital, is very similar for all the complexes, with the second-order perturbation contribution to the complex stabilization ranging between 12.5 and 13.5 kJ mol^{-1}.

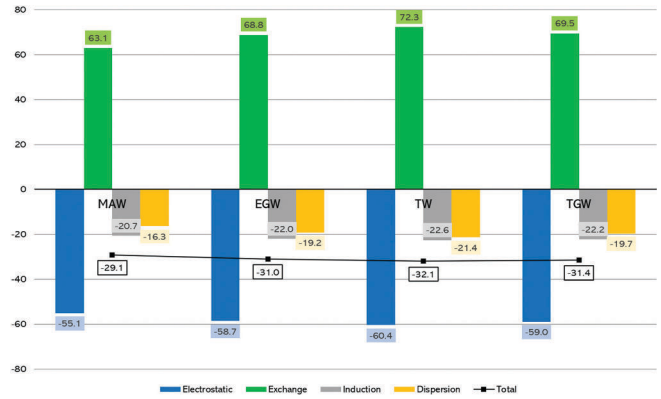


Fig. 5 Summary of the SAPT energy contributions (in kJ mol^{-1}) for the methylamine–water (MAW), ethylamine–water (EGW), iso-propylamine–water (TW), and *n*-propylamine–water (TGW) complexes. When applies, the most stable isomer is considered.

This contribution is accompanied, except for TTW and GTW, by weaker interactions between the oxygen lone-pair(s) and CH antibonding orbital(s).

Finally, starting from the interaction energy (E_{int}), the inclusion of the deformation contribution (Δ_{def}) leads to the evaluation of the dissociation energy ($D_e = E_{int} - \Delta_{def}$; see *e.g.* ref. 76). For the most stable isomers, *i.e.* TGW and TW, the Δ_{def} is −1.48 and −1.56 kJ mol^{-1}, respectively. Using the data (at the ChS level) reported in Table S25 (ESI†), this results in D_e being −32.03 and −32.56 kJ mol^{-1} for TGW and TW, respectively.

Conclusions

In the present work, state-of-the-art quantum-chemical methodologies have been combined with rotational spectroscopy in supersonic expansion to unveil the intermolecular interactions established in alkyl amines–water complexes. While the joint experimental–computational approach has been applied to the characterization of the PA–W and IPA–W adducts, the theoretical investigation has been extended to smaller amines, here considering the complexes formed by methylamine and ethylamine (both *gauche* and *trans* conformers) with water. The comparison of the intermolecular parameters and interaction energies pointed out that the length of the alkyl chain has a very limited effect. This result thus suggests the transferability of the outcomes of the present work to larger alkyl amines, possibly of biological interest.

For all the PA–/IPA–W isomers experimentally observed, the recording and assignment of the rotational spectra of different isotopic species allowed the semi-experimental determination of the intermolecular parameters, which also pointed out the reliability and accuracy of B2(-CP) calculations for structural investigations. Therefore, this work proves that state-of-the-art QC methods coupled to rotational spectroscopy experiments can be confidently employed to obtain equilibrium structures of unprecedented accuracy for non-covalent complexes in the gas phase.

Different energy decomposition analyses, namely NBO and SAPT, agreed in pointing out that the primary interaction is a strong O–H···N HB: the repulsive exchange contribution is nearly compensated by the electrostatic term, while those due to induction and dispersion amount together to two-thirds of the latter contribution, thus leading to an overall interaction energy ranging from ∼29 to 32 kJ mol^{-1}. This picture has been complemented by the NCI analysis, which allowed for better visualizing the number of non-covalent interactions occurring and their strength.

In summary, together with the remarkable intrinsic interest of the studied systems, this work confirms the effectiveness and reliability of combined quantum chemistry – rotational spectroscopy investigations to unveil the role of different physical–chemical factors in tuning the structural and energetic features of non-covalent molecular complexes.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Support from the Italian MIUR (PRIN 2015, Grant Number 2015F59J3R; PRIN 2017, Grant Number 2017A4XRCA) is acknowledged. This work has also been supported by the University of

Bologna (RFO funds). We are grateful for support from: National Natural Science Foundation of China (21703021 and U1931104); Fundamental and Frontier Research Fund of Chongqing (cstc2017jcyjAX0068 and cstc2018jcyjAX0050); Venture & Innovation Support Program for Chongqing Overseas Returns (cx2018064); Foundation of 100 Young Chongqing University (0220001104428); Fundamental Research Funds for the Central Universities (2018CDQYHG0009). The authors also thank Dr Silvia Alessandrini for useful discussions.

Notes and references

1 C. Pérez, J. L. Neill, M. T. Muckle, D. P. Zaleski, I. Peña, J. C. Lopez, J. L. Alonso and B. H. Pate, *Angew. Chem., Int. Ed.*, 2015, **54**, 979–982.
2 C. Pérez, D. P. Zaleski, N. A. Seifert, B. Temelso, G. C. Shields, Z. Kisiel and B. H. Pate, *Angew. Chem., Int. Ed.*, 2014, **53**, 14368–14372.
3 J. Wang, L. Spada, J. Chen, S. Gao, S. Alessandrini, G. Feng, C. Puzzarini, Q. Gou, J.-U. Grabow and V. Barone, *Angew. Chem., Int. Ed.*, 2019, **58**, 13935–13941.
4 D. A. Obenchain, L. Spada, S. Alessandrini, S. Rampino, S. Herbers, N. Tasinato, M. Mendolicchio, P. Kraus, J. Gauss, C. Puzzarini, J.-U. Grabow and V. Barone, *Angew. Chem., Int. Ed.*, 2018, **57**, 15822–15826.
5 N. W. Ulrich, T. S. Songer, R. A. Peebles, S. A. Peebles, N. A. Seifert, C. Pérez and B. H. Pate, *Phys. Chem. Chem. Phys.*, 2013, **15**, 18148–18154.
6 J. C. López, R. Sánchez, S. Blanco and J. L. Alonso, *Phys. Chem. Chem. Phys.*, 2015, **17**, 2054–2066.
7 C. Calabrese, W. Li, G. Prampolini, L. Evangelisti, I. Uriarte, I. Cacelli, S. Melandri and E. J. Cocinero, *Angew. Chem., Int. Ed.*, 2019, **58**, 8437–8442.
8 W. Li, L. Evangelisti, Q. Gou, W. Caminati and R. Meyer, *Angew. Chem., Int. Ed.*, 2019, **58**, 859–865.
9 G. Karir, N. O. B. Lgttschwager and M. A. Suhm, *Phys. Chem. Chem. Phys.*, 2019, **21**, 7831–7837.
10 P. Pulay, W. Meyer and J. E. Boggs, *J. Chem. Phys.*, 1978, **68**, 5077–5085.
11 M. Piccardo, E. Penocchio, C. Puzzarini, M. Biczysko and V. Barone, *J. Phys. Chem. A*, 2015, **119**, 2058–2082.
12 V. Barone, M. Biczysko, J. Bloino, P. Cimino, E. Penocchio and C. Puzzarini, *J. Chem. Theory Comput.*, 2015, **11**, 4342–4363.
13 M. C. McCarthy, L. Cheng, K. N. Crabtree, O. Martinez, T. L. Nguyen, C. C. Womack and J. F. Stanton, *J. Phys. Chem. Lett.*, 2013, **4**, 4133–4139.
14 O. Martinez, K. N. Crabtree, C. A. Gottlieb, J. F. Stanton and M. C. McCarthy, *Angew. Chem., Int. Ed.*, 2015, **54**, 1808–1811.
15 U. Spoerel and W. Stahl, *J. Mol. Spectrosc.*, 1998, **190**, 278–289.
16 S. Melandri, A. Maris, B. M. Giuliano, L. B. Favero and W. Caminati, *Phys. Chem. Chem. Phys.*, 2010, **12**, 10210–10214.
17 I. Yoon, K. Seo, S. Lee, Y. Lee and B. Kim, *J. Phys. Chem. A*, 2007, **111**, 1800–1807.
18 R. Brause, H. Fricke, M. Gerhards, R. Weinkauf and K. Kleinermanns, *Chem. Phys.*, 2006, **327**, 43–53.
19 N. Mayorkas, S. Cohen, H. Sachs and I. Bar, *RSC Adv.*, 2014, **4**, 58752–58757.
20 E. G. Robertson and J. P. Simons, *Phys. Chem. Chem. Phys.*, 2001, **3**, 1–18.
21 R. D. Hoehn, M. A. Carignano, S. Kais, C. Zhu, J. Zhong, X. C. Zeng, J. S. Francisco and I. Gladich, *J. Chem. Phys.*, 2016, **144**, 214701.
22 P. J. Silva, M. E. Erupe, D. Price, J. Elias, Q. G. J. Malloy, Q. Li, B. Warren and D. R. Cocker, *Environ. Sci. Technol.*, 2008, **42**, 4689–4696.
23 J.-S. Youn, E. Crosbie, L. C. Maudlin, Z. Wang and A. Sorooshian, *Atmos. Environ.*, 2015, **122**, 250–258.
24 A. Lavi, N. Bluvshtein, E. Segre, L. Segev, M. Flores and Y. Rudich, *J. Phys. Chem. C*, 2013, **117**, 22412–22421.
25 C. Qiu and R. Zhang, *Phys. Chem. Chem. Phys.*, 2013, **15**, 5738–5752.
26 G. A. Jeffray, T. H. Jordan and R. K. Mullan, *Science*, 1967, **155**, 689–691.
27 F. H. Herbstein, *Crystalline Molecular Complexes and Compounds, Structures and Principles*, IUCr Monographs on Crystallography 18, Oxford University Press, 2005, vol. 1, p. 389.
28 O. Mousis, J. I. Lunine, S. Picaud and D. Cordier, *Faraday Discuss.*, 2010, **147**, 509–525.
29 K. E. Mandt, O. Mousis, B. Marty, T. Cavalié, W. Harris, P. Hartogh and K. Willacy, *Space Sci. Rev.*, 2015, **197**, 297–342.
30 *Nanoscience and Nanoengineering: Advances and Applications*, ed. A. D. Kelkar, D. J.C. Herr and J. G. Ryan, CRC Press, 2014.
31 B. H. Atak, B. Buyuk, M. Huysal, S. Isik, M. Senel, W. Metzger and G. Cetin, *Carbohydr. Polym.*, 2017, **164**, 200–213.
32 W. L. Jorgensen and J. Tirado-Rives, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 6665–6670.
33 M. M. Lencka, J. J. Kosinski, P. Wang and A. Anderko, *Fluid Phase Equilib.*, 2016, **418**, 160–174.
34 W. Li, L. Spada, N. Tasinato, S. Rampino, L. Evangelisti, A. Gualandi, P. G. Cozzi, S. Melandri, V. Barone and C. Puzzarini, *Angew. Chem., Int. Ed.*, 2018, **57**, 13853–13857.
35 S. Grimme, *J. Chem. Phys.*, 2006, **124**, 034108.
36 S. Grimme, J. Anthony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
37 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.
38 E. Papajak, H. R. Leverentz, J. Zheng and D. G. Truhlar, *J. Chem. Theory Comput.*, 2009, **5**, 1197–1202.
39 T. Fornaro, M. Biczysko, J. Bloino and V. Barone, *Phys. Chem. Chem. Phys.*, 2016, **18**, 8479–8490.
40 S. F. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553–566.
41 P. Salvador, B. Paizs, M. Duran and S. Suhai, *J. Comput. Chem.*, 2001, **22**, 765–786.
42 M. Biczysko, G. Scalmani, J. Bloino and V. Barone, *J. Chem. Theory Comput.*, 2010, **6**, 2115–2125.
43 C. Puzzarini and V. Barone, *Phys. Chem. Chem. Phys.*, 2011, **13**, 7158–7166.
44 C. Puzzarini, M. Biczysko, V. Barone, I. Peña, C. Cabezas and J. L. Alonso, *Phys. Chem. Chem. Phys.*, 2013, **15**, 16965–16975.

