SCUOLA NORMALE SUPERIORE DI PISA

# GAS-PHASE FORMATION OF COMPLEX ORGANIC MOLECULES IN INTERSTELLAR MEDIUM: COMPUTATIONAL INVESTIGATIONS

by Fanny Vazart

under the supervision of

Prof. Vincenzo Barone & Dr. Dimitrios Skouteris

A thesis presented for the degree of

*Doctor of Philosophy*

in Chemistry

2017

# *Abstract*

In the field of astro- and prebiotic chemistry, the building blocks of life, which are molecules composed of more than 6 atoms, are called Complex Organic Molecules (COMs). Their appearances on the early inorganic Earth is therefore one of the major issues faced by researchers interested in the origin of life. In this thesis, split into three parts, the main purpose is to show how different COMs are formed in interstellar medium (ISM), using computational chemistry.

The first part focuses mainly on preliminary studies aiming at evaluating the appropriate level of theory to use to perform studies of formation reactions. First, a comprehensive benchmark of C≡N stretching vibrations computed at harmonic and anharmonic levels is reported with the goal of proposing and validating a reliable computational strategy to get accurate results for this puzzling vibrational mode, involved in biological molcules, without any *ad hoc* scaling factor. Anharmonic calculations employing second-order vibrational perturbation theory provide very good results when performed using the B2PLYP double-hybrid functional, in conjunction with an extended basis set and supplemented by semiempirical dispersion contributions. For larger systems, B2PLYP harmonic frequencies, together with B3LYP anharmonic corrections, offer a very good compromise between accuracy and computational cost without the need of any empirical scaling factor. As a second step, transition metal complexes are considered. Even though they have not been of primary importance in ISM so far, it is always useful to investigate a various range of compounds. Here, a quantum mechanical investigation on the luminescence properties of several mono- and dinuclear platinum(II) complexes and on an iridium(III) ($d^6$ metal) complex is exhibited. The electronic structures and geometric parameters are briefly analyzed together with the absorption bands of all complexes. In all cases agreement with experiment is remarkable. Emission (phosphorescence) spectra from the first triplet states have also been investigated by comparing different computational approaches and taking into account also vibronic effects. Once again, agreement with experiment is good, especially using unrestricted electronic computations coupled to vibronic contributions. Together with the intrinsic interest of the results, the robustness and generality of the approach open the

opportunity for computationally oriented chemists to provide accurate results for the screening of large targets which could be of interest in molecular materials design.

In the second part, new insights into the formation of several interstellar species of great relevance in prebiotic chemistry are provided by electronic structure and kinetic calculations. Cyanomethanimine isomers, formamide, glycolaldehyde and acetic acid are considered. (*i*) Concerning the first one, a full thermodynamic and vibrational investigation of C-cyanomethanimine isomers rooted into the Density Functional Theory (DFT) and the second-order vibrational perturbation theory (VPT2) is reported first. It is shown that an anharmonic treatment affects dramatically the vibrational behavior of the molecules, especially thanks to the inclusion of interaction terms between the various modes. Furthermore, the equilibrium constant between the isomers, as well as the rate constant, have been obtained at both harmonic and anharmonic levels showing, as expected, slight but non-negligible differences. Then, in order to solve the formation issue, the reaction CN + $CH_2$=NH has been envisaged. This reaction is a facile formation route of *Z*,*E*-C-cyanomethanimine, even under the extreme conditions of density and temperature typical of cold interstellar clouds. *E*-C-cyanomethanimine has been recently identified in Sgr B2(N) in the Green Bank Telescope (GBT) PRIMOS survey and no efficient formation routes had been envisaged so far. The rate coefficients for the reaction channel show that the *E*-C-cyanomethanimine should be formed with a slightly lower yield than *Z*-C-cyanomethanimine. As the detection of *E*-isomer is favored due to its larger dipole moment, the missing detection of the *Z*-isomer can be due to the sensitivity limit of the GBT PRIMOS survey and the detection of the *Z*-isomer should be attempted with more sensitive instrumentation. The CN + $CH_2$=NH reaction can also play a role in the chemistry of the upper atmosphere of Titan where the cyanomethanimine products can contribute to the buildup of the observed nitrogenrich organic aerosols that cover the moon. (*ii*) Concerning formamide, several pathways of the OH + $CH_2$=NH and $NH_2$ + $H_2CO$ reaction channels have been investigated. The results show that both reaction channels are essentially barrierless (in the sense that all relevant transition states lie below or only marginally above the reactants) and once tunneling is taken into proper account indicate that the reaction can occur under the low temperature conditions of interstellar environments. However, the OH + $CH_2$=NH reaction tends to lead preferentially

to other products of prebiotic interest. Molecular deuteration has proven to be, in other cases, an efficient way to validate a synthesis route for molecules. For formamide, new published observations show that its three deuterated forms have all the same deuteration ratio. Following the work on the gas-phase formamide formation via the reaction $NH_2$ + $H_2CO$, new calculations of the rate cofficients for the production of monodeuterated formamide through the same reaction are presented, starting from monodeuterated $NH_2$ or $H_2CO$. The results of this new computations show that, at the 100 K temperature of the spot of interest, the rate of deuteration of the three forms is the same, within 20%. On the contrary, the reaction between non-deuterated species proceeds three times faster than that with deuterated ones. These results confirm that a gas-phase route for the formation of formamide is perfectly in agreement with the available observations. In order to further corroborate the results, new interferometric observations of formamide towards a recently shocked region near a young solar-type protostar is also reported. The new images of the molecular line emission show the spatial segregation of formamide with respect to other organic species. These observations, coupled with a chemical modelling analysis, provide new evidence that this potentially crucial brick of life can be efficiently formed in gas-phase around Sun-like protostars. (*iii*) Concerning glycolaldehyde and acetic acid, the precursors $O(^3P)$ and hydroxyethyl radicals were considered ($^{\bullet}CH_2CH_2OH$ for glycolaldehyde and $CH_3{}^{\bullet}CHOH$ for acetic acid). Two reaction paths were obtained and the resulting rate constants show not only that they are viable in ISM, but also that when included in an astrochemical model, the obtained abundances for glycolaldehyde also match well the observed ones. In the case of acetic acid, the predicted abundances are too high compared to the observed ones.

To finish, the third part brings several miscellaneous studies together. The first one consists in a quick investigation on the glutamine rotamers, including the interconversion paths from one to the another. As far as it is concerned, the second one focuses on esterification in gas-phase. By mixing primary and secondary alcohols with carboxylic acids just before the supersonic expansion within pulsed Fourier transform microwave experiments, only the rotational spectrum of the ester was observed. However, when formic acid was mixed with tertiary alcohols, adducts were formed and their rotational spectra could be easily measured. Quantum mechanical calculations were therefore performed to

interpret the experimental evidence and the results are exhibited here. Finally, in the last study, the 1:1 complexes of ammonia with pyridine (*i*) and of tert-butyl alcohol with difluoromethane (*ii*) have been characterized by using state-of-the-art quantum-chemical computations combined with pulsed-jet Fourier-Transform microwave spectroscopy. (*i*) The computed potential energy landscape pointed out the formation of a stable σ-type complex, which has been confirmed experimentally: the analysis of the rotational spectrum showed the presence of only one 1:1 pyridine – ammonia adduct. The rotational spectrum provided the proof that the two molecules form a σ-type complex and that they are held together by both a N-H···N and a CH···N bond. This work represents the first application of an accurate and yet efficient computational scheme, designed for the investigation of small biomolecules, to a molecular cluster. Among the results obtained, the dissociation energy (BSSE and ZPE corrected) has been estimated to be 10.9 $kJ \cdot mol^{-1}$. (*ii*) The computed potential energy landscape pointed out the formation of three stable isomers. However, the very low interconversion barriers explain why only one isomer, showing one O-H···F and two C-H···O weak hydrogen bonds, has been experimentally characterized.

## *Acknowledgements*

I would like to thank all the people who contributed in some way to the work described in this thesis. First and foremost, I thank my academic advisor, Professor Vincenzo Barone, for accepting me into his group. During my tenure, he contributed to a rewarding PhD experience by giving me intellectual freedom in my work together with support and advice, supporting my attendance at various conferences and demanding a high quality of work in all my endeavors. Additionally, I would like to thank my committee members for their interest in my work.

Every result described in this thesis was accomplished with the help and support of fellow labmates and collaborators. Dr. Camille Latouche and I worked together on several different projects, particularly treating transition metal complexes, and without his efforts and pieces of advice, my job would have undoubtedly been more difficult. I greatly benefited from his keen scientific insight and his ability to introduce me to practical work. I was also fortunate to have the chance to work with Dr. Dimitrios Skouteris, who was the first person to suggest I work on astrochemistry. He was an extremely reliable source of practical scientific knowledge and I am grateful for his collaboration, particularly regarding the kinetics calculations described in this thesis. He also gave me the opportunity to meet Professors Nadia Balucani and Cecilia Ceccarelli whose knowledge in astronomy and astrochemistry were essential, and whose warm personalities were refreshing. They allowed me to attend a highly profitable and fascinating astrochemistry summer school and I am grateful for it. Last but not least, I would like to thank Professor Cristina Puzzarini for her help and collaboration in performing high-level computations and experiments related to the projects exhibited in this thesis.

Additionally, I would like to thank the various members of the DREAMS lab chemistry department with whom I had the opportunity to work and have not already mentioned: Dr. Nicola Tasinato, Dr. Serena Manti, Dr. Danilo Calderini, Dr. Julien Bloino and Nicola Fusè. They provided a friendly and cooperative atmosphere at work and also useful feedback and insightful comments on my work. I also had the opportunity to help a fellow PhD student, Dr. Ilaria Domenichelli, with some computations related to her project. I would be ingrate if I did not thank Monica Sanna, who deserves credit for providing much needed assistance with

administrative tasks and keeping our work running smoothly. Additionally, the Avogadro staff provided immediate support for any computer problems we encountered.

On another hand, I would like to acknowledge the Scuola Normale Superiore. My PhD experience benefitted greatly from the high-level courses I took and the high-quality seminars that the department organized.

Finally, I would like to acknowledge friends and family who supported me during my time here. First of all, I would like to thank David for his constant love and support, even when I had no time for him. I am also forever indebted to my parents for giving me the opportunities and experiences that have made me who I am. They selflessly encouraged me to explore new directions in life and this journey would not have been possible if not for them.

# *Contents*

# *List of acronyms*

| | |
|---|---|
| COM | Complex Organic Molecule |
| DFT | Density Functional Theory |
| FA | Formic Acid |
| ISM | InterStellar Medium |
| PES | Potential Energy Surface |
| PCM | Polarizable Continuum Model |
| TD-DFT | Time-Dependent Density Functional Theory |
| TS | Transition State |

# *Introduction*

How did life appear on an originally inorganic Earth? This is the main question which is faced by prebiotic chemists and astrochemists. As a rocky planet, Earth is mostly composed of rocks and metals and at its early stages, no organic compound could be found at its surface. However, between 3.8 and 4.1 billion years ago, life arose. In order for it to appear, the presence of water was mandatory. This explains why its origin on Earth has been widely investigated.[1–4] Two main scenarios emerged from these studies: an extraterrestrial origin hypothesis and an out-gassing scenario. The first one would consist in an arrival of water on Earth carried by comets or asteroids while the second implies that water was already present inside Earth during its formation and was released by volcanic activity. In the field of prebiotic chemistry and astrochemistry, the simplest organic building blocks for life, such as amino or nucleic acids and their possible precursors, are called Complex Organic Molecules (COMs) and usually contain up to around 10 atoms. Considering that many of these species have been detected in the interstellar medium (ISM), the most widespread hypothesis to explain how they were brought to Earth involves a transportation of these compounds on comets or asteroids,[5–8] like in the first scenario for the origin of water. The Rosetta spacecraft was recently able to detect some COMs in the dust surrounding its target comet (67p), which strengthens the argument that the building blocks of life itself may have come from icy space rocks.[9] As an example of these COMs, one could cite formamide, the simplest amide, which can be considered a precursor in the abiotic amino acid synthesis and perhaps also that of nucleic acid bases.[10–12] It can be therefore considered as a central compound that could connect both metabolism –conversion of energy, ruled by proteins– and genetics –passage of information, ruled by DNA and RNA–. Glycolaldehyde and glycine are also of major importance as the first detected interstellar sugar molecule and amino acid, respectively. Indeed, as glycolaldehyde has been detected in ISM,[13] glycine has been found in comets and asteroids.[14]

The question of their formation is thus a burning issue. As a matter of fact, two main theories are now discussed to explain how chemical reactions can occur at very low densities –pressure– and temperature environments: dust grain and gas phase chemistry. The first one relies on surface reaction mechanisms and can allow to lower some reaction barriers –which

could have kept the reaction from happening at very low temperatures– and to ease the encounter between reactants –which can be difficult at very low pressures–. However, this theory faces some major issues, such as the adsorption/desorption mechanisms. Indeed, if we take a look at the formation of methanol for example, it is commonly believed that it occurs on the surface of dust grains.[15,16] Yet, this molecule was detected also in gas phase in several cold and dark zones where no thermal desorption or photodissociation can follow.[17,18] This problem led to doubt this surface formation mechanism. In order to avoid those problems, one can consider the second theory: gas-phase reaction routes. In that case, there is not any desorption to consider but the proposed reaction mechanisms cannot involve any barrier and the only available energy is the one issuing from the reactants. Furthermore, no three body collision can be envisaged. This work will focus exclusively on gas-phase reactions.

The formation of COMs in ISM is a very complicated topic that requires three different types of expertise: observation, experiments and theoretical calculations. Astronomers, experimental and theoretical chemists should therefore work together in order to match the different locations where some compounds can be detected with the possible formation mechanisms. Another obstacle that can arise is the difficulty of reproducing experimentally the harsh conditions of the interstellar medium; accordingly, theoretical calculations are mandatory. It is then possible, once the reaction channels issuing from reactants are determined, to calculate as well the reaction rates for the formation of the products depending on the temperature. With these reaction rates, some astrochemical models are set up in order to obtain calculated abundances of the products, which are then feasibly compared with measured abundances in zones that can exhibit different temperatures.

## Kinetics and transition state theory

Chemical kinetics is the study of rates of chemical processes. These reaction rates for a reactant or a product can explain how quickly a reaction takes place. They strongly depend on several factors, including temperature, pressure and concentration.

When considering the basic reaction:

$$aA + bB \rightarrow cC + dD$$

The reaction rate can be defined as:

$$r = -\frac{1}{a}\frac{d[A]}{dt} = -\frac{1}{b}\frac{d[B]}{dt} = \frac{1}{c}\frac{d[C]}{dt} = \frac{1}{d}\frac{d[D]}{dt} \quad (1)$$

where [X] stands for the concentration of X.

It is also possible to consider the rate equation, a mathematical expression that links the rate of a reaction to the concentrations of the reactants, of the form:

$$r = k(T)[A]^n[B]^m \quad (2)$$

here, $n$ and $m$ are the reaction orders, which can be different from the stoichiometric coefficients and depend on the reaction mechanism. $k(T)$, as far as it is concerned, is the reaction rate coefficient, or rate constant. It is not really a constant because it depends on the temperature, often described by the Arrhenius equation:

$$k(T) = Ae^{-\frac{E_a}{RT}} \quad (3)$$

where $A$ is a pre-exponential factor, $R$ is the gas constant and $E_a$ the activation energy of the reaction.

In order to understand the physical meaning of $E_a$, one should consider the transition state theory. This theory is based on three basic ideas:

*(i)* Rates of reaction can be studied by examining activated complexes near the saddle point of a potential energy surface, the details of how these complexes are formed not being important. The saddle point is called the transition state (TS).

*(ii)* The activated complex is in quasi-equilibrium with the reactant molecules.

*(iii)* The activated complex can convert into products, and kinetic theory can be used to calculate the rate of this conversion.

In the particular case exhibited in Figure 1a, this single-step reaction is exothermic since the energy of the products is lower than the energy of the reactants. The endothermic reaction (products higher in energy than reactants) could have been considered as well. In both cases, the TS is always higher in energy than both reactants and products. One can also consider multi-step reactions, such as shown in Figure 1b. Here, all TSs are more energetic than both minima (reactants, reaction intermediates, products) they link.

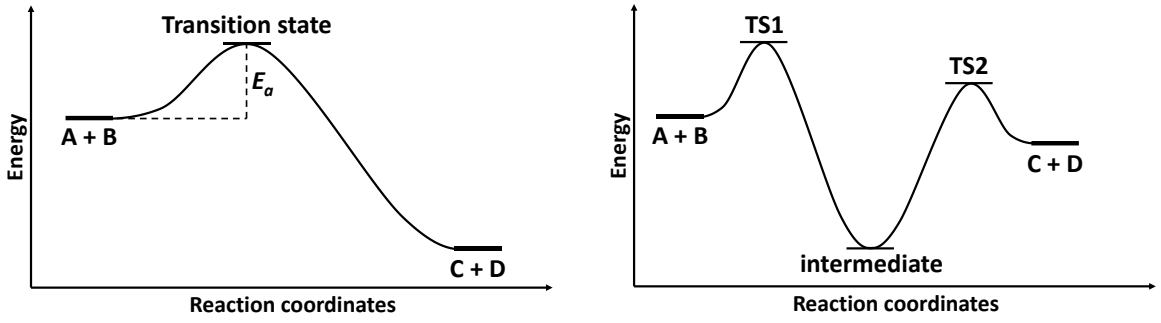



**Figure 1.** Basic reaction coordinates diagrams.

## Theoretical details

Most of the calculations presented here are based on the Density Functional Theory (DFT).

### *The Schrödinger equation*

In quantum chemistry, the ultimate goal is to solve the time-independent non-relativistic Schrödinger equation:

$$\hat{H}\Psi_i\left(\overrightarrow{x_1}, \overrightarrow{x_2}, \dots, \overrightarrow{x_N}, \overrightarrow{R_1}, \overrightarrow{R_2}, \dots, \overrightarrow{R_M}\right) = E_i\Psi_i(\overrightarrow{x_1}, \overrightarrow{x_2}, \dots, \overrightarrow{x_N}, \overrightarrow{R_1}, \overrightarrow{R_2}, \dots, \overrightarrow{R_M}) \quad (1)$$

$\Psi$ being the wave function and $\hat{H}$ the Hamiltonian for a system that consists in $M$ nuclei and $N$ electrons. Starting from here, atomic units are used.

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \frac{1}{2}\sum_{A=1}^{M}\frac{1}{M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_E}{R_{AB}} \quad (2)$$

here, the first two terms describe the kinetic energy of the electrons and nuclei while the three following terms represent the attractive electrostatic interaction between the nuclei and the electrons, the repulsive potential due to the electron-electron and nucleus-nucleus interactions.

In order to simplify Eq. 2, one can use the Born-Oppenheimer approximation that considers the electrons moving in the field of fixed nuclei. It is based on the fact that the masses of electrons are much lower than the masses of nuclei, which leads the latter to move much slower than the former. Thanks to this approximation, the nuclear kinetic energy is

zero and the potential energy of nuclei is merely a constant, which reduces the electronic Hamiltonian to:

$$\hat{H}_{elec} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N}\frac{1}{r_{ij}} = \hat{T} + \hat{V}_{Ne} + \hat{V}_{ee} \qquad (3)$$

The solution of the Schrödinger equation using $\hat{H}_{elec}$ is the electronic wave function $\Psi_{elec}$ and the electronic energy $E_{elec}$. The total energy $E_{tot}$ is then the sum of $E_{elec}$ and $E_{nuc}$, the constant nuclear repulsion term.

$$\hat{H}_{elec}\Psi_{elec} = E_{elec}\Psi_{elec} \qquad (4)$$

$$E_{tot} = E_{elec} + E_{nuc} \qquad \text{where} \qquad E_{nuc} = \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \qquad (5)$$

### *The variational principle for the ground state*

When a system is in the state $\Psi$, the expectation value of the energy is given by:

$$E[\Psi] = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} \qquad \text{where} \qquad \langle\Psi|\hat{H}|\Psi\rangle = \int \Psi^* \hat{H}\Psi d\vec{x} \qquad (6)$$

The variational principle states that the energy computed from a guessed $\Psi$ is an upper bound to the true ground-state energy $E_0$. Full minimization of the functional $E[\Psi]$ with respect to all allowed $N$-electron wave functions will give the true ground state $\Psi_0$ and energy $E[\Psi_0] = E_0$, that is:

$$E_0 = min_{\Psi\to N}E[\Psi] = min_{\Psi\to N}\langle\Psi|\hat{T} + \hat{V}_{Ne} + \hat{V}_{ee}|\Psi\rangle \qquad (7)$$

For a system of $N$ electrons and given nuclear potential $V_{ext}$, the variational principle defines a procedure to determine the ground-state wave function $\Psi_0$, the ground-state energy $E_0[N,V_{ext}]$, and other properties of interest. It needs to be noted that $\Psi_0$ must be antisymmetric with respect to electrons. In other words, the ground state energy is a functional of the number of electrons $N$ and the nuclear potential $V_{ext}$:

$$E_0 = E[N, V_{ext}] \qquad (8)$$

### *The Hartree-Fock approximation*

Supposing that $\Psi_0$ is approximated as an antisymmetrized product of N orthonormal spin orbitals $\psi_i(\vec{x})$, each a product of a spatial orbital $\Phi_k(\vec{r})$ and a spin function σ(s) = α(s) or β(s), the following Slater determinant is obtained:

$$\Psi_0 \approx \Psi_{HF} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\vec{x}_1) & \psi_2(\vec{x}_1) & \dots & \psi_N(\vec{x}_1) \\ \psi_1(\vec{x}_2) & \psi_2(\vec{x}_2) & \dots & \psi_N(\vec{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\vec{x}_N) & \psi_2(\vec{x}_N) & \dots & \psi_N(\vec{x}_N) \end{vmatrix} \quad (9)$$

The Hartree-Fock approximation is the method whereby the orthogonal orbitals $\psi_i$ are searched to minimize the energy for this determinantal form of $\Psi_0$:

$$E_{HF} = min_{E_{HF} \to N} E[\Psi_{HF}] \quad (10)$$

The expected value of energy $E_{HF}$ using the Hamiltonian operator is given by:

$$E_{HF} = \langle \Psi_{HF} | \hat{H} | \Psi_{HF} \rangle = \sum_{i=1}^{N} H_i + \frac{1}{2} \sum_{i,j=1}^{N} (J_{ij} + K_{ij}) \quad (11)$$

with:

$$H_i = \int \psi_i^*(\vec{x}) \left[ -\frac{1}{2} \nabla^2 + V_{ext}(\vec{x}) \right] \psi_i(\vec{x}) d\vec{x} \quad (12)$$

which defines the contribution due to the kinetic energy and the electron-nucleus attraction.

$$J_{ij} = \int \int \psi_i(\overrightarrow{x_1}) \psi_i^*(\overrightarrow{x_1}) \frac{1}{r_{12}} \psi_j^*(\overrightarrow{x_2}) \psi_j(\overrightarrow{x_2}) d\overrightarrow{x_1} d\overrightarrow{x_2} \quad (13)$$

are the Coulomb integrals and

$$K_{ij} = \int \int \psi_i^*(\overrightarrow{x_1}) \psi_j^*(\overrightarrow{x_1}) \frac{1}{r_{12}} \psi_i(\overrightarrow{x_2}) \psi_j(\overrightarrow{x_2}) d\overrightarrow{x_1} d\overrightarrow{x_2} \quad (14)$$

are the exchange integrals.

$$J_{ij} \geq K_{ij} \geq 0 \text{ and } J_{ii} = K_{ii}$$

The minimization of the energy functional of Eq. 11, together with the normalization conditions $\int \psi_i^*(\vec{x}) \psi_j(\vec{x}) \, d\vec{x} = \delta_{ij}$ lead to the Hartree-Fock differential equations:

$$\hat{f}\psi_i = \epsilon_i\psi_i \text{ , } i\text{=1, 2, ..., } N \quad (15)$$

The Fock operator $\hat{f}$ is an effective one-electron operator defined as:

$$\hat{f} = -\frac{1}{2}\nabla_i^2 - \sum_A^M \frac{Z_A}{r_{iA}} + \hat{V}_{HF}(i) \quad (16)$$

The first two terms represent the kinetic energy and the potential energy due to the electron-nucleus attraction. $\hat{V}_{HF}(i)$, as far as it is concerned, is the Hartree-Fock potential operator, the average repulsive potential experienced by the $i^{th}$ electron due to the remaining $N$-1 electrons, and is given by:

$$\hat{V}_{HF}(\overrightarrow{x_1}) = \sum_j^N (\hat{J}_j(\overrightarrow{x_1}) - \hat{K}_j(\overrightarrow{x_1})) \quad (17)$$

with the Coulomb operator $\hat{J}$ that represents the potential an electron at position $\vec{x}_1$experiences due to the average charge distribution of another electron in spin orbital $\psi_j$, defined as:

$$\hat{J}_j(\overrightarrow{x_1}) = \int \left|\psi_j(\overrightarrow{x_2})\right|^2 \frac{1}{r_{12}} d\overrightarrow{x_2} \quad (18)$$

and $\hat{K}$ that is the exchange contribution to the HF potential and is defined through its effect when operating on a spin orbital:

$$\hat{K}_j(\overrightarrow{x_1})\psi_i(\overrightarrow{x_1}) = \int \psi_j^*(\overrightarrow{x_2}) \frac{1}{r_{12}} \psi_i(\overrightarrow{x_2}) dx_2 \psi_j\,(\overrightarrow{x_1}) \quad (19)$$

The HF potential is non-local and depends on the spin orbital. Therefore, the HF equations must be solved self-consistently. Moreover, the Koopman's theorem provides a physical interpretation of the orbital energies: it states that the orbital energy $\varepsilon_i$ is an approximation of minus the ionization energy associated with the removal of an electron from the orbital $\psi_i$, *i.e.* $\epsilon_i \approx E_N - E_{N-1}^i = -IE(i)$.

## *The electron density*

The electron density $\rho(\vec{r})$ is a central property in DFT and is defined as the integral over the spin coordinates of all electrons and over all but one spatial variables. It determines the probability of finding an electron (among the N) within the volume element $d\vec{r}$.

$$\rho(\vec{r}) = N \int \dots \int |\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)|^2 ds_1 d\overrightarrow{x_2} \dots d\overrightarrow{x_N} \quad (20)$$

$\rho(\vec{r})$ is a non-negative function of only the three spatial variables (as opposite to $\Psi$, that depends on 4N variables) that vanishes at infinity and integrates to the total number of electrons:

$$lim_{\vec{r}\to\infty}\rho(\vec{r}) = 0 \qquad \int \rho(\vec{r})d\vec{r} = N \quad (21)$$

At any position of an atom, the gradient of $\rho(\vec{r})$ has a discontinuity and a cusp results:

$$lim_{r_iA\to 0}[\nabla_r + 2Z_A]\bar{\rho}(\vec{r}) = 0 \quad (22)$$

where $Z$ is the nuclear charge and $\bar{\rho}(\vec{r})$ the spherical average of $\rho(\vec{r})$.

The asymptotic exponential decay for large distance from all nuclei leads to this approximation:

$$\rho(\vec{r}) \sim \exp[-2\sqrt{2I}|\vec{r}| \qquad \text{where I is the exact ionization energy} \quad (23)$$

### *The Thomas-Fermi model*

This model can be considered the first density functional theory. It is based on the uniform electron gas and uses the following functional for the kinetic energy:

$$T_{TF}[\rho(\vec{r})] = \frac{3}{10}(3\pi^2)^{2/3} \int \rho^{5/3}(\vec{r})d\vec{r} \quad (24)$$

The approximate energy of an atom is then obtained using the classical expression for the nuclear-nuclear potential and the electron-electron potential:

$$E_{TF}[\rho(\vec{r})] = \frac{3}{10}(3\pi^2)^{2/3} \int \rho^{5/3}(\vec{r})d\vec{r} - Z\int \frac{\rho(\vec{r})}{r} d\vec{r} + \frac{1}{2}\int\int \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{r_{12}} d\vec{r}_1 d\vec{r}_2 \quad (25)$$

This energy is completely given in terms of electron density.

In order to determine the correct density to be included in Eq. 25, a variational principle is employed. The ground state of the system is assumed to be connected to the $\rho(\vec{r})$ for which the energy is minimized under the constraint of $\int \rho(\vec{r})d\vec{r} = N$. However, this variational principle is not necessarily justified.

### *The first Hohenberg-Kohn theorem*

This theorem demonstrates that the electron density uniquely determines the Hamiltonian operator and thus all the properties of the system. It states that the external potential $V_{ext}(\vec{r})$ is (to within an additive constant) a unique functional of $\rho(\vec{r})$; since $V_{ext}(\vec{r})$ fixes $\widehat{H}$, the full many particle ground state is then a unique functional of $\rho(\vec{r})$.