45 C. Puzzarini and M. Biczysko, *J. Phys. Chem. A*, 2015, **119**, 5386–5395.
46 (*a*) K. Raghavachari, G. W. Trucks and J. A. Pople, *Chem. Phys. Lett.*, 1989, **157**, 479–483; (*b*) G. D. Purvis III and R. J. Bartlett, *J. Chem. Phys.*, 1982, **76**, 1910–1918.
47 T. H. Dunning Jr., *J. Chem. Phys.*, 1989, **90**, 1007–1023.
48 C. Møller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618–622.
49 T. Helgaker, W. Klopper, H. Koch and J. Noga, *J. Chem. Phys.*, 1997, **106**, 9639–9646.
50 D. E. Woon and T. H. Dunning, *J. Chem. Phys.*, 1995, **103**, 4572–4585.
51 (*a*) D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652; (*b*) J. P. Perdew, K. Burke and Y. Wang, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**, 16533–16539.
52 Double and Triple-Zeta basis sets of the SNS family, are available in the Download section: http://smart.sns.it/.
53 I. M. Mills in *Molecular Spectroscopy: Modern Research*, ed. K.N. Rao and C.W. Mathews, Academic Press, 1972.
54 V. Barone, *J. Chem. Phys.*, 2005, **122**, 014108.
55 P. Pyykkö, *Mol. Phys.*, 2008, **106**, 1965–1974.
56 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 and D. J. Fox, *Gaussian 16, Revision B.01*, Gaussian, Inc., Wallingford CT, 2016.
57 T. J. Balle and W. H. Flygare, *Rev. Sci. Instrum.*, 1981, **52**, 33–45.
58 J.-U. Grabow, W. Stahl and H. Dreizler, *Rev. Sci. Instrum.*, 1996, **67**, 4072–4084.
59 J.-U. Grabow, Q. Gou and G. Feng, *72nd International Symposium on Molecular Spectroscopy*, TH03, Champaign-Urbana, 2017.
60 T. N. Wassermann, P. Zielke, J. J. Lee, C. Cézard and M. A. Suhm, *J. Phys. Chem. A*, 2007, **111**, 7437–7448.
61 Z. Kisiel, O. Dorosh, A. Maeda, I. R. Medvedev, F. C. De Lucia, E. Herbst, B. J. Drouin and J. C. Pearson, *Phys. Chem. Chem. Phys.*, 2010, **12**, 8329–8339.
62 G. J. Mead, E. R. Alonso, I. A. Finneran, P. B. Carroll and G. A. Blake, *J. Mol. Spectrosc.*, 2017, **335**, 68–73.
63 H. M. Pickett, *J. Mol. Spectrosc.*, 1991, **148**, 371–377.
64 J. K. G. Watson, in *Vibrational Spectra and Structure*, ed. J. R. Durig, Elsevier, 1977, p. 1.
65 F. Pawlowski, P. Jørgensen, J. Olsen, F. Hegelund, T. Helgaker, J. Gauss, K. L. Bak and J. F. Stanton, *J. Chem. Phys.*, 2002, **116**, 6482–6496.
66 C. Puzzarini and V. Barone, *Acc. Chem. Res.*, 2018, **51**, 548–556.
67 R. A. Kendall, T. H. Dunning Jr. and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796.
68 M. Mendolicchio, E. Penocchio, D. Licari, N. Tasinato and V. Barone, *J. Chem. Theory Comput.*, 2017, **13**, 3060–3075.
69 E. R. Johnson, S. Keinan, P. Mori-Sanchez, J. Contreras-Garcia, A. J. Cohen and W. Yang, *J. Am. Chem. Soc.*, 2010, **132**, 6498–6506.
70 R. F. W. Bader, *Chem. Rev.*, 1991, **91**, 893–928.
71 T. Lu and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580–592.
72 B. Jeziorski, R. Moszynski and K. Szalewicz, *Chem. Rev.*, 1994, **94**, 1887–1930.
73 T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno and C. D. Sherrill, *J. Chem. Phys.*, 2014, **140**, 094106.
74 R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince III, E. G. Hohenstein, U. Bozkaya, A. Yu. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer III, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford and C. D. Sherrill, *J. Chem. Theory Comput.*, 2017, **13**, 3185–3197.
75 E. D. Glendening, C. R. Landis and F. Weinhold, *WIREs Comput. Mol. Sci.*, 2012, **2**, 1–42.
76 S. Alessandrini, V. Barone and C. Puzzarini, *J. Chem. Theory Comput.*, 2020, **16**, 988–1006.

Appendix C

Tool for LSCD

When dealing with vibro-rotational spectra, spectroscopists might have to deal with thousands or even tens of thousands of lines. By using the lower state combination of differences (LSCD), the analysis of dense spectra can be eased, with the only prerequisite being the accurate knowledge of the rotational energies of the lower state. This is often achieved thanks to pure rotational measurements, especially if the lower state coincides with the vibrational ground state. Throughout the course of the Ph.D. project, a Python program based on LSCD has been written and tested, in order to facilitate the analysis of vibro-rotational spectra of linear molecules.

This Appendix is organized as follows. Section C.1 consists in a brief theoretical introduction, while Section C.2 shows the algorithms implemented in the program. Then, a description of the user interaction and of the input files is found in Section C.3. Finally, a test case is shown in Section C.4.

C.1 Theoretical introduction

The starting point is the equation for the rotational energy (E_{rot}) of a linear molecule in the vibrational ground state (g.s.), namely

$$E_{\text{rot}} = B_0 J(J+1) - D_0 J^2 (J+1)^2 + H_0 J^3 (J+1)^3 \,, \tag{C.1}$$

where B_0 is the g.s. rotational constant, D_0 and H_0 are the quartic and sextic centrifugal distortion terms, and J is the total angular momentum quantum number.[13] Given the selection rule on J,

$$\Delta J = \pm 1 \,, \tag{C.2}$$

and considering an absorption transition ($\Delta J = +1$), the frequency ν for a rotational transition can be expressed as[13]

$$\nu = 2B_0(J+1) - 4D_0(J+1)^3 + H_0(J+1)^3\left[(J+2)^3 - J^3\right]. \qquad \text{(C.3)}$$

However, also excited vibrational states can be populated and therefore have to be considered. The vibration affects the value of the rotational constant as shown in Equation (2.51) and the related note; a similar expression is found in the values of D_v and H_v.[13]

There is another effect of the interaction between rotation and vibration which takes place when considering degenerate bending modes. In a non-rotating molecule, the bending motions can be described as two movements occurring on orthogonal planes. Therefore, these vibrations can be considered equivalent to a rotation along the common axis. A degenerate bending vibrational state with vibrational quantum number v_r can be regarded as having components of angular momentum with rotational quantum numbers[13]

$$l_r = v_r, v_r - 2, v_r - 4, \ldots, -v_r\,. \qquad \text{(C.4)}$$

When more than one bending vibrational state is excited, l represents the total vibrational angular momentum quantum number ($l = \sum_r l_r$). Therefore, the rotational energy of a linear molecule in the v vibrational state is expressed as[130]

$$E_{\text{rot}} = B_v[J(J+1) - l^2] - D_v[J(J+1) - l^2]^2 + H_v[J(J+1) - l^2]^3\,, \qquad \text{(C.5)}$$

where J includes the vibrational angular momentum[130]

$$J = |l|, |l|+1, |l|+2, \ldots\,. \qquad \text{(C.6)}$$

Finally, considering $|l| \geq 1$, including the l-doubling due to a Coriolis interaction which separates the $\pm l$ degeneracy, and introducing the q coupling constant, one can obtain the expression for the rotational frequencies for an absorption transition with $\Delta l = 0$

$$\nu_{\pm} = 2B_v(J+1) \pm \frac{1}{2}q_r(v_r+1)(J+1) - 4D_v(J+1)\left[(J+1)^2 - l^2\right], \qquad \text{(C.7)}$$

where the sextic term has been neglected for the sake of conciseness.[130]

Having introduced the expressions for the rotational frequencies, a particular relationship between them can be obtained, based on the lower state combination of differences. From now on, the lower state is set to

coincide with the ground state. Indeed, considering transitions happening from the g.s. to an excited vibrational state v:

$$\nu_R : v, J, l \leftarrow 0, J-1\,; \tag{C.8}$$
$$\nu_P : v, J, l \leftarrow 0, J+1\,, \tag{C.9}$$

it follows

$$\nu_R = E_{\mathrm{rot}}(v, J, l) - E_{\mathrm{rot}}(0, J-1)\,; \tag{C.10}$$
$$\nu_P = E_{\mathrm{rot}}(v, J, l) - E_{\mathrm{rot}}(0, J+1)\,, \tag{C.11}$$

therefore

$$\Delta\nu = \nu_R - \nu_P = E_{\mathrm{rot}}(0, J+1) - E_{\mathrm{rot}}(0, J-1)\,. \tag{C.12}$$