*Proof*: if we assume that there were two external potentials $V_{ext}(\vec{r})$ and $V'_{ext}(\vec{r})$ differing by more than a constant, each giving the same $\rho(\vec{r})$ for its ground state, we would have two Hamiltonians $\widehat{H}$ and $\widehat{H}'$ whose ground-state densities were the same although the normalized wave functions $\Psi$ and $\Psi'$ would be different. Taking $\Psi'$ as a trial wave function for the $\widehat{H}$ problem, we get:

$$\begin{aligned} E_0 < \langle \Psi' | \widehat{H} | \Psi' \rangle &= \langle \Psi' | \widehat{H'} | \Psi' \rangle + \langle \Psi' | \widehat{H} - \widehat{H'} | \Psi' \rangle \\ &= E'_0 + \int \rho(\vec{r})[V_{ext}(\vec{r}) - V'_{ext}(\vec{r})]d\vec{r} \end{aligned} \tag{26}$$

where $E_0$ and $E'_0$ are the ground-state energies for $\widehat{H}$ and $\widehat{H}'$, respectively. Similarly, taking $\Psi$ as a trial wave function for the $\widehat{H'}$ problem, we get:

$$\begin{aligned} E'_0 < \langle \Psi | \widehat{H'} | \Psi \rangle &= \langle \Psi | \widehat{H} | \Psi \rangle + \langle \Psi | \widehat{H'} - \widehat{H} | \Psi \rangle \\ &= E_0 + \int \rho(\vec{r})[V_{ext}(\vec{r}) - V'_{ext}(\vec{r})]d\vec{r} \end{aligned} \tag{27}$$

Adding Eq. 26 and 27, we get $E_0 + E'_0 < E'_0 + E_0$, which is a contradiction. Therefore, there cannot be two different $V_{ext}(\vec{r})$ that give the same $\rho(\vec{r})$ for their ground states.
Thus, $\rho(\vec{r})$ determines $N$ and $V_{ext}(\vec{r})$ and hence all the properties of the system, for example the kinetic energy $T[\rho]$, the potential energy $V[\rho]$ and the total energy $E[\rho]$. Now, the total energy can be written as:

$$E[\rho] = E_{Ne}[\rho] + T[\rho] + E_{ee}[\rho] = \int \rho(r) V_{Ne}(r) dr + F_{HK}[\rho] \tag{28}$$

$$F_{FK}[\rho] = T[\rho] + E_{ee} \tag{29}$$

This functional $F_{FK}[\rho]$ is the holy grail of density functional theory. If it were known, the Schrödinger equation could be solved exactly. And, since it is a universal functional completely independent of the system at hand, it applies equally well to the hydrogen atom

as to gigantic molecules. $F_{FK}[\rho]$ contains the functional for the kinetic energy $T[\rho]$ and that for the electron-electron interaction, $E_{ee}[\rho]$. The explicit form of both these functional is completely unknown. However, from the latter it is possible to extract at least the classical part $J[\rho]$:

$$E_{ee}[\rho] = \frac{1}{2}\int\int \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{r_{12}} d\vec{r}_1 d\vec{r}_2 + E_{ncl} = J[\rho] + E_{ncl}[\rho] \quad (30)$$

$E_{ncl}$ is the non-classical contribution to the electron-electron interaction: self-interaction correction, exchange and Coulomb correlation.

The explicit forms of the functionals $T[\rho]$ and $E_{ncl}[\rho]$ is the major challenge of DFT.

### *The second Hohenberg-Kohn theorem*

The second H-K theorem states that $F_{FK}[\rho]$, the functional that delivers the ground state energy of the system, delivers the lowest energy if and only if the input density is the true ground state density. This is also the variational principle:

$$E_0 \leq E[\tilde{\rho}] = T[\tilde{\rho}] + E_{Ne}[\tilde{\rho}] + E_{ee}[\tilde{\rho}] \quad (31)$$

In other words, for any trial density $\tilde{\rho}(\vec{r})$, which satisfies the necessary boundary conditions such as $\tilde{\rho}(\vec{r}) \geq 0$, $\int \tilde{\rho}(\vec{r})d\vec{r} = N$, and which is associated with some external potential $\tilde{V}_{ext}$, the energy obtained from the functional of Eq. 28 represents an upper bound to the true ground state energy $E_0$.

*Proof*: it makes use of the variational principle established for wave functions. We recall that any trial density $\tilde{\rho}$ defines its own Hamiltonian $\hat{H}$ and hence its own wave function $\tilde{\Psi}$. This wave function can now be taken as the trial wave function for the Hamiltonian generated from the true external potential $V_{ext}$. Thus,

$$\langle\tilde{\Psi}|\hat{H}|\tilde{\Psi}\rangle = T[\tilde{\rho}] + E_{ee}[\tilde{\rho}] + \int \tilde{\rho}(\vec{r})V_{ext}d\vec{r} = E[\tilde{\rho}] \geq E_0[\tilde{\rho}] = \langle\tilde{\Psi}_0|\hat{H}|\tilde{\Psi}_0\rangle \quad (32)$$

### *The Kohn-Sham equations*

The ground state energy of a system can be written as:

$$E_0 = min_{\rho\to N}(F[\rho] + \int \rho(\vec{r})V_{Ne}d\vec{r}) \quad (33)$$

where the universal functional $F[\rho]$ contains the contributions of the kinetic energy, the classical Coulomb interaction and the non-classical portion:

$$F[\rho] = T[\rho] + J[\rho] + E_{ncl}[\rho] \quad (34)$$

Here, only $J[\rho]$ is known. The main problem is to find the expressions for $T[\rho]$ and $E_{ncl}[\rho]$.

The Thomas-Fermi model described above provides an example of density functional theory. However, its performance is really low due to the poor approximation of the kinetic energy. To solve this problem, Kohn and Sham proposed an alternative approach.

They suggested to calculate the exact kinetic energy of a non-interacting reference system with the same density as the real, interacting one:

$$T_S = -\frac{1}{2}\sum_i^N \langle\psi_i|\nabla^2|\psi_i\rangle \qquad \rho_S(\vec{r}) = \sum_i^N \sum_s |\psi_i(\vec{r},s)| = \rho(\vec{r}) \quad (35)$$

where the $\psi_i$ are the orbitals of the non-interacting system. $T_S$ is not equal to the true kinetic energy of the system but Kohn and Sham accounted for that by introducing the following separation of the functional $F[\rho]$ :

$$F[\rho] = T_S[\rho] + J[\rho] + E_{XC}[\rho] \quad (36)$$

where $E_{XC}$, the so-called exchange-correlation energy is defined, thanks to Eq. 36, as:

$$E_{XC}[\rho] = (T[\rho] - T_S[\rho]) + (E_{ee}[\rho] - J[\rho]) \quad (37)$$

This $E_{XC}$ is the functional that contains everything that is unknown.

Now, the question is: how can we uniquely determine the orbitals in our non-interacting reference system? In other words, how can we define a potential $V_S$ such that it provides us with a Slater determinant which is characterized by the same density as our real system? In order to solve this problem, the expression for the energy of the interacting system in terms of the separation described in Eq. 36 is written:

$$E[\rho] = T_S[\rho] + J[\rho] + E_{XC}[\rho] + E_{Ne}[\rho] \quad (38)$$

$$E[\rho] = T_S[\rho] + \frac{1}{2}\int\int \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{r_{12}} d\vec{r}_1 d\vec{r}_2 + E_{XC}[\rho] + \int V_{Ne}\rho(\vec{r}) d\vec{r} =$$
$$-\frac{1}{2}\sum_i^N \langle\psi_i|\nabla^2|\psi_i\rangle$$
$$+\frac{1}{2}\sum_i^N\sum_j^N \int\int |\psi_i(\vec{r}_1)|^2 \frac{1}{r_{12}} |\psi_j(\vec{r}_2)|^2 d\vec{r}_1 d\vec{r}_2 + E_{XC}[\rho] \quad (39)$$
$$-\sum_i^N \int \sum_A^M \frac{Z_A}{r_{1A}} |\psi_i(\vec{r}_1)|^2 d\vec{r}_1$$

The only term for which no explicit form can be given is $E_{XC}$. It is possible to apply the variational principle and look for the conditions that the orbitals $\{\psi_i\}$ must fulfill in order to minimize this energy expression under the usual constraint $\langle\psi_i|\psi_j\rangle = \delta_{ij}$. The resulting equations are the Kohn-Sham equations:

$$\left(-\frac{1}{2}\nabla^2 + \left[\int \frac{\rho(\vec{r}_2)}{r_{12}} + V_{XC}(\vec{r}_1) - \sum_A^M \frac{Z_A}{r_{1A}}\right]\right)\psi_i = \left(-\frac{1}{2}\nabla^2 + V_S(\vec{r}_1)\right)\psi_i = \epsilon_i\psi_i \quad (40)$$

$$V_S(\vec{r}_1) = \int \frac{\rho(\vec{r}_2)}{r_{12}} d\vec{r}_2 + V_{XC}(\vec{r}_1) - \sum_A^M \frac{Z_A}{r_{1A}} \quad (41)$$

Once the various contributions in Eq. 40 and 41 are known, a grip is obtained on the potential $V_S$ which is needed to insert into the one-particle equations, which then determine the orbitals and hence the ground state density and the ground state energy employing Eq. 39. It is noted that $V_S$ depends on the density, and therefore the Kohn-Sham equations have to be solved iteratively. Moreover, the exchange-correlation potential, $V_{XC}$ is defined as the functional derivative of $E_{XC}$ with respect to ρ, i.e. $V_{XC} = \delta E_{XC}/\delta\rho$.

It is important to realize that if the exact forms of $E_{XC}$ and $V_{XC}$ were known, the Kohn-Sham strategy would lead to the exact energy.

Strictly speaking, the Kohn-Sham orbitals have no physical significance, except for the highest occupied orbital, $\varepsilon_{max}$, which equals the negative of the exact ionization energy.

### *The local density approximation (LDA)*

The LDA is the basis of all approximate exchange-correlation functionals. This model is based on the idea of a uniform electron gas. This is a system in which electrons move on a positive background charge distribution such that the total ensemble is neutral.
The central idea behind LDA is the assumption that we can write $E_{XC}$ in the form:

$$E_{XC}^{LDA}[\rho] = \int \rho(\vec{r})\epsilon_{XC}\big(\rho(\vec{r})\big)d\vec{r} \tag{42}$$

Here, $\epsilon_{XC}\big(\rho(\vec{r})\big)$ is the exchange-correlation energy per particle of a uniform electron gas of density $\rho(\vec{r})$. This energy per particle is weighted with the probability $\rho(\vec{r})$ that there is an electron at this position. The quantity $\epsilon_{XC}\big(\rho(\vec{r})\big)$ can be further split into exchange and correlation contributions,

$$\epsilon_{XC}\big(\rho(\vec{r})\big) = \epsilon_X\big(\rho(\vec{r})\big) + \epsilon_C\big(\rho(\vec{r})\big) \tag{43}$$

The exchange part, $\epsilon_X$,which represents the exchange energy of an electron in a uniform electron gas of a particular density, was originally derived by Bloch and Dirac:

$$\epsilon_X = -\frac{3}{4}\left(\frac{3\rho(\vec{r})}{\pi}\right)^{1/3} \tag{44}$$

No such explicit expression is known for the correlation part, $\epsilon_C$. However, highly accurate numerical quantum Monte-Carlo simulations of the homogeneous electron gas are available.[19]

The accuracy of the LDA for the exchange energy is typically within 10%, while the normally much smaller correlation energy is generally overestimated by up to a factor 2. The two errors typically cancel partially. Moreover, experience has shown that the LDA gives ionization energies of atoms, dissociation energies of molecules and cohesive energies with a fair accuracy of typically 10-20%. However, the LDA gives bond lengths of molecules and solids typically with an astonishing accuracy of ~2%. This moderate accuracy that LDA delivers is certainly insufficient for most applications in chemistry.

### *The generalized gradient approximation (GGA)*

The first logical step to go beyond LDA is the use of not only the information about the density $\rho(\vec{r})$ at a particular point $\vec{r}$, but to supplement the density with information about the gradient of the charge density, $\nabla\rho(\vec{r})$ in order to account for the non-homogeneity of the true electron density. Thus, the exchange-correlation energy can be written in the following form termed generalized gradient approximation (GGA):

$$E_{XC}^{GGA}[\rho_\alpha, \rho_\beta] = \int f(\rho_\alpha, \rho_\beta, \nabla\rho_\alpha, \nabla\rho_\beta) d\vec{r} \quad (45)$$

In another approach A. Becke introduced a successful hybrid functional:[20]

$$E_{XC}^{hyb} = \alpha E_X^{KS} + (1-\alpha) E_{XC}^{GGA} \quad (46)$$

where $E_X^{KS}$ is the exchange calculated with the exact KS wave function, $E_{XC}^{GGA}$ is an appropriate GGA, and α a fitting parameter.

GGA and hybrid approximations has reduced the LDA errors of atomization energies of standard set of small molecules by a factor 3-5. This improved accuracy has made DFT a significant component of quantum chemistry. On the other hand, all the present functionals are inadequate for situations where the density is not a slowly varying function. However, this does not keep DFT with appropriate approximations from successfully dealing with such problems.

### *The LCAO ansatz in the Kohn-Sham equations*

The central ingredient of the Kohn-Sham approach to density functional theory are the one-electron KS equations:

$$\left( -\frac{1}{2}\nabla^2 + \left[ \sum_j^N \int \frac{|\psi_j(\vec{r}_2)|^2}{r_{12}} d\vec{r}_2 + V_{XC}(\vec{r}_1) - \sum_A^M \frac{Z_A}{r_{1A}} \right] \right) \psi_i = \epsilon_i \psi_i \quad (47)$$

The term in square brackets defines the Kohn-Sham one-electron operator and Eq. 47 can be written more compactly as:

$$\hat{f}^{KS}\psi_i = \epsilon_i \psi_i \quad (48)$$

Most of the applications in chemistry of the Kohn-Sham density functional theory make use of the LCAO expansion of the Kohn-Sham orbitals. In this approach, a set of $L$ predefined basis functions $\{\eta_\mu\}$ are introduced and the K-S orbitals are linearly expanded as:

$$\psi_i = \sum_{\mu=1}^{L} c_{\mu i}\eta_\mu \tag{49}$$

When inserting Eq. 43 into Eq. 42 a new equation in very close analogy to the Hartree-Fock case is obtained:

$$\hat{f}^{KS}(\vec{r}_1)\sum_{\nu=1}^{L} c_{\nu i}\eta_\nu(\vec{r}_1) = \epsilon_i \sum_{\nu=1}^{L} c_{\nu i}\eta_\nu(\vec{r}_1) \tag{50}$$

If this equation is multiplied with an arbitrary basis function $\eta_\mu^*$ and integrated over space, $L$ equations are obtained

$$\sum_{\nu=1}^{L} c_{\nu i}\int \eta_\mu^*(\vec{r}_1)\hat{f}^{KS}(\vec{r}_1)\eta_\nu(\vec{r}_1)d\vec{r}_1 = \epsilon_i \sum_{\nu=1}^{L} c_{\nu i}\int \eta_\mu^*(\vec{r}_1)\eta_\nu(\vec{r}_1)d\vec{r}_1 \tag{51}$$
$$1 \le i \le L$$

The integrals on both sides of this equation define a matrix:

$$F_{\mu\nu}^{KS} = \int \eta_\mu^*(\vec{r}_1)\,\hat{f}^{KS}(\vec{r}_1)\eta_\nu(\vec{r}_1)d\vec{r}_1 \qquad S_{\mu\nu} = \int \eta_\mu^*(\vec{r}_1)\,\eta_\nu(\vec{r}_1)d\vec{r}_1 \tag{52}$$

which are the elements of the Kohn-Sham matrix and the overlap matrix, respectively. Both matrices are L x L dimensional. Eq. 51 can be re-written compactly as a matrix equation

$$\hat{F}^{KS}\hat{C} = \hat{S}\hat{C}\hat{\epsilon} \tag{53}$$

Hence, through the LCAO expansion, the non-linear optimization problem has been translated into a linear one, which can be expressed in the language of standard algebra.
By expanding $\hat{f}^{KS}$ into its components, the individual elements of the KS matrix become:

$$F_{\mu\nu}^{KS} = \int \eta_\mu^*(\vec{r}_1)\left(-\frac{1}{2}\nabla^2 - \sum_{A}^{M}\frac{Z_A}{r_{1A}} + \int \frac{\rho(\vec{r}_2)}{r_{12}}d\vec{r}_2 + V_{XC}(\vec{r}_1)\right)\eta_\nu(\vec{r}_1)d\vec{r}_1 \tag{54}$$