Finally, Equation (C.12) can be rearranged to obtain

$$\nu_R = \nu_P + \Delta\nu = \nu_P + \Delta E_{0,J}\,, \tag{C.13}$$

where:

$$\Delta E_{0,J} = E_{\mathrm{rot}}(0, J+1) - E_{\mathrm{rot}}(0, J-1)\,. \tag{C.14}$$

C.2 Groundwork of the program

Given the availability of accurate estimates of $\Delta E_{0,J}$ for different values of J, this information can be used to estimate the ν_R position starting from the ν_P value and *vice versa*. Obviously, the lower state can be generalized to be different from the vibrational ground state. Starting from a group of R transitions which have been assigned to their value of J (namely, $\nu_{R,J'}$), the set of $\nu_{P,J'}$ can be built by augmenting each transition of the corresponding $\Delta E_{0,J}$ according to Equation (C.13). The frequencies of the newly built $\nu_{P,J'}$ set will be compared with a peak list provided by the user, to create the $\rho_{P,J'}$ set, in which the frequencies calculated using Equation (C.13) are replaced with the nearest frequencies belonging to the peak list. Finally, we can compute the root mean square (rms) error for a tentative assignment of i lines

$$\mathrm{rms} = \sqrt{\frac{\sum_i \left[\left(\left|\nu_{R,J',i} - \rho_{P,J',i}\right|\right) - \Delta E_{0,J,i}\right]^2}{i}}\,. \tag{C.15}$$

The value of the rms error is a measure of the accuracy of the tentative assignment. Furthermore, a number of additional features have been implemented:

i. the possibility to scan a series of J values starting from the initial one

$$J' \Rightarrow J' - x, \ldots, J', \ldots, J' + x\,; \tag{C.16}$$

ii. the possibility to scan over all known lower states v

$$\Delta E_{0,J} \Rightarrow \Delta E_{v,J}\,, \forall v\,; \tag{C.17}$$

iii. the counting of the number of calculated transitions which are close to the experimental ones within a threshold t

$$|\nu_{P,J'} - \rho_{P,J'}| \leq t\,. \tag{C.18}$$

If more states and/or J values are tested, the program will order the results by rms error values. The output of the program will include the tentative assignments and the best-fitting lower state.

Finally, the program calculates the difference between the best $\rho_{P,J'}$ set (the one which minimizes the rms of Equation C.15) and the initial $\nu_{R,J'}$ set (eventually corrected for J variations). Then, this is compared to the $\Delta E_{v,J}$ for each vibrational state v, in order to find the best candidate for the upper state, according to

$$\mathrm{rms}_v = \sqrt{\frac{\sum_i \left[\left(|\nu_{R,J',i} - \rho_{P,J',i}|\right) - \Delta E_{v,J,i}\right]^2}{i}}\,. \tag{C.19}$$

C.3 User interaction

The program can be used thorough the commandline or *via* a graphical user interface (GUI).

Input

The user will be able to provide the following information:

i. the `.egy` file, which is used to determine the values of $\Delta E_{0,J}$;

ii. the `.list` file, used to determine the set of states and parities;

iii. an experimental peak list;

iv. an assignment file, which includes a list of initial assignments (pairs of J and ν values) and their branch (R, P), a plausible lower state and its parity, and eventually a range of J values to scan.

Additional input can be provided:

v. the experimental spectrum from which the peak list was generated.

The `.egy` file

```
  IBLK INDX         EGY         ERR        PMIX          QNs
   int  int       float       float       float  :    3*int
1    1    1    0.000000    0.000000    1.000000  :  0  0  0
2    3    1  612.526871    0.042275    0.863569  :  0  0  1
3    3    2  499.753299    0.000012    0.941312  :  0  0  3
4 ...
```

The `.egy` file is generated by running the SPCAT program (see Ref. 202). It collects the values of the energy of all the vibro-rotational levels of the molecule of interest, ordered by quantum number value. The relevant part of this file is in the *EGY* and *QNs* columns, where the energy is expressed in wavenumbers and the corresponding quantum numbers are listed. For a linear molecule, the *QNs* section includes only three values, namely J, l and v.

The `.list` file

```
 #STATES   ···    NAME      TERM
     int   ···  !  str  ::  str
1      20   ···  !   GS  ::    0
2       1   ···  !   v6  ::    0
3       2   ···  !   v8  ::  1ef
4 ...
```

The list file is generated by the LINFIT1 pipeline for the fit of a linear rotor by L. Bizzocchi. It includes the number of total states in the first line (*#STATES*), their name and the term (in the "e, f" notation). After the first line, *#STATES* is the cardinal number of the state, which is set to 0 for the ground state. The format of the `.list` lines is identical to the third line of the `.par` file from the SPFIT program.

The peak list file

```
          FREQ          INT
         float        float
1  200.000000000 0.150000000
2  200.020150484 0.204012012
3  200.060480121 0.604801312
4 ...
```

The peak list file contains all the peaks of a portion of the vibro-rotational spectrum of the species of interest. It can be generated with any program, as long as it contains an x, y form of data which refer to the peak frequency and intensity.

The assignment file

```
 BRANCH STATE PARITY J_RANGE
    str   str    str     int
1     P   v10     1e      10
```

The first line of the assignment file contains the suggested identification for a group of lines. It must include the branch (P or R) of the starting assignments, the state and the parity (depending on the existing states for the molecule of interest) of the assumed lower state, and the range of values of increments of J to be tested.

```
 J_L           FREQ
 int          float
2  20     250.000000
3  21     250.084012
4  22     250.161312
5 ...
```

In the following lines, the suggested assignments for a group of transitions is reported: the value of the starting J and the frequency of the line are listed.

Output

By default, the program output is printed to screen and it consists in:

i. the rms error of the test(s);

ii. J and frequency values of the generated and original assignments of the test showing the lower rms error (with the correct adjustments on the J value, if required).

Additional output can be required:

iii. the intensity over frequency stem plot of the best generated and original assignments (where an experimental spectrum can be overlaid);

iv. the collection of all states read in the `.list` file.

Graphical user interface

A simple GUI has been realized using PyQt5 in order to ease the use of the program, and it is shown in Figure C.1. File dialogs help in the input of the different files. The test options are visible before executing the test. The text output is printed in three panels, while the stem plot (and, eventually, the experimental spectrum) have a dedicated and interactive panel. The value of the rms can be initialized, as well as the result panels. All information submitted in the first line of the assignment file can be modified afterwards.

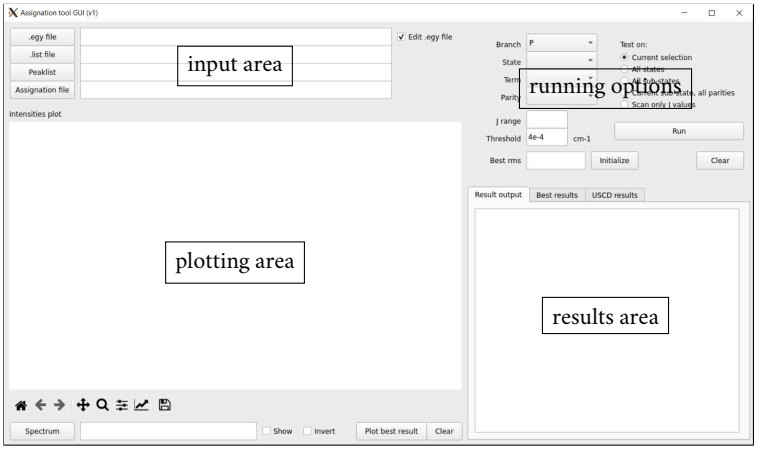


Figure C.1: GUI for the program, realized with PyQt5 and QtDesigner.

C.4 Case test

The extensive study on the vibro-rotational spectrum of DC_3N by Melosso *et al.* in Ref. 312 represents an ideal candidate for testing the program. For the sake of brevity, only one test example is here reported. For the assignment file, eleven lines belonging to the *R* branch of the $2\nu_7 \leftarrow \nu_7$ band have been chosen, and their initial value of J has been incremented by one. Having access to the experimental spectrum, a peak list file has been generated in the 180–330 cm^{-1} range. Finally, the `.egy` and `.list` files have been obtained, resorting respectively to the Spcat and the Linfil1 programs.

The LSCD analysis has been conducted over all 20 available vibrational states and over three possible variations of J $(0, \pm 1)$, leading to a total of 168 tests. The first five results, sorted by increasing rms value, are represented in Table C.1. The best guess, *i.e.*, the correct result, performs

significantly better than the rest, with the rms error more than 40 times lower than the second best guess. The last column of Table C.1 shows that all tentative assignments have a correspondence in the peak list within the threshold (4×10^{-4} cm^{-1}) defined in Equation (C.18). From this result, the upper state combination of differences (USCD) procedure has been carried out according to Equation (C.19), and the results are reported in Table C.2. Again, the correct result is evidently better than the rest (the rms error is more than 60 times lower). In conclusion, the program is able to identify the lower and the upper state for a vibro-rotational band, easing the process of assignment and analysis of dense spectra. The results have been obtained using the GUI (see Figure C.2), which allows to show the experimental spectrum and the position of the suggested transitions.

Table C.1: LSCD results for the $2\nu_7 \leftarrow \nu_7$ band of DC_3N. The last column indicates the number of transitions below threshold.

#	Lower state	Term	Parity	ΔJ	rms error	# lines
1	ν_7	P	1e	−1	3.994×10^{-5}	11/11
2	ν_6	P	1f	−1	1.835×10^{-3}	10/11
3	ν_4	S+	0	0	2.376×10^{-3}	9/11
4	$4\nu_7$	S+	0	0	4.390×10^{-3}	6/11
5	$\nu_6 + 2\nu_7$	P	(1,–2)f	+1	4.671×10^{-3}	4/11

Table C.2: USCD results for the $2\nu_7 \leftarrow \nu_7$ band of DC_3N, based on LSCD best result.