The first two terms describe the kinetic energy and the electron-nuclear interaction, and they are usually combined in one-electron integrals:

$$h_{\mu\nu} = \int \eta_\mu^*(\vec{r}_1)\left(-\frac{1}{2}\nabla^2 - \sum_A^M \frac{Z_A}{r_{1A}}\right)\eta_\nu(\vec{r}_1)d\vec{r}_1 \quad (55)$$

For the third term, the electron density $\rho$ is required and takes the following form in the LCAO scheme:

$$\rho(\vec{r}) = \sum_i^L |\psi_i(\vec{r})|^2 = \sum_i^N \sum_\mu^L \sum_{nu}^L c_{\mu i}^* c_{\nu i} \eta_\mu^*(\vec{r})\eta_\nu(\vec{r}) \quad (56)$$

The expansion coefficients are usually collected in the so-called density matrix $\hat{P}$ with elements:

$$P_{\mu\nu} = \sum_i^N c_{\mu i}^* c_{\nu i} \quad (57)$$

Thus, the Coulomb contribution in Eq. 54 can be expressed as:

$$J_{\mu\nu} = \sum_\lambda^L \sum_\sigma^L P_{\lambda\sigma} \int\int \eta_\mu^*(\vec{r}_1)\eta_\nu(\vec{r}_1)\frac{1}{r_{12}}\eta_\lambda^*(\vec{r}_2)\eta_\sigma(\vec{r}_2)d\vec{r}_1 d\vec{r}_2 \quad (58)$$

Up to this point, exactly the same formulae also apply in the Hartree-Fock case. The difference is only in the exchange-correlation part. In the Kohn-Sham scheme this is represented by the integral:

$$V_{\mu\nu}^{XC} = \int \eta_\mu^*(\vec{r}_1)V_{XC}(\vec{r}_1)\eta_\nu(\vec{r}_1)d\vec{r}_1 \quad (59)$$

whereas the Hartree-Fock exchange integral is given by:

$$K_{\mu\nu} = \sum_\lambda^L \sum_\sigma^L P_{\lambda\sigma} \int\int \eta_\mu^*(\vec{x}_1)\eta_\lambda(\vec{x}_1)\frac{1}{r_{12}}\eta_\nu(\vec{x}_2)\eta_\sigma(\vec{x}_2)d\vec{x}_1 d\vec{x}_2 \quad (60)$$

The calculation of the $L^2/2$ one-electron integrals contained in $h_{\mu\nu}$ can be fairly easily computed. The computational bottle-neck is the calcalution of the $\sim L^4$ two-electron integrals in the Coulomb term.

## *Basis sets*

*Slater-type-orbitals (STO):* they seem to be the natural choice for basis functions. They are exponential functions that mimic the exact eigen functions of the hydrogen atom. A typical STO is expressed as:

$$\eta^{STO} = Nr^{n-1}e^{-\beta r}Y_{lm}(\theta, \Phi) \tag{61}$$

Here, *n* corresponds to the principal quantum number, the orbital exponent is termed $\beta$ and $Y_{lm}$ are the usual spherical harmonics. Unfortunately, many-center integrals are very difficult to compute with STO basis, and they do not play a major role in quantum chemistry. *Gaussian-type-orbitals (GTO)*: they are the usual choice in quantum chemistry. They have the following general form:

$$\eta^{GTO} = Nx^{l}y^{m}z^{n}e^{-\alpha r^2} \tag{62}$$

*N* is a normalization factor which ensures that $\langle \eta_\mu | \eta_\mu \rangle = 1$, α represents the orbital exponent. $L = l + m + n$ is used to classify the GTO as s-functions ($L = 0$), p-functions ($L = 1$), etc.

*Contracted Gaussian functions (CGF)*: basis sets in which several primitive Gaussian functions are combined in a fixed linear combination. They are rarely used and are of the form:

$$\eta_\tau^{CGF} = \sum_a^A d_{a\tau}\eta_a^{GTO} \tag{63}$$

# PART I. Preliminary Studies

Before starting actual studies on some COMs formations, it is important to define which level of theory to use. To do so, some models were compared on both organic and transition metal containing compounds.

## *Computational details*

All calculations have been performed with a development version of the Gaussian suite of programs, using DFT methods.[21] Different levels of theory have been employed in this part: (i) In order to study organic compounds, the global hybrid B3LYP[22,23] and the double-hybrid B2PLYP[24,25] density functionals have been used, in conjunction with the m-aug-cc-pVTZ triple-$\xi$ basis set,[26,27] where *d* functions on the hydrogens have been removed. Semiempirical dispersion contributions were also included in DFT computations by means of the D3 model of Grimme.[28,29] (ii) For the study of larger transition metal containing systems, the B3PW91 functional has been chosen[20,30,31] in conjunction with the so-called LANL2DZ basis set, which includes a pseudopotential for describing inner electrons of large atoms, with polarization functions on C (d; 0.587), N (d; 0.736), O (d; 0.961), Ru (f; 1.235), Cl (d; 0.648) Pt (f; 0.8018) and Ir (f; 0.938).[32,33] The choice of this couple (functional/basis set) is based on previous results pointing out the robustness of this computational model.[34,35]

Full geometry optimizations have been performed for all the compounds and the optimized structures have been checked to be true minima on the PES by diagonalizing their Hessians.

For the organic species, after optimizations, cubic and semidiagonal quartic force constants have been next computed by finite differences of analytical Hessians and used to obtain anharmonic frequencies with the GVPT2 model taking into proper account possible resonances for frequencies[36] together with IR intensities with the DVPT2 model including both mechanical and electrical anharmonicities.[37]

At this level, the vibrational energy of asymmetric top molecules can be written

$$E_{vib} = E_0 + \sum_i \frac{h}{2\pi}\omega_i n_i + \sum_{i\leq j} X_{ij}\left(\frac{1}{2}n_i + \frac{1}{2}n_j + n_i n_j\right) \qquad (1)$$

where $E_0$ is the zero-point corrected energy, h is the Planck constant, and $n_i$ and $\omega_i$ are the vibrational quantum numbers and harmonic wavenumbers, respectively; $X_{ij}$ is the anharmonic coupling between modes i,j, which is an explicit function of cubic and semidiagonal quartic force constants.[36–38]

In the case of transition metal complexes, excitation energies from the electronic ground state and equilibrium geometries for excited states were evaluated by means of TD-DFT and unrestricted DFT. One-photon emission (OPE) spectra were also simulated within the Born−Oppenheimer and harmonic approximations by a time independent approach, which effectively takes into account the transitions from the ground vibrational state of the initial electronic state to all the vibrational states of the final electronic state.[39] Both vertical (same geometries for both electronic states) and adiabatic (optimized geometry for each electronic state) models were considered together with inclusion (vertical Hessian, VH, or adiabatic Hessian, AH) or not (vertical gradient, VG, or adiabatic shift, AS) of frequency changes and mode mixing between the two electronic states. In all cases, the Franck−Condon approximation was enforced, namely that the transition moment is only marginally affected by (small) geometry modifications. Further details can be found in ref [39]. For some specific cases, the lowest normal modes were removed from the vibronic treatment in view of their marginal role in the spectrum and poor description at the harmonic level. Composition and plot of frontier orbitals were determined respectively thanks to the Gaussview and Chemissian packages.[40,41]

Solvent effects have been modelled using the Polarizable Continuum Model (PCM),[42–44] and all the spectra have been generated and managed by the VMS-draw graphical user interface.[45]

# *Chapter 1. Accurate Infrared (IR) Spectra for Molecules Containing the C≡N Moiety by Anharmonic Computations with the Double Hybrid B2PLYP Density Functional.*

Over the past few years, the C=N and C≡N moieties have been deeply investigated,[46–49] in connection with their intrinsic role as intermediates in reactions involving purines and proteins.[50] From another point of view, the simplest form of this compound, HCN, plays a remarkable role in the interstellar space as a widespread small molecule[51–53] of relevance for prebiotic chemistry.[54] Furthermore, cyanocarbons (*i.e.*, organic compounds bearing enough cyano functional groups to significantly alter their chemical properties) are considered a classical example of discovery-driven research,[55] tetracyanoethylene (TCNE) and its derivatives playing a prominent role in the field of magnetic materials.[56,57] Finally, the vibration of the cyano group in organic and biological molecules has been studied as an infrared (IR) probe of the local environment in biological systems.[58,59] Several studies have shown that the CN stretching is highly localized and its frequency is very sensitive to environmental changes.[60] Therefore, a molecular level understanding of the tuning of frequency shift, peak width, and intensity change by stereoelectronic and environmental effects is of increasing importance. For this purpose, quantum chemical computations are playing an increasing role toward disentanglement of intrinsic and environmental effects in determining specific spectroscopic signatures.[61,62] Methods based on density functional theory (DFT) have been instrumental, especially for infrared and Raman spectra of medium- and large sized molecules, because of their reliability, coupled with favorable scaling with the number of electrons.[38,63] Several benchmarks have shown that harmonic computations performed by global hybrid functionals (*e.g.*, B3LYP),[22,23] in conjunction with medium-sized basis sets and scaled by a reduced number of empirical factors, lead to remarkable agreement with experimental frequencies and intensities of fundamental bands.[64,65] More recently, second-order vibrational perturbation theory (VPT2) generalized to include a variational treatment of leading resonances (GVPT2)[36,66,67] has been shown to deliver accurate frequencies and intensities for fundamentals, overtones, and combination bands without the need of any scaling factor.[37,68,69] Unfortunately, global hybrid functionals provide

disappointing results for some moieties involving multiple bonds such as C≡N, unless specific scaling factors are introduced.[46,70] The recent development of double hybrid functionals and of their analytical second derivatives[25] has led to more reliable geometries and vibrational spectra for a large panel of organic and biological systems.[37,38,67] On these grounds, a systematic analysis of representative molecules containing the C≡N moiety by both the double-hybrid B2PLYP functional[24,25] and its parent B3LYP global hybrid[22,23] was performed. The main objective of this investigation is to perform direct vis-à-vis comparisons of computed and experimental spectra by including the leading contributions of both mechanical and electrical anharmonicity in a general and robust strategy. To this end, the so-called hybrid approach will also be considered, in which harmonic and anharmonic contributions are computed at different levels of sophistication and the reduced dimensionality model,[38] in which anharmonicity is introduced only for a limited number of normal modes. To this end, the seven molecules shown in Figure 1.1 were selected, which are representative of quite different situations including a single C≡N moiety in different environments (**I**, **II**, **III**), two geminal (**IV**) or vicinal (**V**, **VI**) C≡N moieties, and four coupled C≡N moieties (**VII**) for our study.

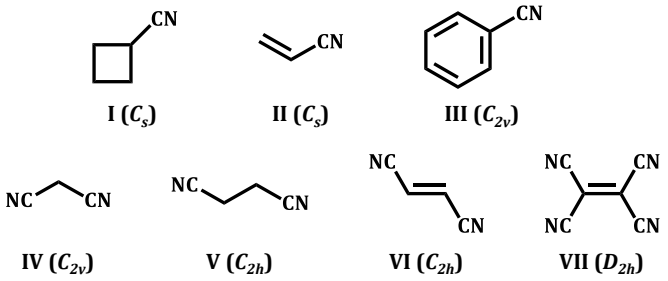



**Figure 1.1.** Investigated compounds.

## *1.1. Specific computational details*

Here, the calculations have been performed using the B3LYP-D3[22,23] and the B2PLYP-D3[24,25] functionals, in conjunction with the m-aug-cc-pVTZ basis set,[26,27] where *d* functions on the hydrogens have been removed. All calculations have been performed in the same medium employed for recording the corresponding experimental spectra (gas phase or different solvents), including bulk solvent effects, when needed. The overall molecular

symmetry and, possibly, mode degeneration were taken into full account in all computational steps including finite differentiations.

## *1.2. Preliminary study*

As a first step, the intrinsic frequency of the C≡N stretching was investigated employing HCN as a template. Contrary to the general trend observed for several other functional groups, the B3LYPD3 computational model overestimates the experimental value by about 100 $cm^{-1}$ at the harmonic level ($\omega_{B3}$) and by 70 $cm^{-1}$ at the anharmonic level ($\nu_{B3}$). As suggested in a previous study,[70] agreement with experiment can be restored by scaling the harmonic frequency without modifying the anharmonic contribution ($\Delta_{B3} = \nu_{B3} - \omega_{B3}$), *i.e.*

$$\nu_S = K\omega_{B3} + \Delta_{B3} \qquad (1)$$

For HCN we get K=0.966. On the other hand, the B2PLYPD3 computational model overestimates the experimental frequency by about 35 $cm^{-1}$ at the harmonic level ($\omega_{B2}$), but inclusion of anharmonicity ($\nu_{B2}$), decreases the discrepancy to less than 5 $cm^{-1}$, a value allowing quantitative comparisons with experiment. Furthermore, anharmonic corrections evaluated at the B3LYPD3 and B2PLYPD3 levels (hereafter $\Delta_{B3}$ and $\Delta_{B2}$, respectively) are very close, thus suggesting that a hybrid method estimating the anharmonic frequency by

$$\nu_H = \omega_{B2} + \Delta_{B3} \qquad (2)$$

could represent the best compromise between accuracy and computational efficiency avoiding at the same time the use of any empirical scaling factor.

Starting from these results, we have then investigated the effect of different chemical environments and of coupling between several moieties in tuning the characteristic frequency of the C≡N stretching together with the corresponding IR intensity thanks to the selection of species (Figure 1.1).

Some of the systems are particularly demanding since they require not only a good description of inductive and conjugative effects on the C≡N vibrational frequency, but also strong couplings between different C≡N moieties. The harmonic and anharmonic vibrational wavenumbers of the C≡N stretching for all the studied molecules are compared in Table 1.1 with experimental data and the general trends are sketched in Figure 1.2.

## *1.3. Full anharmonic treatment*

The experimental data (black line in Figure 1.2) show that the presence of more than one CN moiety increases the average vibrational frequency, but leads to negligible splitting, except for compound **VII**. Among the different computational methods, only the full B2PLYP anharmonic treatment (yellow line in Figure 1.2) fully reproduces this trend also from a quantitative point of view, whereas all the other methods give only qualitatively correct trends for compounds **I**–**VI** and fail to provide significant splittings for compound **VII**. As a matter of fact, the experimental splitting between the lowest and highest CN stretchings of **VII** amounts to 48 $cm^{-1}$, whereas the computed anharmonic value is 18 $cm^{-1}$ at the B3LYPD3 level, which is increased to 30 $cm^{-1}$ at the B2PLYP level.

Furthermore, the difference between experimental and B2PLYPD3 anharmonic frequencies is constant (15 $cm^{-1}$) for the three highest frequencies and drops to just 3 $cm^{-1}$ for the lowest frequency.

It appears that the electronic structure of tetracyanoethylene (**VII**) is particularly challenging and, contrary to the other studied molecules can be hardly interpreted in terms of localized CN moieties. In more general terms, the B3LYPD3 anharmonic results (Table 1.1 and blue line in Figure 1.2) provide systematic errors on the vibrational frequencies of C≡N stretchings ranging between 60 $cm^{-1}$ and 100 $cm^{-1}$, with respect to the experiment (corresponding to ~3%) and a mean absolute error (MAE) of ~70 $cm^{-1}$. These results confirm that quantitative agreement with experiment cannot be reached by either harmonic or anharmonic computations, even in the absence of strong couplings without using specific scaling factors. The results collected in Table 1.1 and the yellow line in Figure 1.2 show that the B2PLYPD3 computed anharmonic frequencies for the C≡N stretchings are, instead, in remarkable agreement with the experimental data. The discrepancy between theory and experiment ranges between 1 $cm^{-1}$ and 15 $cm^{-1}$ (corresponding to an average error of 0.26%) and the MAE is reduced to a fully satisfactory value (5.8 $cm^{-1}$). Furthermore, all the experimental trends are correctly reproduced, including the trend for the strongly coupled C≡N moieties of compound **VII**.

**Table 1.1.** Experimental *vs.* computed vibrational frequencies ($cm^{-1}$) for the molecules shown in Figure 1.1.

| Molecule | Mode | $\nu_{exp.}$(C≡N) | B3LYPD3/m-aug-cc-pVTZ | | | | B2PLYPD3/m-aug-cc-pVTZ | | | | $\nu_H = \omega_{B2} + \Delta_{B3}$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $\omega_{B3}$ | $\nu_{B3}$ | $\Delta_{B3}$ | $\lvert\nu_{B3}-\nu_{exp}\rvert$ | $\omega_{B2}$ | $\nu_{B2}$ | $\Delta_{B2}$ | $\lvert\nu_{B2}-\nu_{exp}\rvert$ | $\nu_H$ | $\lvert\nu_H-\nu_{exp}\rvert$ |
| **I** | 1 | 2240[a] | 2342 | 2310 | 32 | 70 | 2275 | 2239 | 36 | 1 | 2243 | 3 |
| **II** | 2 | 2228[b] | 2335 | 2306 | 29 | 78 | 2268 | 2230 | 38 | 2 | 2239 | 11 |
| **III** | 3 | 2235[c] | 2335 | 2302 | 33 | 67 | 2269 | 2229 | 40 | 6 | 2236 | 1 |
| **IV** | 4 | 2265[d] | 2377 | 2346 | 31 | 81 | 2305 | 2272 | 34 | 6 | 2274 | 9 |
| | 5 | 2265[d] | 2369 | 2337 | 32 | 72 | 2307 | 2263 | 44 | 2 | 2275 | 10 |
| **V** | 6 | 2257[e] | 2354 | 2318 | 36 | 61 | 2291 | 2255 | 36 | 2 | 2255 | 2 |
| | 7 | 2257[e] | 2355 | 2326 | 29 | 69 | 2292 | 2257 | 35 | 0 | 2263 | 6 |
| **VI** | 8 | 2239[f] | 2342 | 2309 | 33 | 70 | 2276 | 2238 | 38 | 1 | 2243 | 4 |
| | 9 | 2230[f] | 2327 | 2294 | 37 | 64 | 2261 | 2223 | 38 | 7 | 2228 | 6 |
| **VII** | 10 | 2262[f] | 2365 | 2329 | 36 | 67 | 2284 | 2247 | 37 | 15 | 2248 | 14 |
| | 11 | 2248[f] | 2344 | 2319 | 25 | 71 | 2265 | 2233 | 32 | 15 | 2240 | 8 |
| | 12 | 2237[f] | 2340 | 2315 | 25 | 78 | 2261 | 2222 | 39 | 15 | 2236 | 1 |
| | 13 | 2214[f] | 2340 | 2311 | 29 | 97 | 2269 | 2217 | 52 | 3 | 2240 | 26 |
| ***MAE*** | | | | | | 72.4 | | | | 5.8 | | 7.6 |

[a]In gas phase from Ref. [71]; [b]In gas phase from Ref. [72]; [c]In gas phase from Ref. [73]; [d]In solution (DMSO) from Ref. [74]; [e] In solid phase (Neat) from Ref. [75]; [f]In solution (bromoform) from Ref. [76].

We can thus conclude that the B2PLYPD3 computational model is able to reproduce at the same time intrinsic, environmental, and coupling effects on the CN stretching frequency with remarkable accuracy, whereas this is not the case for the B3LYPD3 computational model, irrespective of inclusion (or exclusion) of anharmonic contributions.

## *1.4. Scaling factor and hybrid approach*

In previous works, it has been showed that the most critical step in the computation of accurate vibrational frequencies and intensities is a good description at the harmonic level, whereas several methods deliver comparable anharmonic corrections.[37,38,67] This finding is fully confirmed for C≡N stretchings, since the average anharmonic corrections for the seven studied compounds are quite similar at both B3LYPD3 and B2PLYPD3 levels (~35 $cm^{-1}$), whereas the harmonic values are strongly different.

**Table 1.2.** B3LYP-D3 Scale Factors (K).

| Molecule | Mode | $\upsilon_{exp}$ | $\omega_{B3}$ | $\upsilon_{B3}$ | $\Delta_{B3}$ | $\upsilon_{exp} + \Delta_{B3}$ | $K_i$ | $\upsilon_{corr}$ |
|---|---|---|---|---|---|---|---|---|
| **I** | 1 | 2240 | 2342 | 2310 | 32 | 2272 | 0.970111 | 2237 |
| **II** | 2 | 2228 | 2335 | 2306 | 29 | 2257 | 0.966595 | 2234 |
| **III** | 3 | 2235 | 2335 | 2302 | 33 | 2268 | 0.971306 | 2230 |
| **IV** | 4 | 2265 | 2377 | 2346 | 31 | 2296 | 0.965923 | 2272 |
| | 5 | 2265 | 2369 | 2337 | 32 | 2297 | 0.969607 | 2264 |
| **V** | 6 | 2257 | 2354 | 2318 | 36 | 2293 | 0.974087 | 2245 |
| | 7 | 2257 | 2355 | 2326 | 29 | 2286 | 0.970701 | 2253 |
| **VI** | 8 | 2239 | 2342 | 2309 | 33 | 2272 | 0.970111 | 2236 |
| | 9 | 2230 | 2327 | 2294 | 33 | 2263 | 0.972497 | 2222 |
| **VII** | 10 | 2262 | 2365 | 2329 | 36 | 2298 | 0.97167 | 2256 |
| | 11 | 2248 | 2344 | 2319 | 25 | 2273 | 0.96971 | 2246 |
| | 12 | 2237 | 2340 | 2315 | 25 | 2262 | 0.966667 | 2243 |
| | 13 | 2214 | 2340 | 2311 | 29 | 2243 | 0.958547 | 2239 |

For instance, the computed frequencies for the C≡N stretching of compound **I** at the harmonic levels are 2342 and 2275 $cm^{-1}$ for B3LYPD3 and B2PLYPD3, respectively. We have investigated if the application of a constant scaling factor (K) to the B3LYPD3 harmonic frequencies could lead to improved agreement with experimental frequencies. Equation 1 leads to very similar scale factors for all the CN stretchings of the studied molecules (see Table 1.2) with an average value ($K_{B3}$ = 0.969) that is very similar to that discussed above for the prototypical HCN molecule (K = 0.966). The use of $K_{B3}$ in eq. 1 leads to remarkable agreement with experiment for all the studied compounds (green line in Figure 1.2) and points out the transferability of the scaling and its "intrinsic" nature, since environmental effects tuning the final frequency value are well reproduced at the B3LYPD3 level, except for the extreme case of compound **VII**.

A more satisfactory way of reducing the computational effort involved in full anharmonic computations at the B2PLYPD3 level (which become rapidly prohibitive for larger systems) is to couple harmonic B2PLYPD3 vibrational frequencies and intensities with B3LYPD3 anharmonic corrections without introducing any scaling factor. The results (red line in Figure 1.2) show that, while the overall computational effort becomes reasonable, the final accuracy remains fully satisfactory, with the MAE and maximum error being 7.6 and 26 $cm^{-1}$, respectively.

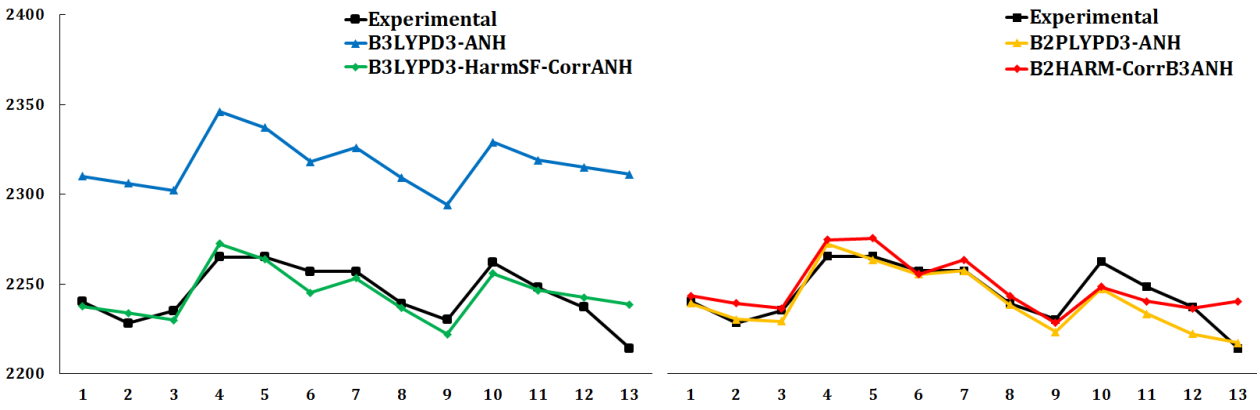



**Figure 1.2.** Comparison between experimental frequencies of C≡N stretchings (black) and their counterparts issuing from different computational models: blue = $\upsilon_{B3}$; green = $k_{B3}\omega_{B3}$ + $\Delta_{B3}$; orange = $\upsilon_{B2}$; red = $\omega_{B2}$ + $\Delta_{B3}$; mode numbering (abscissa) is given in Table 1.1.

## *1.5. Reduced dimensionality approach*

Another way of obtaining accurate anharmonic frequencies for selected vibrations is the so-called reduced dimensionality approach (RDA) in which displacements are performed only along a selected number of normal modes in the numerical evaluation of third and fourth energy derivatives.[38] For purposes of illustration, we selected compounds **I**, **II**, and **III**, which contain a single C≡N moiety, so that a one-dimensional model (hereafter referenced as 1-RDA) is sufficient to obtain all the leading anharmonic couplings between CN stretching and the other normal modes. Table 1.3 shows that the results issuing from the 1-RDA approach yield an absolute deviation of < 3 $cm^{-1}$, with respect to the results issuing from a full anharmonic treatment. The strength of the reduced dimensionality approach is, of course, the tremendous saving of computational resources (*e.g.*, by a factor of 20 for compound **III**) coupled to a negligible reduction of the overall accuracy.

**Table 1.3.** Experimental *vs.* simulated (B3LYPD3, B2PLYPD3 / m-aug-cc-pVTZ) vibrational frequencies ($cm^{-1}$) for compounds containing a single C≡N moiety.

| Molecule | Exp | B3LYPD3/m-aug-cc-pVTZ | | | B2PLYPD3/m-aug-cc-pVTZ | | |
|---|---|---|---|---|---|---|---|
| | | *Harm* | *Full Anh* | *1-RDA* | *Harm* | *Full Anh* | *1-RDA* |
| **I** | 2240[a] | 2342 | 2310 | 2310 | 2275 | 2239 | 2238 |
| **II** | 2228[b] | 2335 | 2306 | 2303 | 2268 | 2230 | 2231 |
| **III** | 2235[c] | 2335 | 2302 | 2300 | 2269 | 2229 | 2227 |

[a]In gas phase from Ref. [71]; [b]in gas phase from Ref. [72]; [c]in gas phase from Ref. [73].

## *1.6. IR spectral shape*

Let us now focus our attention on the reproduction of complete IR spectra. As an example, the IR spectrum of compound **I** is shown in Figure 1.3.a., obtained by B2PLYPD3 (red) and B3LYPD3 (blue) computations, including both mechanical and electrical anharmonicity. Each computed line was convoluted by a Lorentzian function with half-weight at half-maximum (HWHM) of 5 $cm^{-1}$, chosen on the ground of systematic studies aimed to simulate medium resolution IR spectra. It is apparent that the overall shape of the spectrum simulated with the B3LYPD3 functional is in good agreement with its experimental

counterpart, but the position of the C≡N stretching peak is strongly shifted and its relative height overestimated. In the spectrum simulated using hybrid vibrational frequencies ($\nu_H$ = $\omega_{B2}$ + $\Delta_{B3}$, green line), band positions are nicely reproduced, but the relative intensities remain unsatisfactory.

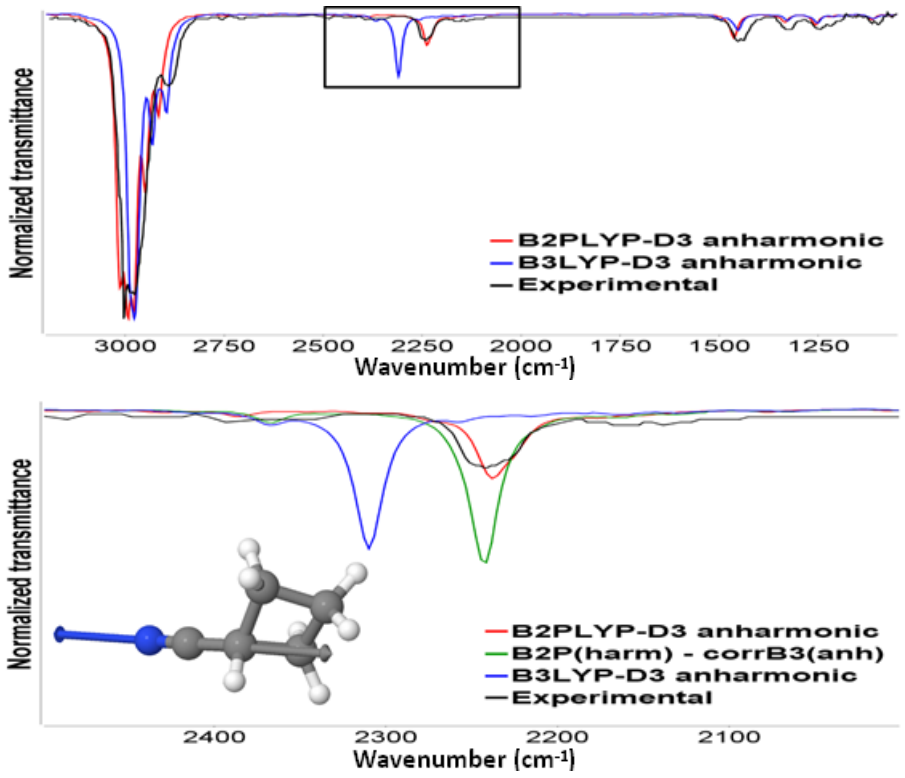



**Figure 1.3.** (a) Simulated anharmonic (red = B2PLYPD3, blue = B3LYPD3) and experimental (black) spectra of **I** (1000-3200 $cm^{-1}$ range); (b) enlarged view of the region centered on the CN stretching peak (green = B2PLYPD3 harmonic + correction from anharmonic B3LYPD3). Theoretical spectra convoluted with a Lorentzian distribution function with HMHW = 5$cm^{-1}$.