#	Upper state	Term	Parity	rms error
1	$2\nu_7$	S+	0	3.525×10^{-5}
2	$2\nu_7$	D	2e	2.408×10^{-3}
3	$2\nu_7$	D	2f	4.934×10^{-3}
4	$\nu_5 + 2\nu_7$	P	(1,0)e	9.593×10^{-3}
5	$\nu_6 + \nu_7$	S-	0f	1.098×10^{-2}

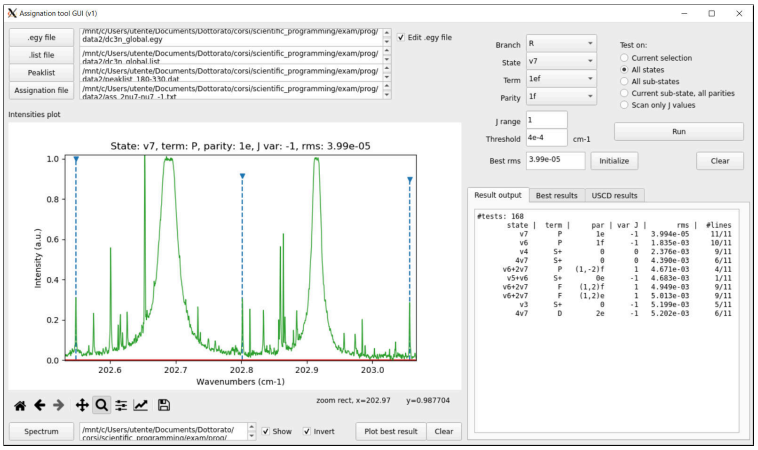


Figure C.2: GUI visualization of the LSCD results for the $2\nu_7 \leftarrow \nu_7$ band of DC_3N. A portion of the experimental spectrum (green line) is overlaid to the tentative assignments (dashed blu line).

Bibliography

[1] A. Dalgarno, "Astrochemistry – a summary," in *Symposium – International Astronomical Union*, vol. 120, pp. 577–581, Cambridge University Press, 1987.

[2] S. Yamamoto, *Introduction to Astrochemistry: Chemical Evolution from Interstellar Clouds to Star and Planet Formation*. Springer, 2017.

[3] A. McKellar, *Publications of the Astronomical Society of the Pacific*, vol. 52, no. 307, pp. 187–192, 1940.

[4] A. Douglas and G. Herzberg, *The Astrophysical Journal*, vol. 94, p. 381, 1941.

[5] C. H. Townes, "Microwave and radio-frequency resonance lines of interest to radio astronomy," in *Symposium – International Astronomical Union*, vol. 4, pp. 92–103, Cambridge University Press, 1957.

[6] A. Cheung, D. M. Rank, C. Townes, *et al.*, *Physical Review Letters*, vol. 21, no. 25, p. 1701, 1968.

[7] L. E. Snyder, D. Buhl, B. Zuckerman, and P. Palmer, *Physical Review Letters*, vol. 22, no. 13, p. 679, 1969.

[8] B. A. McGuire, *The Astrophysical Journal Supplement Series*, vol. 259, no. 2, p. 30, 2022.

[9] E. Herbst and E. F. Van Dishoeck, *Annual Review of Astronomy & Astrophysics*, vol. 47, pp. 427–480, 2009.

[10] C. Puzzarini, *International Journal of Quantum Chemistry*, vol. 117, no. 2, pp. 129–138, 2017.

[11] H. S. Müller, S. Thorwirth, D. Roth, and G. Winnewisser, *Astronomy & Astrophysics*, vol. 370, no. 3, pp. L49–L52, 2001.

[12] H. S. Müller, F. Schlöder, J. Stutzki, and G. Winnewisser, *Journal of Molecular Structure*, vol. 742, no. 1-3, pp. 215–227, 2005.

[13] W. Gordy and R. L. Cook, *Microwave molecular spectra*. Wiley, 1984.

[14] C. Puzzarini and V. Barone, *Physics of Life Reviews*, vol. 34, no. 35, pp. 143–146, 2020.

[15] C. Puzzarini and V. Barone, *Physics of Life Reviews*, vol. 32, pp. 59–94, 2020.

[16] R. L. Brown, W. Wild, and C. Cunningham, *Advances in Space Research*, vol. 34, no. 3, pp. 555–559, 2004.

[17] F. Tercero, J. López-Pérez, J. Gallego, *et al.*, *Astronomy & Astrophysics*, vol. 645, p. A37, 2021.

[18] M. Melosso, L. Dore, F. Tamassia, *et al.*, *The Journal of Physical Chemistry A*, vol. 124, no. 1, pp. 240–246, 2019.

[19] M. Agúndez, N. Marcelino, C. Cabezas, *et al.*, *Astronomy & Astrophysics*, vol. 657, p. A96, 2022.

[20] L. Dore, *Il linguaggio della simmetria. La teoria dei gruppi. Un'introduzione per chimici.* Pitagora Editrice, 2015.

[21] R. C. Fortenberry, *International Journal of Quantum Chemistry*, vol. 117, no. 2, pp. 81–91, 2017.

[22] J. Cernicharo, N. Marcelino, J. Pardo, *et al.*, *Astronomy & Astrophysics*, vol. 641, p. L9, 2020.

[23] J. Cernicharo, N. Marcelino, M. Agúndez, *et al.*, *Astronomy & Astrophysics*, vol. 642, p. L17, 2020.

[24] V. Barone, S. Alessandrini, M. Biczysko, *et al.*, *Nature Reviews Methods Primers*, vol. 1, no. 1, p. 38, 2021.

[25] B. F. Burke, F. Graham-Smith, and P. N. Wilkinson, *An introduction to radio astronomy*. Cambridge University Press, 2019.

[26] H. Karttunen, P. Kröger, H. Oja, *et al.*, *Fundamental astronomy*. Springer, 2007.

[27] S. M. Fortman, J. P. McMillan, C. F. Neese, *et al.*, *Journal of Molecular Spectroscopy*, vol. 280, pp. 11–20, 2012.

[28] C. Puzzarini, *Frontiers in Astronomy and Space Sciences*, vol. 7, p. 19, 2020.

[29] A. Pross and R. Pascal, *Open Biology*, vol. 3, no. 3, p. 120190, 2013.

[30] L. T. Troland, *The American Naturalist*, vol. 51, no. 606, pp. 321–350, 1917.

[31] W. Gilbert, *Nature*, vol. 319, no. 6055, pp. 618–618, 1986.

[32] H. S. Bernhardt, *Biology Direct*, vol. 7, no. 1, pp. 1–10, 2012.

[33] S. A. Kauffman, *Investigations*. Oxford University Press, 2000.

[34] G. Wächtershäuser, *Science*, vol. 289, no. 5483, pp. 1307–1308, 2000.

[35] B. K. Pearce, R. E. Pudritz, D. A. Semenov, and T. K. Henning, *Proceedings of the National Academy of Sciences*, vol. 114, no. 43, pp. 11327–11332, 2017.

[36] C. Chyba and C. Sagan, *Nature*, vol. 355, no. 6356, pp. 125–132, 1992.

[37] A. M. Shaw, *Astrochemistry: From astronomy to astrobiology*. John Wiley & Sons, 2007.

[38] A. Bellili, R. Linguerri, M. Hochlaf, and C. Puzzarini, *The Journal of Chemical Physics*, vol. 143, no. 18, p. 184314, 2015.

[39] J. M. Hollis, F. J. Lovas, and P. R. Jewell, *The Astrophysical Journal*, vol. 540, no. 2, p. L107, 2000.

[40] V. M. Rivilla, L. Colzi, I. Jiménez-Serra, *et al.*, *The Astrophysical Journal Letters*, vol. 929, no. 1, p. L11, 2022.

[41] A. Belloche, R. Garrod, H. Müller, *et al.*, *Astronomy & Astrophysics*, vol. 628, p. A10, 2019.

[42] B. Turner, H. Liszt, N. Kaifu, and A. Kisliakov, *The Astrophysical Journal*, vol. 201, pp. L149–L152, 1975.

[43] V. M. Rivilla, J. Martín-Pintado, I. Jiménez-Serra, *et al.*, *The Astrophysical Journal Letters*, vol. 899, no. 2, p. L28, 2020.

[44] B. Turner, *Astrophysical Journal, vol. 163, p. L35*, vol. 163, p. L35, 1971.

[45] R. A. Loomis, D. P. Zaleski, A. L. Steber, *et al.*, *The Astrophysical Journal Letters*, vol. 765, no. 1, p. L9, 2013.

[46] D. Navarro-Almaida, A. Fuente, L. Majumdar, *et al.*, *Astronomy & Astrophysics*, vol. 653, p. A15, 2021.

[47] Y. Hirahara, H. Suzuki, S. Yamamoto, *et al.*, *The Astrophysical Journal*, vol. 394, pp. 539–551, 1992.

[48] J. G. de la Concepción, C. Puzzarini, V. Barone, *et al.*, *The Astrophysical Journal*, vol. 922, no. 2, p. 169, 2021.

[49] D. C. Lis and P. F. Goldsmith, *The Astrophysical Journal*, vol. 356, pp. 195–210, 1990.

[50] Á. Sánchez-Monge, P. Schilke, A. Schmiedeke, *et al.*, *Astronomy & Astrophysics*, vol. 604, p. A6, 2017.

[51] D. M. Mehringer, L. E. Snyder, Y. Miao, and F. J. Lovas, *The Astrophysical Journal*, vol. 480, no. 1, p. L71, 1997.

[52] A. Belloche, K. Menten, C. Comito, *et al.*, *Astronomy & Astrophysics*, vol. 492, no. 3, pp. 769–773, 2008.

[53] D. Goldhaber and A. Betz, *The Astrophysical Journal*, vol. 279, pp. L55–L58, 1984.

[54] J. Cernicharo, C. Cabezas, J. Pardo, *et al.*, *Astronomy & Astrophysics*, vol. 672, p. L13, 2023.

[55] L. Zack, D. Halfen, and L. Ziurys, *The Astrophysical Journal Letters*, vol. 733, no. 2, p. L36, 2011.

[56] L. M. Ziurys, C. Savage, J. Highberger, *et al.*, *The Astrophysical Journal*, vol. 564, no. 1, p. L45, 2001.

[57] A. G. Tielens, *The physics and chemistry of the interstellar medium*. Cambridge University Press, 2005.

[58] W. Klemperer, *Nature*, vol. 227, pp. 1230–1230, 1970.

[59] E. Herbst and W. Klemperer, *The Astrophysical Journal*, vol. 185, pp. 505–534, 1973.

[60] W. D. Watson, *Astrophysical Journal, Vol. 188, pp. 35-42 (1974)*, vol. 188, pp. 35–42, 1974.

[61] E. Herbst, H.-H. Lee, D. Howe, and T. Millar, *Monthly Notices of the Royal Astronomical Society*, vol. 268, no. 2, pp. 335–344, 1994.

[62] A. Tielens, *Reviews of Modern Physics*, vol. 85, no. 3, p. 1021, 2013.

[63] A. G. Tielens, *Annual Review of Astronomy & Astrophysics*, vol. 46, pp. 289–337, 2008.

[64] R. Gredel, Y. Carpentier, G. Rouillé, *et al.*, *Astronomy & Astrophysics*, vol. 530, p. A26, 2011.

[65] W. Watson and E. Salpeter, *The Astrophysical Journal*, vol. 174, p. 321, 1972.

[66] W. Watson and E. Salpeter, *The Astrophysical Journal*, vol. 175, p. 659, 1972.

[67] W. D. Watson, *Reviews of Modern Physics*, vol. 48, no. 4, p. 513, 1976.

[68] T. Iguchi, *Publications of the Astronomical Society of Japan*, vol. 27, pp. 515–532, 1975.

[69] M. Allen and G. W. Robinson, *The Astrophysical Journal*, vol. 212, pp. 396–415, 1977.

[70] A. Tielens and W. Hagen, *Astronomy & Astrophysics*, vol. 114, pp. 245–260, 1982.

[71] P. Caselli, C. Walmsley, M. Tafalla, *et al.*, *The Astrophysical Journal*, vol. 523, no. 2, p. L165, 1999.

[72] T. I. Hasegawa, E. Herbst, and C. M. Leung, *The Astrophysical Journal Supplement Series*, vol. 82, pp. 167–195, 1992.

[73] R. Garrod, V. Wakelam, and E. Herbst, *Astronomy & Astrophysics*, vol. 467, no. 3, pp. 1103–1115, 2007.

[74] Y. Aikawa, E. Herbst, H. Roberts, and P. Caselli, *The Astrophysical Journal*, vol. 620, no. 1, p. 330, 2005.

[75] J. Cernicharo, C. Cabezas, M. Agúndez, *et al.*, *Astronomy & Astrophysics*, vol. 647, p. L3, 2021.

[76] N. Marcelino, J. Cernicharo, M. Agúndez, *et al.*, *The Astrophysical Journal*, vol. 665, no. 2, p. L127, 2007.

[77] J. Cernicharo and M. Guélin, *Astronomy & Astrophysics*, vol. 309, pp. L27–L30, 1996.

[78] A. J. Remijan, J. Hollis, F. Lovas, *et al.*, *The Astrophysical Journal*, vol. 664, no. 1, p. L47, 2007.

[79] K. Kawaguchi, R. Fujimori, S. Aimi, *et al.*, *Publications of the Astronomical Society of Japan*, vol. 59, no. 5, p. L47, 2007.

[80] M. Guélin, J. Cernicharo, M. Travers, *et al.*, *Astronomy & Astrophysics*, vol. 317, pp. L1–L4, 1997.

[81] M. McCarthy, C. Gottlieb, H. Gupta, and P. Thaddeus, *The Astrophysical Journal*, vol. 652, no. 2, p. L141, 2006.

[82] H. Suzuki, M. Ohishi, N. Kaifu, *et al.*, *Publications of the Astronomical Society of Japan*, vol. 38, pp. 911–917, 1986.

[83] M. Guélin, J. Cernicharo, C. Kahane, *et al.*, *Astronomy & Astrophysics*, vol. 175, pp. L5–L8, 1987.

[84] J. Cernicharo, C. Kahane, J. Gomez-Gonzalez, and M. Guélin, *Astronomy & Astrophysics*, vol. 164, pp. L1–L4, 1986.

[85] J. Cernicharo, C. Kahane, J. Gomez-Gonzalez, and M. Guélin, *Astronomy & Astrophysics*, vol. 167, no. 1, pp. L5–L7, 1986.

[86] J. Cernicharo, M. Guélin, and C. Walmsley, *Astronomy & Astrophysics*, vol. 172, p. L5, 1987.

[87] M. Guélin, S. Green, and P. Thaddeus, *The Astrophysical Journal*, vol. 224, pp. L27–L30, 1978.

[88] P. Thaddeus, C. Gottlieb, A. Hjalmarson, *et al.*, *The Astrophysical Journal*, vol. 294, pp. L49–L53, 1985.

[89] K. Tucker, M. Kutner, and P. Thaddeus, *The Astrophysical Journal*, vol. 193, pp. L115–L119, 1974.

[90] J. Cernicharo, M. Agúndez, C. Cabezas, *et al.*, *Astronomy & Astrophysics*, vol. 647, p. L2, 2021.

[91] H. E. Matthews and T. J. Sears, *The Astrophysical Journal*, vol. 272, pp. 149–153, 1983.

[92] J. Cernicharo, M. Agúndez, C. Cabezas, *et al.*, *Astronomy & Astrophysics*, vol. 649, p. L15, 2021.

[93] B. A. McGuire, A. M. Burkhardt, S. Kalenskii, *et al.*, *Science*, vol. 359, no. 6372, pp. 202–205, 2018.

[94] V. Kofman, P. Sarre, R. Hibbins, *et al.*, *Molecular Astrophysics*, vol. 7, pp. 19–26, 2017.

[95] J. Cernicharo, A. M. Heras, A. Tielens, *et al.*, *The Astrophysical Journal*, vol. 546, no. 2, p. L123, 2001.

[96] J. Cernicharo, M. Agúndez, R. Kaiser, *et al.*, *Astronomy & Astrophysics*, vol. 652, p. L9, 2021.

[97] B. A. McGuire, R. A. Loomis, A. M. Burkhardt, *et al.*, *Science*, vol. 371, no. 6535, pp. 1265–1269, 2021.

[98] M. C. McCarthy, K. L. K. Lee, R. A. Loomis, *et al.*, *Nature Astronomy*, vol. 5, no. 2, pp. 176–180, 2021.

[99] K. L. K. Lee, P. B. Changala, R. A. Loomis, *et al.*, *The Astrophysical Journal Letters*, vol. 910, no. 1, p. L2, 2021.

[100] M. L. Sita, P. B. Changala, C. Xue, *et al.*, *The Astrophysical Journal Letters*, vol. 938, no. 2, p. L12, 2022.

[101] J. Cernicharo, M. Agúndez, R. Kaiser, *et al.*, *Astronomy & Astrophysics*, vol. 655, p. L1, 2021.

[102] J. Cernicharo, R. Fuentetaja, M. Agúndez, *et al.*, *Astronomy & Astrophysics*, vol. 663, p. L9, 2022.

[103] K. L. K. Lee, R. A. Loomis, A. M. Burkhardt, *et al.*, *The Astrophysical journal letters*, vol. 908, no. 1, p. L11, 2021.

[104] K. L. K. Lee, B. A. McGuire, and M. C. McCarthy, *Physical Chemistry Chemical Physics*, vol. 21, no. 6, pp. 2946–2956, 2019.

[105] P. Theulé, F. Borget, F. Mispelaer, *et al.*, *Astronomy & Astrophysics*, vol. 534, p. A64, 2011.

[106] S. Kwok, *Astrophysics and Space Science*, vol. 319, pp. 5–21, 2009.

[107] B. J. Esselman, M. A. Zdanovskaia, W. H. Styers, *et al.*, *The Journal of Physical Chemistry A*, vol. 127, no. 8, pp. 1909–1922, 2023.

[108] D. Tokaryk, S. Culligan, B. Billinghurst, and J. Van Wijngaarden, *Journal of Molecular Spectroscopy*, vol. 270, no. 1, pp. 56–60, 2011.

[109] T. J. Barnum, M. A. Siebert, K. L. K. Lee, *et al.*, *The Journal of Physical Chemistry A*, vol. 126, no. 17, pp. 2716–2728, 2022.

[110] T. J. Barnum, K. L. K. Lee, and B. A. McGuire, *ACS Earth and Space Chemistry*, vol. 5, no. 11, pp. 2986–2994, 2021.

[111] R. Simbizi, D. Nduwimana, J. Niyoncuti, *et al.*, *RSC advances*, vol. 12, no. 39, pp. 25332–25341, 2022.

[112] W. D. Fuller, R. A. Sanchez, and L. E. Orgel, *Journal of Molecular Biology*, vol. 67, no. 1, pp. 25–33, 1972.

[113] O. Leslie E, *Critical Reviews in Biochemistry and Molecular Biology*, vol. 39, no. 2, pp. 99–123, 2004.

[114] M. W. Powner, B. Gerland, and J. D. Sutherland, *Nature*, vol. 459, no. 7244, pp. 239–242, 2009.

[115] S. Becker, J. Feldmann, S. Wiedemann, *et al.*, *Science*, vol. 366, no. 6461, pp. 76–82, 2019.

[116] Q. P. Tran, Z. R. Adam, and A. C. Fahrenbach, *Life*, vol. 10, no. 12, p. 352, 2020.

[117] F. M. Kruse, J. S. Teichert, and O. Trapp, *Chemistry – A European Journal*, vol. 26, no. 65, pp. 14776–14790, 2020.

[118] L. E. Snyder, F. J. Lovas, J. Hollis, *et al.*, *The Astrophysical Journal*, vol. 619, no. 2, p. 914, 2005.

[119] P. Jones, M. Cunningham, P. D. Godfrey, and D. M. Cragg, *Monthly Notices of the Royal Astronomical Society*, vol. 374, no. 2, pp. 579–589, 2007.

[120] M. Cunningham, P. Jones, P. D. Godfrey, *et al.*, *Monthly Notices of the Royal Astronomical Society*, vol. 376, no. 3, pp. 1201–1210, 2007.

[121] N. Kaifu, M. Morimoto, K. Nagane, *et al.*, *The Astrophysical Journal*, vol. 191, pp. L135–L137, 1974.

[122] N. Fourikis, K. Takagi, and M. Morimoto, *The Astrophysical Journal*, vol. 191, p. L139, 1974.

[123] P. D. Holtom, C. J. Bennett, Y. Osamura, *et al.*, *The Astrophysical Journal*, vol. 626, no. 2, p. 940, 2005.