Once again, the B2PLYPD3 functional performs a remarkable job, restoring full agreement with the experiment concerning both the position and the relative intensity of this specific band. Furthermore, a Fermi resonance between the C≡N stretching vibration and the fisrt overtone of the 15$^{th}$ normal mode (corresponding to the torsion of the four membered ring) is now correctly recognized. As a matter of fact, C–H stretchings have similar intensities (51.8 and 47.3 $km.mol^{-1}$ at the B3LYPD3 and B2PLYPD3 level, respectively), whereas this is

not the case for the C≡N stretching (18.7 and 2.4 km.mol$^{-1}$ at the B3LYPD3 and B2PLYPD3 level, respectively). Once again, the problem can be traced back to the harmonic contribution (22.2 and 9.5 km.mol$^{-1}$ at the B3LYPD3 and B2PLYPD3 level, respectively), so that the discrepancy can be strongly reduced by adding B3LYPD3 anharmonic contributions to B2PLYPD3 harmonic intensities.This hybrid intensity (6.0 km.mol$^{-1}$) shows a difference (2.6 km.mol$^{-1}$) from the full B2PLYPD3 value of the same order of magnitude as those observed for "well-behaving" modes, like C–H stretchings. Since this model does not involve any additional computational burden, we strongly suggest using the hybrid approach for both frequencies and intensities possibly in connection with reduced dimensionality models.

## *1.7. Partial conclusion*

The accuracy and robustness of different computational approaches to vibrational frequencies and IR intensities for several compounds including one, two, or four C≡N moieties was tested. Together with conventional harmonic contributions, mechanical and electrical anharmonicities have been taken into account in the framework of generalized second-order vibrational perturbation theory by means of global hybrid B3PLYP or double-hybrid B2PLYP density functionals, semiempirical dispersion contribution, and an extended basis set. Anharmonic vibrational calculations including all the normal modes, and selecting only the normal mode corresponding to C≡N stretching were separately performed. As a matter of fact, the C≡N stretching is very sensitive to the functional used, and, especially, to the proper choice of a correlation part. Therefore, it is not surprising that the B2PLYP level of theory (which includes MP2 corrections) delivers more accurate results than B3LYP, rivaling the performance of the very efficient CCSD(T) with complete basis set extrapolation method, offering a strongly reduced computational cost. Moreover, an integrated strategy combining B2PLYP harmonic results and B3LYP anharmonic corrections was further validated. This technique dramatically reduced the computational effort without any significant loss of accuracy, except for the extreme case of compound **VII**, involving four strongly coupled CN moieties. It is noteworthy, in this connection, that, besides mechanical effects, electrical effects (here, IR intensities) also are significantly improved including B2PLYP harmonic values.

# *Chapter 2. Transition Metal Complexes*

Even though right now, no information on metal complexes in ISM can be found, it is not inconceivable that one day some of them could be detected. It is therefore always useful to perform theoretical studies on this type of species.

## 2.1. Vibronic coupling investigation to compute phosphorescence spectra of Pt(II)

Compounds exhibiting emission in the near-infrared are mainly of interest for the development of organic light-emitting diodes (OLED) or light-emitting cells (LEC).[77–79] Platinum(II) complexes seem particularly promising candidates for their strong and tunable phosphorescence, which could be of interest in a wide spectrum of applications such as sensors or storage data, especially thanks to the possibility of emitting in the 500−800 nm range.[80–82] The emission wavelength can be tuned by decreasing the LUMO energy or by raising the HOMO energy. Other possible strategies include the decrease of the HOMO−LUMO gap by increasing the intra- and interligand conjugation[83–85] or by introducing a second metal center into the molecule.[86,87] In a recent paper,[87] a combined experimental and computational investigation was performed on the geometric and electronic structures of several mono- and dinuclear platinum(II) complexes with a square planar $d^8$ configuration of the metal (Figure 2.1.1). The ligands consist of one or two acetylacetonate (acac) moiety(ies) and a biphenylpyrimidine (**I**, **IV**) or biphenylpyrazine (**II**, **III**, **V**, **VI**) bridging fragment. The number of metal atoms and/or of ligands could allow a fine-tuning of emission wavelengths and therefore represent crucial targets to test the robustness of new computational methods. This is why an extension of previous computational investigations (limited to electronic ground states) to emission (phosphorescence) spectra was carried out. This task is not trivial due to the presence of closely spaced multiplets and to the difficulty of managing triplet states by conventional methods. The dimensions (and the number) of the systems suggest the use of effective computational approaches rooted into density functional theory (DFT) and its time-dependent extension (TD-DFT). As a matter of fact, the first triplet states of molecular systems can be considered either ground high spin electronic states (being thus amenable to standard unrestricted DFT computations) or excited electronic

states requiring TD-DFT computations starting from the singlet ground state. A comparison between these two routes is one of the aims of the present investigation.

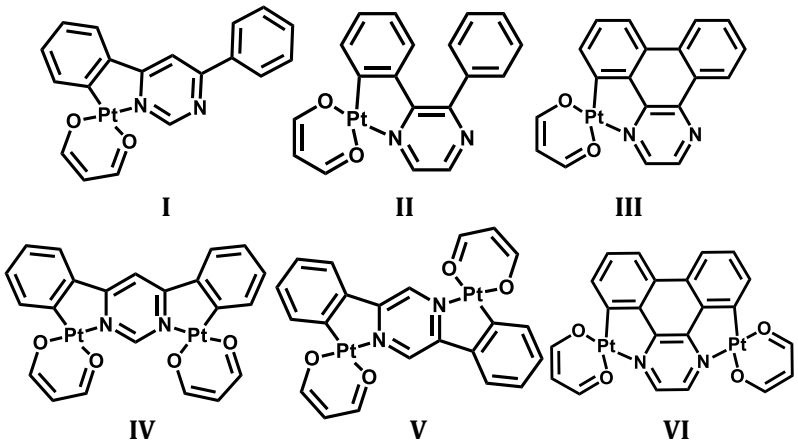



**Figure 2.1.1.** Investigated compounds.

A second point of interest is related to geometry relaxation effects (difference between vertical and adiabatic transitions) and to vibrational modulation of electronic spectra (the so-called vibronic couplings). Thanks to the development of effective analytical derivatives for both DFT and TD-DFT methods, together with powerful and general tools for the evaluation of vibronic couplings (including Franck-Condon, Herzberg-Teller, and mode mixing effects), it was possible to perform a comprehensive investigation of those aspects which have a non-negligible impact on a proper comparison with experiments.

## *2.1.1. Specific computational details*

The B3PW91 functional[23,30,31] has been used for these calculations, in conjunction with the LANL2DZ basis set. On the basis of test computations, the VG approach will be our standard in the following, except for some selected cases.

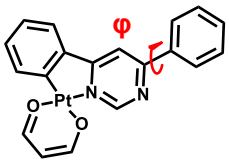
## *2.1.2. Structural investigation*

Table 2.1.1 lists some structural parameters for the fundamental singlet and first triplet states of compounds **I**–**VI**. The bond lengths (C–N, C–C, C–O) around the platinum atom are given for all compounds, together with the dihedral angles formed between the phenyl

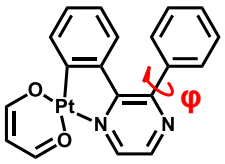
and the aza moiety($\varphi$) for compounds **I** and **II**. Except for the possible torsion of phenyl group (which is much larger in **II** than in **I** due to the stronger hindrance it experiences in the former case) in the ground electronic state, the geometrical parameters of all the molecules are quite similar. Indeed the Pt–N distances always remain between 1.98 and 2.00 Å.

**Table 2.1.1.** Calculated structural data for **I-VI**: Bond lengths (Å), angles (°), (B3PW91/LANL2DZ+pol./PCM).

| | | **I** | **II** | **III** | **IV** | | **V** | **VI** |
|---|---|---|---|---|---|---|---|---|
| **Ground State** | *Pt-N* | 1.996 | 1.982 | 1.996 | 1.996 | | 1.983 | 1.994 |
| | *Pt-C* | 1.964 | 1.959 | 1.975 | 1.962 | | 1.966 | 1.986 |
| | *Pt-O₁* | 2.022 | 2.025 | 2.015 | 2.019 | | 2.019 | 2.015 |
| | *Pt-O₂* | 2.128 | 2.144 | 2.127 | 2.128 | | 2.134 | 2.115 |
| | $\varphi$ | 21 | 55 | | | | | |
| **Triplet State (unr)** | *Pt-N* | 1.993 | 1.967 | 1.995 | 1.995 | | 1.976 | 1.998 |
| | *Pt-C* | 1.919 | 1.930 | 1.925 | 1.946 | | 1.946 | 1.970 |
| | *Pt-O₁* | 2.046 | 2.044 | 2.044 | 2.022 | | 2.035 | 2.026 |
| | *Pt-O₂* | 2.113 | 2.142 | 2.104 | 2.119 | | 2.133 | 2.101 |
| | $\varphi$ | 6 | 42 | | | | | |
| **Triplet State (TD-DFT)** | *Pt-N* | 1.990 | 1.963 | 1.989 | 1.965 | 1.989 | 1.977 | 2.001 |
| | *Pt-C* | 1.924 | 1.946 | 1.930 | 1.970 | 1.937 | 1.950 | 1.957 |
| | *Pt-O₁* | 2.043 | 2.037 | 2.042 | 2.016 | 2.035 | 2.030 | 2.031 |
| | *Pt-O₂* | 2.123 | 2.148 | 2.113 | 2.143 | 2.133 | 2.135 | 2.108 |
| | $\varphi$ | 6 | 37 | | | | | |




**I**




**II**

The Pt–C bond lengths have the same behavior except for complexes **III** and **VI**, which exhibit larger distances. This trend can be explained by the strong delocalization inside the ligand. Compared to the ground states, a few general trends can be observed in the excited states. In particular, one of the Pt–O bond lengths increases when going from $S_0$ to the triplet, whereas the other Pt–O and, especially, the Pt–C bond lengths decrease, the latter showing a stronger shortening for monometallic complexes (0.02–0.05 Å) than in the case of bimetallic

ones (0.01−0.04 Å). At the same time, the $\varphi$ angle decreases dramatically in the excited electronic states of **I** and **II** due to an increase of the electronic delocalization.

### *2.1.3. Absorption*

TD-DFT computations have been performed at the optimized geometries of the $S_0$ electronic states to investigate the absorption properties of the targeted complexes and compared to the available experimental data (Table 2.1.2 and Figure 2.1.2).[87]

All computed excitations are in reasonable agreement with the experimental data, confirming this computational method, in line with previous investigations.[34,35] A more detailed analysis can be based on a direct vis-à-vis comparison of the available experimental spectra[87] with those issuing from vibronic simulations. Figure 2.1.2 shows that the methodology used is robust. Indeed, for all the compounds, all the absorption bands are reproduced in our simulations, and also the peak intensities match their experimental counterparts.

Thanks to this accuracy, it is possible to assign unambiguously the absorption bands (Table 2.1.2). As already discussed in a previous paper by Williams et al,[87] the first transition is mainly a HOMO to LUMO transition with a strong metal-to-ligand charge transfer (MLCT) character. Indeed, as expected, the HOMOs involved in the transitions exhibit a strong platinum character (especially its d orbitals), leading to the so-called MLCT. The *acac* moiety plays a negligible role in the transitions due to its marginal contributions in the frontier orbitals. The $C_4N_2$ ring possesses the strongest accepting ability, and therefore the LUMO and LUMO+1 of each compound are strongly localized on this moiety and, in the case of the LUMO, more precisely on both nitrogen atoms.

For the bimetallic complexes, symmetry constraints impose equal contributions of both Pt atoms and both *acac* moieties in the MOs. As suggested by Williams and co-workers, a spin-forbidden transition (singlet to triplet) occurs for some biplatinum complexes and is well reproduced in our computations. However, nothing related to this singlet → triplet excitation has been discussed on **VI**, whereas, according to our computations, two excitations also appear at 615 and 573 nm. These values fit nicely with the observed tail around 600 nm in the experimental spectrum.

**Table 2.1.2.** Absorption data: experiment *vs.* TD-DFT calculated values; *t* stands for an absorption peak from a Singlet ⟶ Triplet excitation (B3PW91/LANL2DZ+pol./PCM).

| | $\lambda_{exp.}$[a] **(nm)** | $\lambda_{calc.}$ **(nm)**, ***oscillator strength (f)*** | **Transitions involved, *% weight*** | **Gap HO-LU (eV)** |
|---|---|---|---|---|
| ***I*** | 408 | 428 *f=0.09* | HO → LU *85%* | 3.54 |
| | 338 | 328 *f=0.13* | HO-3 → LU *60%*<br>HO-1 → LU+1 *25%* | |
| ***II*** | 449 | 437 *f=0.03* | HO → LU *92%* | 3.59 |
| | 398 | 372 *f=0.07* | HO-1 → LU *78%* | |
| | 340 | 356 *f=0.19* | HO → LU+1 *82%* | |
| ***III*** | 463 | 457 *f=0.04* | HO → LU *93%* | 3.38 |
| | 407 | 392 *f=0.14* | HO-1 → LU *88%* | |
| | 344 | 348 *f=0.21* | HO → LU+1 *89%* | |
| ***IV*** | 532sh[b] | *t* 541 *f=0.00*<br>*t* 527 *f=0.00* | HO → LU *91%*<br>HO-2 → LU *45%*<br>HO-1 → LU *40%* | 3.24 |
| | 481 | 467 *f=0.27* | HO → LU *96%* | |
| | 430 | 409 *f=0.17* | HO-2 → LU *84%* | |
| ***V*** | 618sh | *t* 628 *f=0.00* | HO → LU *76%* | 3.08 |
| | 503 | 513 *f=0.12* | HO → LU *95%* | |
| | 453 | 425 *f=0.24* | HO-2 → LU *91%* | |
| | 372 | 373 *f=0.48* | HO-6 → LU *49%*<br>HO → LU+1 *41%* | |
| | 333 | 341 *f=0.34* | HO-2 → LU+1 *87%* | |
| ***VI*** | 595sh | *t* 615 *f=0.00*<br>*t* 573 *f=0.00* | HO-1 → LU *96%*<br>HO → LU *77%*<br>HO-2 → LU *13%* | 3.05 |
| | 509 | 509 *f=0.10* | HO → LU *91%* | |
| | 462 | 431 *f=0.25* | HO-2 → LU *87%* | |
| | 428 | 396 *f=0.05* | HO-5 → LU *96%* | |
| | 350 | 349 *f=0.33* | HO-1 → LU+1 *77%* | |

[a] From Ref. [87]**;** [b] shoulder**.**

The computed data reported in Table 2.1.2, and the simulated spectra in Figure 2.1.2 fit nicely with experimental data. One should notice that some asymmetric bands are well reproduced in our computations which involve several electronic states. Vibrational progressions, if not negligible, are thus masked by overlapping Gaussians describing the

average behavior of electronic states: under such circumstances, explicit vibronic computations are of marginal interest.