[124] M. Förstel, A. Bergantini, P. Maksyutenko, *et al.*, *The Astrophysical Journal*, vol. 845, no. 1, p. 83, 2017.

[125] S. Zeng, I. Jiménez-Serra, V. M. Rivilla, *et al.*, *The Astrophysical Journal Letters*, vol. 920, no. 2, p. L27, 2021.

[126] J. M. Brown, *Molecular spectroscopy*, vol. 75. Oxford University Press Oxford, 1998.

[127] C. Puzzarini, J. F. Stanton, and J. Gauss, *International Reviews in Physical Chemistry*, vol. 29, no. 2, pp. 273–367, 2010.

[128] C. Puzzarini, J. Bloino, N. Tasinato, and V. Barone, *Chemical Reviews*, 2019.

[129] I. M. Mills, "Vibration-rotation structure in asymmetric- and symmetric-top molecules," in *Molecular Spectroscopy: Modern Research* (K. N. Rao and C. W. Matthews, eds.), vol. 1, pp. 115–140, Academic Press, 1972.

[130] D. Papoušek and M. Aliev, *Molecular Vibrational-rotational Spectra: Theory and Applications of High Resolution Infrared, Microwave, and Raman Spectroscopy of Polyatomic Molecules*. Elsevier Scientific Publishing Company, 1982.

[131] H. V. L. Nguyen and I. Kleiner, *Physical Sciences Reviews*, 2020.

[132] F. Jensen, *Introduction to computational chemistry*. John Wiley & Sons, 2017.

[133] I. Shavitt and R. J. Bartlett, *Many-body methods in chemistry and physics: MBPT and coupled-cluster theory*. Cambridge University Press, 2009.

[134] C. J. Cramer, *Essentials of computational chemistry: theories and models*. John Wiley & Sons, 2013.

[135] T. H. Dunning Jr, *The Journal of Chemical Physics*, vol. 90, no. 2, pp. 1007–1023, 1989.

[136] R. A. Kendall, T. H. Dunning Jr, and R. J. Harrison, *The Journal of Chemical Physics*, vol. 96, no. 9, pp. 6796–6806, 1992.

[137] D. E. Woon and T. H. Dunning Jr, *The Journal of Chemical Physics*, vol. 103, no. 11, pp. 4572–4585, 1995.

[138] E. Papajak, H. R. Leverentz, J. Zheng, and D. G. Truhlar, *Journal of Chemical Theory and Computation*, vol. 5, no. 5, pp. 1197–1202, 2009.

[139] E. Papajak, J. Zheng, X. Xu, *et al.*, *Journal of Chemical Theory and Computation*, vol. 7, no. 10, pp. 3027–3034, 2011.

[140] E. Papajak and D. G. Truhlar, *Journal of Chemical Theory and Computation*, vol. 7, no. 1, pp. 10–18, 2011.

[141] E. Papajak and D. G. Truhlar, *Journal of Chemical Theory and Computation*, vol. 6, no. 3, pp. 597–601, 2010.

[142] V. Barone and P. Cimino, *Chemical Physics Letters*, vol. 454, no. 1-3, pp. 139–143, 2008.

[143] V. Barone, P. Cimino, and E. Stendardo, *Journal of Chemical Theory and Computation*, vol. 4, no. 5, pp. 751–764, 2008.

[144] V. Barone, M. Biczysko, and J. Bloino, *Physical Chemistry Chemical Physics*, vol. 16, no. 5, pp. 1759–1787, 2014.

[145] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *Chemical Physics Letters*, vol. 157, no. 6, pp. 479–483, 1989.

[146] R. J. Bartlett, J. Watts, S. Kucharski, and J. Noga, *Chemical Physics Letters*, vol. 165, no. 6, pp. 513–522, 1990.

[147] A. D. Becke, *Physical Review A*, vol. 38, no. 6, p. 3098, 1988.

[148] C. Lee, W. Yang, and R. G. Parr, *Physical Review B*, vol. 37, no. 2, p. 785, 1988.

[149] B. Miehlich, A. Savin, H. Stoll, and H. Preuss, *Chemical Physics Letters*, vol. 157, no. 3, pp. 200–206, 1989.

[150] J. P. Perdew, K. Burke, and M. Ernzerhof, *Physical Review Letters*, vol. 77, no. 18, p. 3865, 1996.

[151] J. P. Perdew, K. Burke, and M. Ernzerhof, *Physical Review Letters*, vol. 78, pp. 1396–1396, Feb 1997.

[152] J. P. Perdew, *Physical Review B*, vol. 33, no. 12, p. 8822, 1986.

[153] A. D. Becke, *The Journal of Chemical Physics*, vol. 98, no. 7, pp. 5648–5652, 1993.

[154] S. Grimme, *The Journal of Chemical Physics*, vol. 124, no. 3, p. 034108, 2006.

[155] S. Kozuch and J. M. Martin, *Physical Chemistry Chemical Physics*, vol. 13, no. 45, pp. 20104–20107, 2011.

[156] S. Kozuch, D. Gruzman, and J. M. Martin, *The Journal of Physical Chemistry C*, vol. 114, no. 48, pp. 20801–20808, 2010.

[157] Y. Zhang and W. Yang, *Physical Review Letters*, vol. 80, no. 4, p. 890, 1998.

[158] G. Santra, N. Sylvetsky, and J. M. Martin, *The Journal of Physical Chemistry A*, vol. 123, no. 24, pp. 5129–5143, 2019.

[159] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, *The Journal of Chemical Physics*, vol. 132, no. 15, p. 154104, 2010.

[160] S. Grimme, S. Ehrlich, and L. Goerigk, *Journal of Computational Chemistry*, vol. 32, no. 7, pp. 1456–1465, 2011.

[161] M. Heckert, M. Kállay, and J. Gauss, *Molecular Physics*, vol. 103, no. 15-16, pp. 2109–2115, 2005.

[162] M. Heckert, M. Kállay, D. P. Tew, *et al.*, *The Journal of Chemical Physics*, vol. 125, no. 4, p. 044108, 2006.

[163] D. Feller, *The Journal of Chemical Physics*, vol. 98, no. 9, pp. 7059–7071, 1993.

[164] T. Helgaker, W. Klopper, H. Koch, and J. Noga, *The Journal of Chemical Physics*, vol. 106, no. 23, pp. 9639–9646, 1997.

[165] C. Puzzarini and V. Barone, *Physical Chemistry Chemical Physics*, vol. 13, no. 15, pp. 7189–7197, 2011.

[166] S. Alessandrini, V. Barone, and C. Puzzarini, *Journal of Chemical Theory and Computation*, vol. 16, no. 2, pp. 988–1006, 2019.

[167] V. Barone, *The Journal of Chemical Physics*, vol. 122, no. 1, p. 014108, 2005.

[168] V. Barone, M. Biczysko, and C. Puzzarini, *Accounts of Chemical Research*, vol. 48, no. 5, pp. 1413–1422, 2015.

[169] M. Piccardo, E. Penocchio, C. Puzzarini, *et al.*, *The Journal of Physical Chemistry A*, vol. 119, no. 10, pp. 2058–2082, 2015.

[170] M. Piccardo, E. Penocchio, C. Puzzarini, *et al.*, *The Journal of Physical Chemistry A*, vol. 120, no. 20, pp. 3754–3754, 2016.

[171] E. Penocchio, M. Piccardo, and V. Barone, *Journal of Chemical Theory and Computation*, vol. 11, no. 10, pp. 4689–4707, 2015.

[172] E. Penocchio, M. Piccardo, and V. Barone, *Journal of Chemical Theory and Computation*, vol. 12, no. 6, pp. 3001–3001, 2016.

[173] C. Puzzarini and J. F. Stanton, *Physical Chemistry Chemical Physics*, 2023.

[174] C. Puzzarini, M. Heckert, and J. Gauss, *The Journal of Chemical Physics*, vol. 128, no. 19, p. 194108, 2008.

[175] S. Alessandrini, J. Gauss, and C. Puzzarini, *Journal of Chemical Theory and Computation*, vol. 14, no. 10, pp. 5360–5371, 2018.

[176] V. Barone, M. Biczysko, J. Bloino, *et al.*, *Journal of Chemical Theory and Computation*, vol. 11, no. 9, pp. 4342–4363, 2015.

[177] V. Barone, M. Biczysko, J. Bloino, and C. Puzzarini, *The Journal of Chemical Physics*, vol. 141, no. 3, p. 034107, 2014.

[178] J. Lupi, C. Puzzarini, C. Cavallotti, and V. Barone, *Journal of Chemical Theory and Computation*, vol. 16, no. 8, pp. 5090–5104, 2020.

[179] J. Lupi, S. Alessandrini, C. Puzzarini, and V. Barone, *Journal of Chemical Theory and Computation*, vol. 17, no. 11, pp. 6974–6992, 2021.

[180] J. Kraitchman, *American Journal of Physics*, vol. 21, no. 1, pp. 17–24, 1953.

[181] J. K. Watson, *Journal of Molecular Spectroscopy*, vol. 48, no. 3, pp. 479–502, 1973.

[182] P. Pulay, W. Meyer, and J. E. Boggs, *The Journal of Chemical Physics*, vol. 68, no. 11, pp. 5077–5085, 1978.

[183] F. Pawłowski, P. Jørgensen, J. Olsen, *et al.*, *The Journal of Chemical Physics*, vol. 116, no. 15, pp. 6482–6496, 2002.

[184] J. Pople, R. Krishnan, H. Schlegel, and J. Binkley, *International Journal of Quantum Chemistry*, vol. 14, no. 5, pp. 545–560, 1978.

[185] M. Mendolicchio, E. Penocchio, D. Licari, *et al.*, *Journal of Chemical Theory and Computation*, vol. 13, no. 6, pp. 3060–3075, 2017.

[186] L. Bartell, D. Romenesko, and T. Wong, "Augmented analyses: Method of predicate observations," in *Molecular Structure by Diffraction Methods* (G. Sim and L. Sutton, eds.), vol. 3, pp. 72–80, The Royal Society of Chemistry, 1975.

[187] G. Ceselin, V. Barone, and N. Tasinato, *Journal of Chemical Theory and Computation*, vol. 17, no. 11, pp. 7290–7311, 2021.

[188] A. Melli, F. Tonolo, V. Barone, and C. Puzzarini, *The Journal of Physical Chemistry A*, vol. 125, no. 45, pp. 9904–9916, 2021.