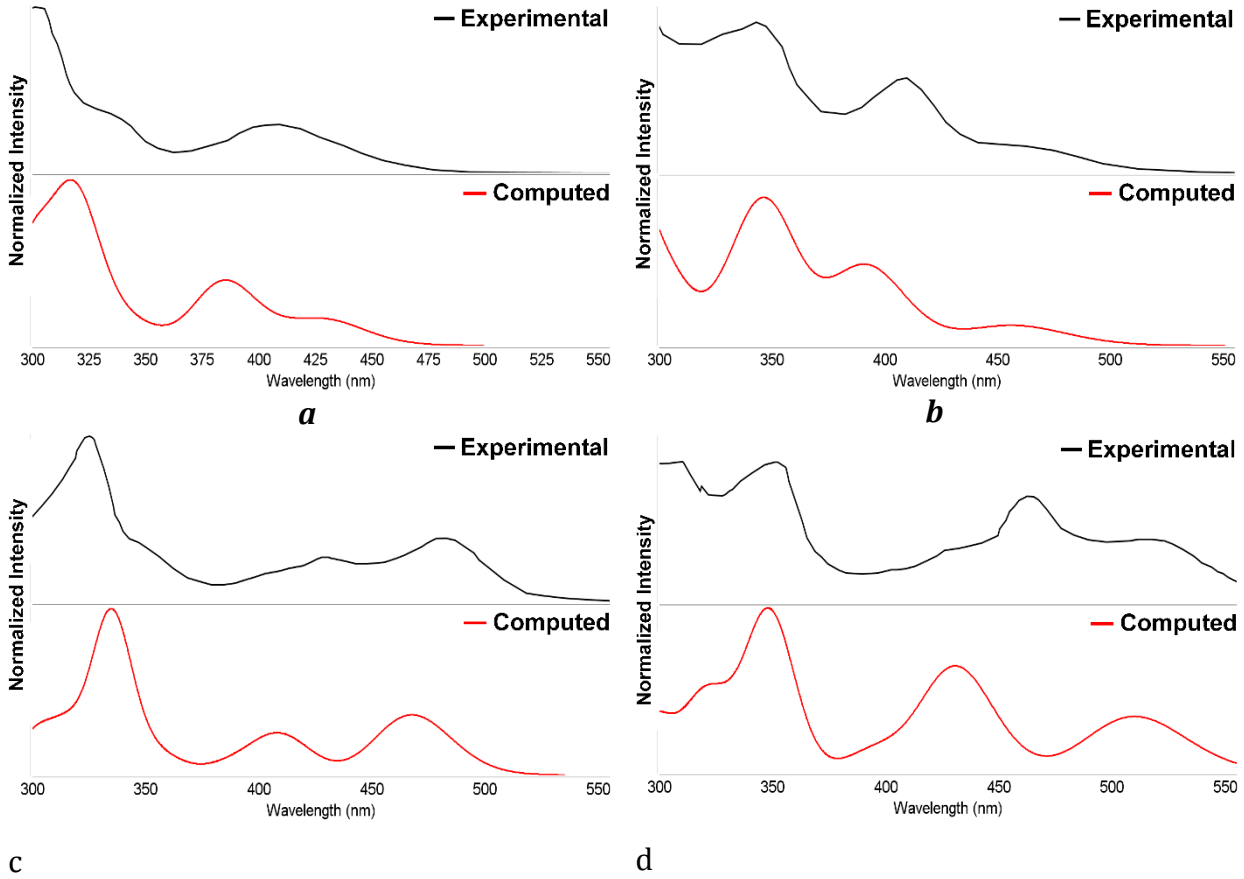



**Figure 2.1.2.** Experimental vs. computed electronic absorption spectra of I (a), III (b), IV (c) and VI (d)(solvent included). (B3PW91/LANL2DZ+pol./PCM).

## *2.1.4. Phosphorescence*

The phosphorescence wavelength was first studied by conventional open-shell DFT computations (unrestricted) of the first triplet state. In this case the electronic wavelength was obtained as the energy difference between the optimized triplet open-shell (ET,r(T)) and a single-point calculation of the singlet closed-shell with the optimized geometry of the triplet (ES,r(T)). This method gives on average an error around 0.15 eV. While the error is higher for complexes II and V, it remains acceptable even in this case (<0.30 eV) (Table 2.1.3).

This preliminary study paves the way to more refined calculations based on vibronic transitions. As one can see in Table 2.1.3, the vibrational contributions affect the position of the peaks. In every case, the unrestricted method provided accurate results in terms of position.

**Table 2.1.3.** Emission data: experiment vs TD-DFT, unrestricted-DFT electronic and vibronic calculated values (B3PW91/LANL2DZ+pol./PCM).

| | ***Exp.***[a] | ***Electronic*** | | ***Vibronic*** | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | *TD-DFT* | *Unrestricted* | *AS (TD)* | *AH (TD)* | *AS (unr)* | *AH (unr)* | *VG (unr)* | *VH (unr)* |
| ***I*** | 521 | 573 | 554 | 566 | 592 | 557 | 582 | 527 | 522 |
| ***II*** | 554<br>578 | 716 | 639 | 616<br>673 | - | 574<br>612 | - | 575<br>604 | - |
| ***III*** | 568 | 629 | 609 | 588 | - | 575 | - | 573 | - |
| ***IV*** | 558<br>594 | 623 | 571 | 580<br>628 | - | 552<br>598 | - | 552<br>599 | - |
| ***V*** | 628<br>685 | 797 | 733 | 705<br>780 | - | 665<br>730 | - | 664<br>715 | - |
| ***VI*** | 628 | 707 | 645 | 670 | 716 | 610 | 624 | 611 | 613 |

[a] From Ref. [87].

Furthermore, the "cheap" AS and VG levels of theory show a substantial agreement with respect to experiment. Indeed, after inclusion of these corrections in the VG case, the average error on each peak falls around 0.06 eV with respect to experiment. Upon additions of Hessian effects, the peaks are all shifted to lower energies, with the exception of **I** (VH). Considering the global band shape and contrary to absorption spectra, the phosphorescence ones involve only one final electronic state ($S_0$), so that any asymmetry (or fine structure) in the spectral shape can only originate from vibronic effects. Figure 2.3.1 shows the overall spectra obtained by unrestricted DFT electronic computations together with VG or VH treatments of vibronic contributions. The band shape for each compound is well reproduced by the computations, giving further support to the reliability of the underlying computational approach. Indeed, it is observed that complex **II** has two intense peaks, with a difference of *ca.* 30 nm, and the first slightly less intense in comparison to the second one. This behavior is also well reproduced by the computations. Furthermore, for complexes **III** and **VI**,

experimental and observed phosphorescence spectra are almost identical. For complexes **IV** and **V**, despite a slight discrepancy on the shoulder intensity of the curve, the localization of the peak is, once again, in good agreement with experimental data. As test cases, the band shape of **I** and **VI** have also been investigated after a Hessian diagonalization on both structures (ET,r(T) and ES,r(T)). As a matter of fact, the resulting spectrum (VH) fits slightly better the experimental one.

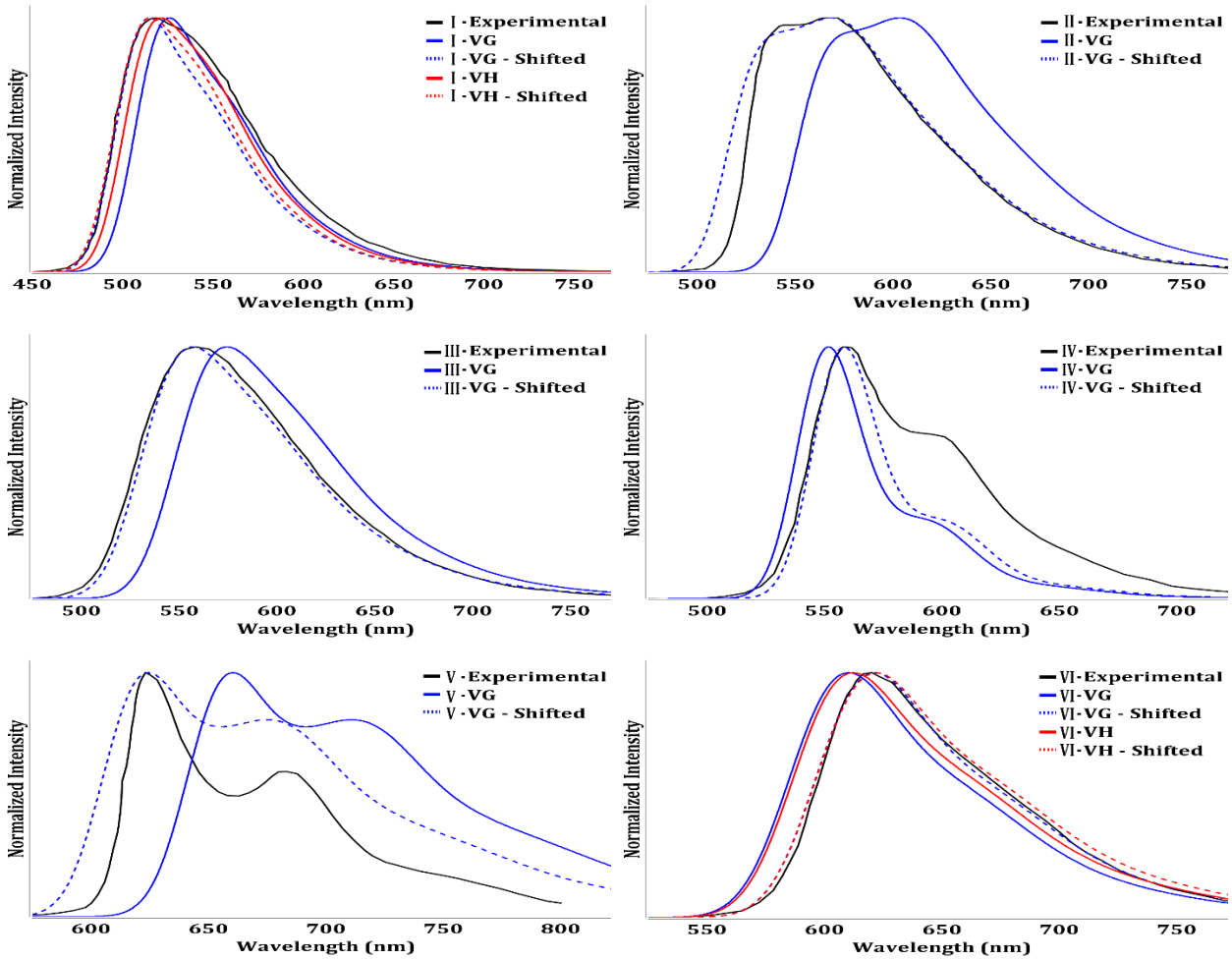



**Figure 2.1.3.** Experimental (black) vs. simulated phosphorescence spectra at the vertical level of theory. VG(blue), VG+Shift(dashed), VH(red), VH+Shift(dashed).

Conventional TD-DFT computations from the $S_0$ ground electronic state led to larger and less systematic errors. In particular, complexes **I**, **III**, **IV**, and **VI** shared a systematic error of *ca.* 10%, which is quite typical for this level of theory. On the other hand, the error for **II**

and **V** reached 25%, a value too high to allow any reliable investigation. If one compares the TD-DFT electronic energies with respect to the experimental ones, the disagreement is always above 0.20 eV, and up to 0.50 eV (**II**), with a mean value of about 0.30 eV (based on the highest experimental energy values). However, upon addition of the vibronic coupling, the error drops, ranging from about 0.07 to 0.30 eV, with an average around 0.20 eV. It is of interest to analyze the main normal modes ruling the vibrational progressions of the emission spectra. In Figure 2.1.4a, the simulated phosphorescence spectrum of complex **VI**, is superimposed to the stick spectrum showing the different vibrational contributions at the VG level of theory. As one can see, the band maximum does not correspond to the 0–0 transition but to excited vibrational levels. To characterize the principal vibrational progressions, the main components of the shift vector (*i.e.*, the gradient of the final state projected on the normal modes of the initial state) is shown in Figure 4b.

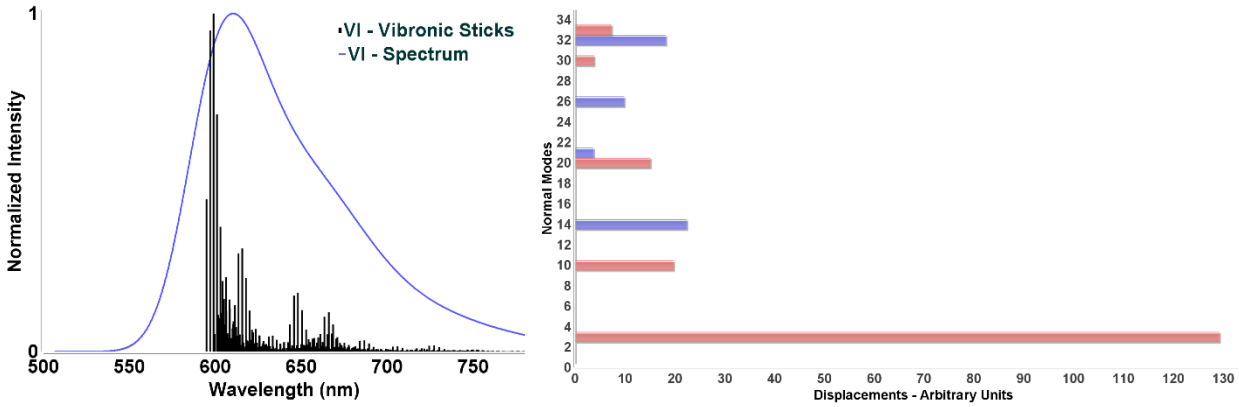



**Figure 2.1.4**. Stick and convoluted phosphorescence spectrum with of **VI** (**a**). Main components of the shift vector along normal modes of the initial electronic state (**b**) (blue and red refer to positive and negative sign, respectively).

One can notice that the third normal mode of the triple state (unrestricted), whose computed harmonic frequency is 56 $cm^{-1}$, dominates the shift vector, followed by non-negligible (although at least 6 times lower) contribution by normal modes 10, 14, 20, 26, and

32, whose computed harmonic frequencies are 133, 202, 263, 361, and 455 $cm^{-1}$, respectively.