[189] K. Rohlfs, T. L. Wilson, and S. Hüttemeister, *Tools of radio astronomy*. Springer Science & Business Media, 2013.

[190] J. Baars, B. Hooghoudt, P. Mezger, and M. de Jonge, *Astronomy & Astrophysics*, vol. 175, no. 1–2, pp. 319–326, 1987.

[191] B. A. McGuire, C. N. Shingledecker, E. R. Willis, *et al.*, *The Astrophysical Journal Letters*, vol. 851, no. 2, p. L46, 2017.

[192] L. Kolesniková, A. Belloche, J. Koucký, *et al.*, *Astronomy & Astrophysics*, vol. 659, p. A111, 2022.

[193] N. Jiang, M. Melosso, F. Tamassia, *et al.*, *Frontiers in Astronomy and Space Sciences*, vol. 8, p. 656295, 2021.

[194] P. F. Goldsmith and W. D. Langer, *The Astrophysical Journal*, vol. 517, no. 1, p. 209, 1999.

[195] M. J. Frisch, G. W. Trucks, H. B. Schlegel, *et al.*, "Gaussian16," 2016. Wallingford CT. See `http://gaussian.com`.

[196] J. F. Stanton, J. Gauss, M. E. Harding, and P. G. Szalay, "Cfour, Coupled-Cluster techiques for Computational Chemistry," 2008. See `http://www.cfour.de`.

[197] M. Kállay, P. R. Nagy, D. Mester, *et al.*, *The Journal of Chemical Physics*, vol. 152, no. 7, p. 074107, 2020.

[198] M. Kállay, P. R. Nagy, D. Mester, *et al.*, "Mrcc, a quantum chemical program suite," 2005. See `http://www.mrcc.hu`.

[199] C. M. Western, *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 186, pp. 221–242, 2017.

[200] C. M. Western and B. E. Billinghurst, *Physical Chemistry Chemical Physics*, vol. 21, no. 26, pp. 13986–13999, 2019.

[201] C. M. Western and B. E. Billinghurst, *Physical Chemistry Chemical Physics*, vol. 19, no. 16, pp. 10222–10226, 2017.

[202] H. M. Pickett, *Journal of Molecular Spectroscopy*, vol. 148, no. 2, pp. 371–377, 1991.

[203] B. A. McGuire and co-workers, "Astromol: A database of molecules detected in space," 2021. See `https://github.com/bmcguir2/astromol`.

[204] J. M. Hollas, *Modern spectroscopy*. John Wiley & Sons, 2004.

[205] L. Dore, L. Bizzocchi, and C. Degli Esposti, *Astronomy & Astrophysics*, vol. 544, p. A19, 2012.

[206] M. Melosso, C. Degli Esposti, and L. Dore, *The Astrophysical Journal Supplement Series*, vol. 233, no. 1, p. 15, 2017.

[207] J.-U. Grabow, W. Stahl, and H. Dreizler, *Review of Scientific Instruments*, vol. 67, no. 12, pp. 4072–4084, 1996.

[208] T. Mandai, T. Nakata, H. Murayama, *et al.*, *Tetrahedron Letters*, vol. 31, no. 49, pp. 7179–7180, 1990.

[209] Y. Bergman, P. Perlmutter, and N. Thienthong, *Green Chemistry*, vol. 6, no. 11, pp. 539–540, 2004.

[210] D. R. Boyd, R. Hamilton, N. T. Thompson, and M. E. Stubbs, *Tetrahedron Letters*, vol. 20, no. 34, pp. 3201–3204, 1979.

[211] S. Gueugnot and G. Linstrumelle, *Tetrahedron Letters*, vol. 34, no. 24, pp. 3853–3856, 1993.

[212] R. J. Halter, R. L. Fimmen, R. J. McMahon, *et al.*, *Journal of the American Chemical Society*, vol. 123, no. 49, pp. 12353–12363, 2001.

[213] R. C. Woods, *Journal of Molecular Structure*, vol. 97, pp. 195–202, 1983.

[214] G. Cazzoli, V. Lattanzi, T. Kirsch, *et al.*, *Astronomy & Astrophysics*, vol. 591, p. A126, 2016.

[215] M. Melosso, L. Bizzocchi, H. Gazzeh, *et al.*, *Chemical Communications*, vol. 58, no. 16, pp. 2750–2753, 2022.

[216] M. Melosso, A. Melli, C. Puzzarini, *et al.*, *Astronomy & Astrophysics*, vol. 609, p. A121, 2018.

[217] A. Melli, M. Melosso, N. Tasinato, *et al.*, *The Astrophysical Journal*, vol. 855, no. 2, p. 123, 2018.

[218] M. Melosso, B. A. McGuire, F. Tamassia, *et al.*, *ACS Earth and Space Chemistry*, vol. 3, no. 7, pp. 1189–1195, 2019.

[219] J. August, H. Kroto, D. McNaughton, *et al.*, *Journal of Molecular Spectroscopy*, vol. 130, no. 2, pp. 424–430, 1988.

[220] M. C. McCarthy, K. L. K. Lee, P. B. Carroll, *et al.*, *The Journal of Physical Chemistry A*, vol. 124, no. 25, pp. 5170–5181, 2020.

[221] J.-C. Guillemin and J.-M. Denis, *Angewandte Chemie International Edition in English*, vol. 21, no. 9, pp. 690–690, 1982.

[222] D. H. Hunter and S. Sim, *Canadian Journal of Chemistry*, vol. 50, no. 5, pp. 669–677, 1972.

[223] R. Kupfer and U. H. Brinker, *The Journal of Organic Chemistry*, vol. 61, no. 12, pp. 4185–4186, 1996.

[224] P. Młynarz, T. Ptak, A. Czernicka, *et al.*, *Journal of Molecular Structure*, vol. 991, no. 1-3, pp. 18–23, 2011.

[225] I. Pecnikaj, F. Foschi, R. Bucci, *et al.*, *European Journal of Organic Chemistry*, vol. 2019, no. 39, pp. 6707–6713, 2019.

[226] C. Fang, M. Li, X. Hu, *et al.*, *Advanced Synthesis & Catalysis*, vol. 358, no. 7, pp. 1157–1163, 2016.

[227] C. Keussen and H. Dreizler, *Zeitschrift für Naturforschung A*, vol. 46, no. 6, pp. 527–534, 1991.

[228] S. C. Mehrotra, L. L. Griffin, C. O. Britt, and J. E. Boggs, *Journal of Molecular Spectroscopy*, vol. 64, no. 2, pp. 244–251, 1977.

[229] S. Thorwirth, M. C. McCarthy, J. B. Dudek, and P. Thaddeus, *Journal of Molecular Spectroscopy*, vol. 225, no. 1, pp. 93–95, 2004.

[230] T. Kojima, T. Ogata, and S. Maeda, *Chemistry Letters*, vol. 5, no. 6, pp. 607–610, 1976.

[231] L. Engelbrecht and D. Sutter, *Zeitschrift für Naturforschung A*, vol. 31, no. 6, pp. 670–672, 1976.

[232] L. Vereecken, P. De Groof, and J. Peeters, *Physical Chemistry Chemical Physics*, vol. 5, no. 22, pp. 5070–5076, 2003.

[233] F. Tonolo, J. Lupi, C. Puzzarini, and V. Barone, *The Astrophysical Journal*, vol. 900, no. 1, p. 85, 2020.

[234] M. R. Dash and B. Rajakumar, *Physical Chemistry Chemical Physics*, vol. 17, no. 5, pp. 3142–3156, 2015.

[235] J. Lupi, C. Puzzarini, and V. Barone, *The Astrophysical Journal Letters*, vol. 903, no. 2, p. L35, 2020.

[236] F. Vazart, D. Calderini, C. Puzzarini, *et al.*, *Journal of Chemical Theory and Computation*, vol. 12, no. 11, pp. 5385–5397, 2016.

[237] S. Alessandrini, F. Tonolo, and C. Puzzarini, *The Journal of Chemical Physics*, vol. 154, no. 5, p. 054306, 2021.

[238] C. Puzzarini and G. Cazzoli, *Journal of Molecular Spectroscopy*, vol. 240, no. 2, pp. 260–264, 2006.

[239] C. Blom and A. Bauder, *Chemical Physics Letters*, vol. 88, no. 1, pp. 55–58, 1982.

[240] C. Puzzarini, E. Penocchio, M. Biczysko, and V. Barone, *Journal of Physical Chemistry A*, vol. 118, no. 33, pp. 6648–6656, 2014.

[241] R. A. Motiyenko, I. A. Armieieva, L. Margulès, *et al.*, *Astron: Astrophys.*, vol. 623, p. A162, 2019.

[242] A. I. Jaman and R. Bhattacharya, *Journal of Atomic and Molecular Physics*, vol. 2012, 2012.

[243] A. Melli, S. Potenti, M. Melosso, *et al.*, *Chemistry – A European Journal*, vol. 26, no. 65, pp. 15016–15022, 2020.

[244] R. D. Brown, P. D. Godfrey, and D. A. Winkler, *Chemical Physics*, vol. 59, no. 3, pp. 243–247, 1981.

[245] B. A. McGuire, A. M. Burkhardt, R. A. Loomis, *et al.*, *The Astrophysical Journal Letters*, vol. 900, no. 1, p. L10, 2020.

[246] O. Desyatnyk, L. Pszczółkowski, S. Thorwirth, *et al.*, *Physical Chemistry Chemical Physics*, vol. 7, no. 8, pp. 1708–1715, 2005.

[247] L. Bizzocchi, D. Prudenzano, V. M. Rivilla, *et al.*, *Astronomy & Astrophysics*, vol. 640, p. A98, 2020.

[248] N. C. Craig, P. Groner, and D. C. McKean, *Journal of Physical Chemistry A*, vol. 110, no. 23, pp. 7461–7469, 2006.

[249] A. Jabri, L. Kolesniková, E. Alonso, *et al.*, *Journal of Molecular Spectroscopy*, vol. 372, p. 111333, 2020.

[250] M. Bogey, C. Demuynck, J. Destombes, and Y. Vallee, *Journal of Molecular Spectroscopy*, vol. 172, no. 2, pp. 344–351, 1995.

[251] W. Caminati, B. Vogelsanger, and A. Bauder, *Journal of Molecular Spectroscopy*, vol. 128, no. 2, pp. 384–398, 1988.

[252] S. Thorwirth and H. Lichau, *Astronomy & Astrophysics*, vol. 398, no. 2, pp. L11–L13, 2003.

[253] S. Thorwirth, M. E. Harding, J. B. Dudek, and M. C. McCarthy, *Journal of Molecular Spectroscopy*, vol. 350, pp. 10–17, 2018.

[254] R. W. Larsen, F. Pawłowski, F. Hegelund, *et al.*, *Physical Chemistry Chemical Physics*, vol. 5, no. 22, pp. 5031–5037, 2003.

[255] H. S. P. Müller, A. Belloche, K. M. Menten, *et al.*, *Journal of Molecular Spectroscopy*, vol. 251, no. 1-2, pp. 319–325, 2008.

[256] V. Barone, G. Ceselin, M. Fusè, and N. Tasinato, *Frontiers in Chemistry*, vol. 8, p. 584203, 2020.

[257] D. McNaughton, O. Osman, and H. Kroto, *Journal of Molecular Structure*, vol. 190, pp. 195–204, 1988.

[258] M. B. Gardner, B. R. Westbrook, R. C. Fortenberry, and T. J. Lee, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 248, p. 119184, 2021.

[259] Y. Ohshima, S. Yamamoto, M. Nakata, and K. Kuchitsu, *Journal of Physical Chemistry*, vol. 91, no. 18, pp. 4696–4700, 1987.

[260] A. Melli, M. Melosso, L. Bizzocchi, *et al.*, *The Journal of Physical Chemistry A*, 2022.

[261] A. Bouchy, J. Demaison, G. Roussy, and J. Barriol, *Journal of Molecular Structure*, vol. 18, no. 2, pp. 211–217, 1973.

[262] G. Schwahn, R. Schieder, M. Bester, and G. Winnewisser, *Journal of Molecular Spectroscopy*, vol. 116, no. 2, pp. 263–270, 1986.

[263] M. Lattelais, F. Pauzat, Y. Ellinger, and C. Ceccarelli, *The Astrophysical Journal Letters*, vol. 696, no. 2, p. L133, 2009.

[264] M. Lattelais, F. Pauzat, Y. Ellinger, and C. Ceccarelli, *Astronomy & Astrophysics*, vol. 519, p. A30, 2010.

[265] M. Lattelais, M. Bertin, H. Mokrane, *et al.*, *Astronomy & Astrophysics*, vol. 532, p. A12, 2011.

[266] G. K. Friestad and A. K. Mathies, *Tetrahedron*, vol. 63, no. 12, pp. 2541–2569, 2007.

[267] X.-Q. Zhu, Q.-Y. Liu, Q. Chen, and L.-R. Mei, *The Journal of Organic Chemistry*, vol. 75, no. 3, pp. 789–808, 2010.

[268] J. Gálvez and A. Guirado, *Journal of Computational Chemistry*, vol. 31, no. 3, pp. 520–531, 2010.

[269] C. Puzzarini, *The Journal of Physical Chemistry A*, vol. 119, no. 47, pp. 11614–11622, 2015.

[270] K. Fukui, *Accounts of Chemical Research*, vol. 14, no. 12, pp. 363–368, 1981.

[271] L. Coudert, F. Lovas, R. Suenram, and J. Hougen, *The Journal of Chemical Physics*, vol. 87, no. 11, pp. 6290–6299, 1987.

[272] J. H. Lee, S. Gupta, W. Jeong, *et al.*, *Angewandte Chemie International Edition*, vol. 51, no. 43, pp. 10851–10855, 2012.

[273] K.-N. T. Tseng, A. M. Rizzi, and N. K. Szymczak, *Journal of the American Chemical Society*, vol. 135, no. 44, pp. 16352–16355, 2013.

[274] T. Achard, J. Egly, M. Sigrist, *et al.*, *Chemistry – A European Journal*, vol. 25, no. 58, pp. 13271–13274, 2019.

[275] M. Nava, M.-A. Martin-Drumel, C. A. Lopez, *et al.*, *Journal of the American Chemical Society*, vol. 138, no. 36, pp. 11441–11444, 2016.

[276] T. I. Crowell and R. K. McLeod, *The Journal of Organic Chemistry*, vol. 32, no. 12, pp. 4030–4033, 1967.

[277] J.-M. Huang, J.-F. Zhang, Y. Dong, and W. Gong, *The Journal of Organic Chemistry*, vol. 76, no. 9, pp. 3511–3514, 2011.

[278] J. van Schijndel, D. Molendijk, H. Spakman, *et al.*, *Green Chemistry Letters and Reviews*, vol. 12, no. 3, pp. 323–331, 2019.

[279] M. Melosso, S. Alessandrini, L. Spada, *et al.*, "Rotational spectra and semi-experimental structures of furonitrile and its water cluster." submitted to *Physical Chemistry Chemical Physics*.

[280] A. Melli, M. Melosso, K. G. Lengsfeld, *et al.*, *Molecules*, vol. 27, no. 10, p. 3278, 2022.

[281] R. Kydd and I. Mills, *Journal of Molecular Spectroscopy*, vol. 42, no. 2, pp. 320–326, 1972.

[282] E. Ye and R. P. Bettens, *Journal of Molecular Spectroscopy*, vol. 223, no. 1, pp. 73–79, 2004.

[283] D. Christen, D. Norbury, D. G. Lister, and P. Palmieri, *Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics*, vol. 71, pp. 438–446, 1975.

[284] H. Møllendal and A. Konovalov, *The Journal of Physical Chemistry A*, vol. 114, no. 5, pp. 2151–2156, 2010.

[285] D. Lister and J. Tyler, *Chemical Communications (London)*, no. 6, pp. 152–153, 1966.

[286] A. Hatta, M. Suzuki, and K. Kozima, *Bulletin of the chemical society of Japan*, vol. 46, no. 8, pp. 2321–2323, 1973.

[287] D. Lister, J. Tyler, J. Høg, and N. W. Larsen, *Journal of Molecular Structure*, vol. 23, no. 2, pp. 253–264, 1974.

[288] G. Roussy and A. Nonat, *Journal of Molecular Spectroscopy*, vol. 118, no. 1, pp. 180–188, 1986.

[289] B. Kleibömer and D. Sutter, *Zeitschrift für Naturforschung A*, vol. 43, no. 6, pp. 561–571, 1988.

[290] R. Brown, P. Godfrey, B. Kleibomer, *et al.*, *Journal of Molecular Spectroscopy*, vol. 142, no. 2, pp. 195–204, 1990.

[291] M. Melosso, A. Melli, L. Spada, *et al.*, *The Journal of Physical Chemistry A*, vol. 124, no. 7, pp. 1372–1381, 2020.

[292] N. Sato, Y. Hamada, and M. Tsuboi, *Spectrochimica Acta Part A: Molecular Spectroscopy*, vol. 43, no. 7, pp. 943–954, 1987.

[293] L. B. De Carvalho, A. A. Da Costa, M. L. Duarte, and J. Teixeira-Dias, *Spectrochimica Acta Part A: Molecular Spectroscopy*, vol. 44, no. 7, pp. 723–732, 1988.

[294] J. R. Durig, I. D. Darkhalil, J. J. Klaassen, *et al.*, *Journal of Raman Spectroscopy*, vol. 43, no. 9, pp. 1329–1336, 2012.

[295] J. Durig, G. Guirgis, and D. Compton, *Journal of Physical Chemistry*, vol. 83, no. 10, pp. 1313–1323, 1979.

[296] J. R. Durig, J. J. Klaassen, I. D. Darkhalil, *et al.*, *Journal of Molecular Structure*, vol. 1009, pp. 30–41, 2012.

[297] P. Palmieri and A. Mirri, *Journal of Molecular Structure*, vol. 37, no. 1, pp. 164–167, 1977.

[298] A. Verma and H. Bernstein, *Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics*, vol. 69, pp. 1586–1589, 1973.

[299] Z. Kisiel, O. Dorosh, A. Maeda, *et al.*, *Physical Chemistry Chemical Physics*, vol. 12, no. 29, pp. 8329–8339, 2010.

[300] J. Pearson, K. Sastry, E. Herbst, and F. C. De Lucia, *Journal of Molecular Spectroscopy*, vol. 175, no. 2, pp. 246–261, 1996.

[301] Z. Kisiel, "Assignment and analysis of complex rotational spectra," in *Spectroscopy from Space* (J. Demaison, K. Sarka, and E. A. Cohen, eds.), vol. 20, pp. 91–106, Springer Science & Business Media, 2001.

[302] J. G. de la Concepción, I. Jiménez-Serra, J. C. Corchado, *et al.*, *The Astrophysical Journal Letters*, vol. 912, no. 1, p. L6, 2021.

[303] J. K. Watson, *The Journal of Chemical Physics*, vol. 98, no. 7, pp. 5302–5309, 1993.

[304] F. Hegelund, J. Duncan, and D. McKean, *Journal of Molecular Spectroscopy*, vol. 65, no. 3, pp. 366–378, 1977.

[305] H. B. Andersen, F. Hegelund, N. R. Zangenberg, and F. Winther, *Journal of Raman Spectroscopy*, vol. 6, no. 5, pp. 238–250, 1977.

[306] J. Plíva and C. Martin, *Journal of Molecular Spectroscopy*, vol. 91, no. 1, pp. 218–237, 1982.

[307] J. Pliva, K. Rousan, A. Merchan, and S. Polo, *Journal of Molecular Spectroscopy*, vol. 99, no. 1, pp. 180–202, 1983.

[308] A. G. Maki and R. A. Toth, *Journal of Molecular Spectroscopy*, vol. 17, no. 1, pp. 136–155, 1965.

[309] A. Melli, V. Barone, and C. Puzzarini, *The Journal of Physical Chemistry A*, vol. 125, no. 14, pp. 2989–2998, 2021.

[310] E. Gougoula, D. J. Cole, and N. R. Walker, *The Journal of Physical Chemistry A*, vol. 124, no. 13, pp. 2649–2659, 2020.

[311] J. Chen, Y. Zheng, A. Melli, *et al.*, *Physical Chemistry Chemical Physics*, vol. 22, no. 9, pp. 5024–5032, 2020.

[312] M. Melosso, L. Bizzocchi, A. Adamczyk, *et al.*, *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 254, p. 107221, 2020.