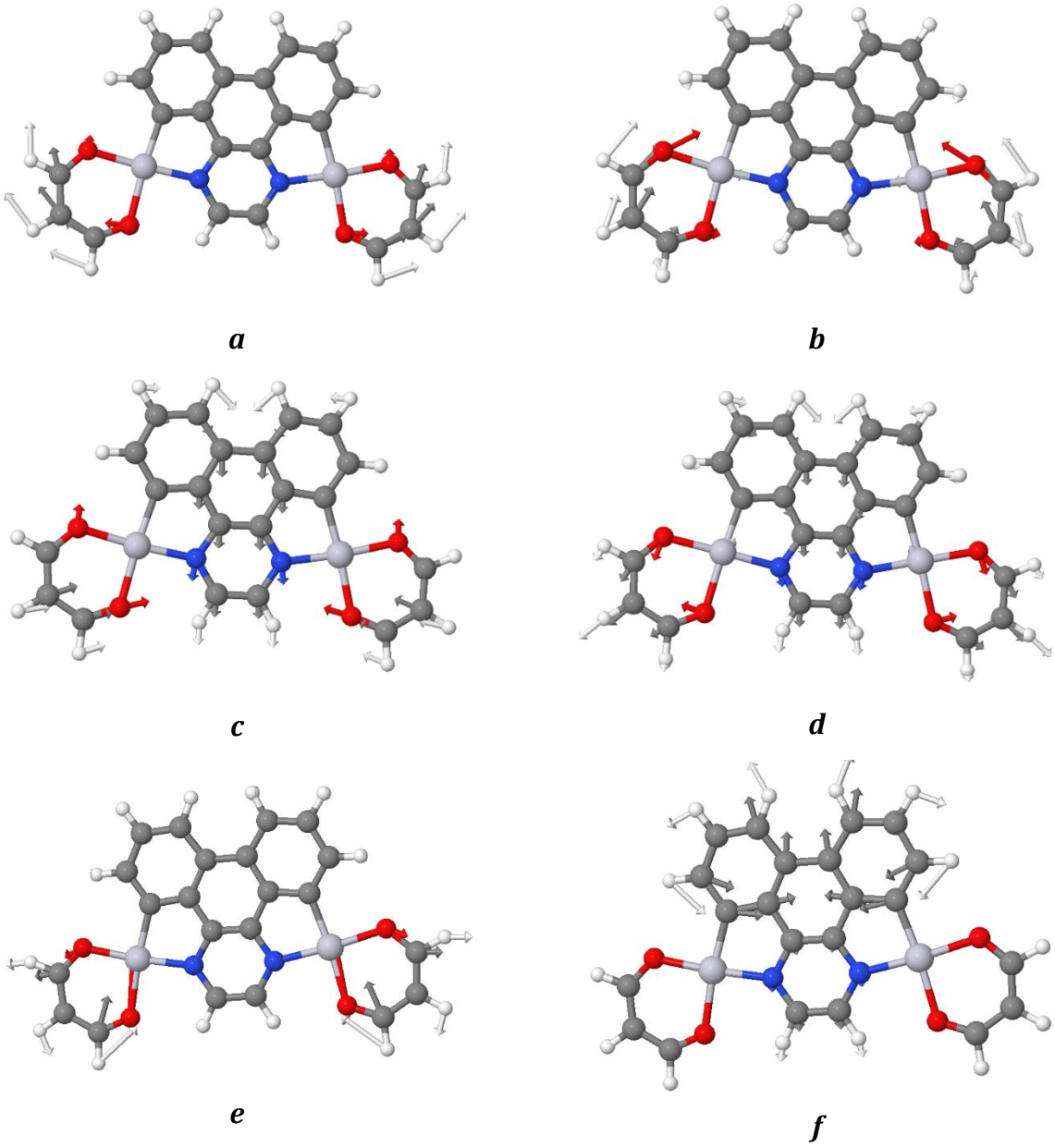



**Figure 2.1.5**. Normal modes 3 (**a**), 10 (**b**), 14 (**c**), 20 (**d**), 26 (**e**), and 32 (**f**) of the triplet state of **VI**.

These six modes (3, 10, 14, 20, 26, 32) are sketched in Figure 2.1.5. It is noteworthy that all of these modes involve out-of-plane displacements and receive significant contributions from the acac moiety. Actually, modes 3, 10, 26, and 32 involve exclusively the

acac moiety. This point is remarkable because this moiety is only weakly involved in the electronic transition (*Cf.* Absorption section), but contributes strongly to tune the overall band shape.

### *2.1.5. Partial conclusion*

From a methodological point of view, the B3PW91 hybrid functional in conjunction with the LANL2DZ valence basis set and pseudo potential confirms its reliability and robustness for analyzing structural and spectroscopic features of transition metal complexes involving, inter alia, platinum and ruthenium, together with different kinds of conjugated ligands. For high spin states, both conventional unrestricted DFT and TD-DFT (starting from the singlet ground state) computations can be performed with the first approach, providing more reliable results. Inclusion of vibronic couplings then leads to remarkable agreement with experiment: positions, intensities, and band shapes are nicely reproduced in the simulated spectra. It is noteworthy that the acac moiety gives a marginal direct contribution to the electronic transition but tunes the band shape through vibronic couplings.

Nevertheless, some limitations are still present especially concerning proper account of the spin-orbit coupling, which would allow the computation of absolute (instead of relative) intensities together with Herzberg-Teller effects. Another significant aspect is the proper treatment of inner electrons beyond the pseudopotential approach employed in the present work.

## 2.2. Validation of a computational protocol to simulate near IR phosphorescence spectra for an Ir(III) metal complex

Transition metal complexes have a large number of applications, particularly thanks to their strong ability to luminesce. In particular, one can cite iridium complexes that have this capacity to emit in a large energetic range. However, it remains difficult to experimentally fully characterize these targets with only one technique. Therefore, computationally oriented chemists can provide crucial information on these molecules, especially thanks to the very recent developments in this area.[25,35,38] Furthermore, Density Functional Theory (DFT) and its Time-Dependent extension (TD-DFT) are increasingly used to get electronic and vibrational information from small- to large-size systems thanks to the strong improvements in hardware and software.[88–94]

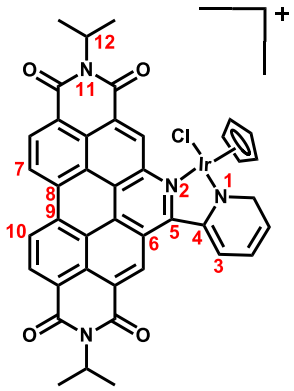



**Figure 2.2.1.** Investigated complex including atom labels discussed later.

In a recent paper, a combined experimental and computational (ground state + vertical electronic absorption) study performed on the geometric and electronic structures of an iridium(III) complex with a $d^6$ configuration on the metal was performed (Figure 2.2.1). The study also reported the observed phosphorescence spectrum of this complex which possesses a strong signature in the near IR.[95] Despite its possible attractive nature for new technological devices, this molecule represents a computational challenge. The properties of the excited state (geometries, phosphorescence spectra) can be obtained considering practically the first triplet state of the system either as a ground high spin state, according to the unrestricted technique, or as the excited electronic state using the TD-DFT formalism

starting from the singlet ground state. However, the size of the system prevents the use of this last case if one wants to obtain the vibrational contribution of the triplet excited state due to the very large computational cost.

### *2.2.1. Specific computational details*

Again here, the B3PW91 functional[23,30,31] has been chosen in conjunction with the LANL2DZ basis set. Vibronic contributions have next been taken into account with the general Adiabatic Hessian approach (AH),[39] in which the PES of the final state is more properly described, including mode mixing as well. In order to reduce the computational cost and improve the vibronic progression, the alkyl terminal moieties in the experimental structure were replaced by isopropyl and the Cp* ligand by a Cp.

### *2.2.2. Structural investigations*

After optimization, the obtained geometries of the complex were compared with available experimental data.[95] As one can see in Table 2.2.1, the bond lengths around the metal are in reasonable agreement with the ones observed using X-ray diffraction.

**Table 2.2.1.** Experimental and computed relevant geometric parameters.

| | **Exp. Ref.** [95] | **DFT Ref.** [95] | **This study** |
|---|---|---|---|
| ***Ir-Cl* (Å)** | 2.388[a] | 2.380 | 2.418 |
| ***Ir-N*$_2$ (Å)** | 2.138(5) | 2.110 | 2.093 |
| ***Ir-N*$_1$ (Å)** | 2.084[a] | 2.057 | 2.060 |
| ***Ir-Cp* (Å)** | 1.803(8) | 1.801 | 1.795 |
| ***C*$_{11}$*-N*$_{12}$ (Å)** | 1.495[a] | 1.310/1.315 | 1.493 |
| ***3-4-5-6* (°)** | 23.7[a] | 21.8 | 22.4 |
| ***7-8-9-10* (°)** | 3.8 | 2.1 | 0.5 |

[a]Average value taken directly from the crystal structure

Indeed the average Ir–C (Cp*) is 1.803(8) Å in the crystal, whereas it is 1.795 Å in the simulations. The computed Ir–Cl bond length slightly overestimates the experimental

distance. On the other hand, the computed Ir–$N_{1/2}$ distances are both underestimated by *ca.* 0.03 Å with respect to experiment. The previous computational results reported together with the experimental ones make appear a very short $C_{12}$–$N_{13}$ distance (around ≈1.313 Å).[95] This result is intriguing when one compares this bond length with the one observed in the crystal (1.495 Å) which fits nicely with our computations (1.493 Å). Finally, examining the different torsions of the molecule, it appears that the computed structure fits nicely with the one obtained using X-ray diffraction. All these results gave confidence for further investigations.

### *2.2.3. Electronic structure and vertical excitations*

On the optimized geometry of the compound, TD-DFT vertical excitations have been performed. Its molecular orbital (MO) diagram is sketched in Figure 2.2.2 and the ten first computed vertical transitions with an oscillator strength higher than 0.01 are given in Table 2.2.2.

**Table 2.2.2.** Ten first computed vertical transitions with f (oscillator strength) >0.01.

| $\lambda_{calc}$ (nm) | f | Transition | Assignment |
|---|---|---|---|
| 503 | 0.6648 | HOMO ⟶ LUMO 96% | |
| 492 | 0.0332 | HOMO -1 ⟶ LUMO 81% | |
| 482 | 0.1339 | HOMO -1 ⟶ LUMO +1 76%<br>HOMO ⟶ LUMO +1 14% | |
| 463 | 0.2307 | HOMO ⟶ LUMO +1 56%<br>HOMO -1 ⟶ LUMO +1 17%<br>HOMO -1 ⟶ LUMO 14%<br>HOMO -2 ⟶ LUMO 8% | MLCT + ILCT |
| 432 | 0.2129 | HOMO -2 ⟶ LUMO 87% | |
| 419 | 0.1182 | HOMO -2 ⟶ LUMO +1 88% | |
| 388 | 0.2658 | HOMO -3 ⟶ LUMO 93% | |
| 367 | 0.1001 | HOMO -3 ⟶ LUMO +1 63%<br>HOMO -4 ⟶ LUMO +1 20% | |

As one can see, the highest occupied frontier molecular orbitals possess a strong metal character, while most LUMOs are predominantly localized on the large π moiety.

As one can see in Table 2.2.2, many electronic transitions occur between 400 and 500 nm. The calculated lowest energy electronic transition is mainly a HOMO⟶LUMO (96 %) excitation. This transition is assigned to a MLCT and ILCT, like all the transitions of interest exhibited on Table 2.2.2.

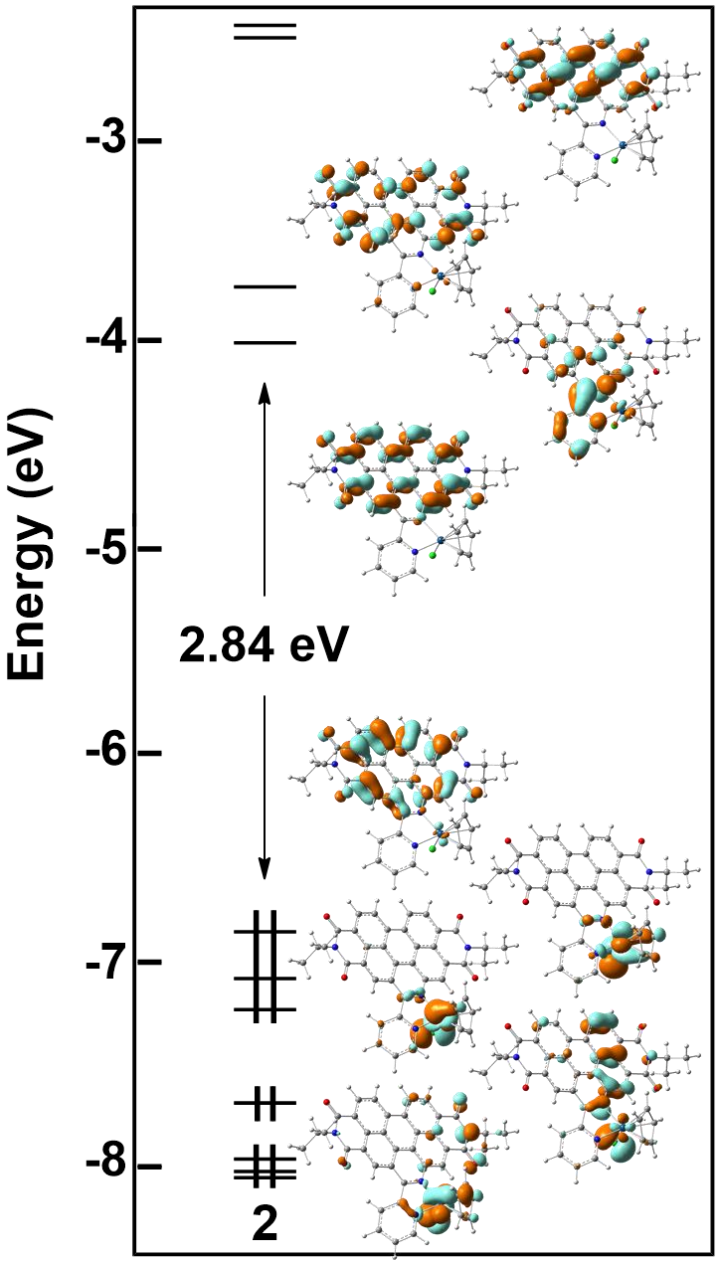



**Figure 2.2.2.** Frontier MOs of the complex at the singlet state (cut-off = 0.04).

### *2.2.4. Excited states*

The excited state of the complex was treated by means of the unrestricted method. The nature of the optimized geometry has been checked to be a true minimum on the PES. The optimized geometries of the fundamental singlet state and of the triplet excited state were compared in Table 2.2.3. As one can see, the bond lengths around the metal stay similar between the two states with variations only at the third decimal place. The torsion angles do not vary either. The variation of the π backbone of the complex was also investigated, and no significant changes were noticed.

**Table 2.2.3.** Differences in some relevant geometric parameters.

| | **Singlet** | **Triplet** |
|---|---|---|
| ***Ir-Cl* (Å)** | 2.418 | 2.417 |
| ***Ir-N*$_2$ (Å)** | 2.093 | 2.089 |
| ***Ir-N*$_1$ (Å)** | 2.060 | 2.059 |
| ***Ir-Cp* (Å)** | 1.795 | 1.797 |
| ***C*$_{11}$*-N*$_{12}$ (Å)** | 1.493 | 1.492 |
| ***3-4-5-6* (°)** | 22.4 | 21.8 |
| ***7-8-9-10* (°)** | 0.5 | 1.1 |

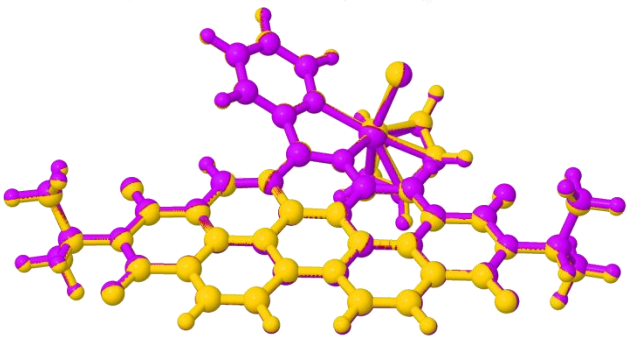


**Figure 2.2.3.** Differences of geometry between the initial and the final state.

In order to use the vibronic coupling methodology, the geometries of the initial and final states need to be close. Therefore, both geometries (initial and final states) are exhibited

in Figure 2.2.3 and they indeed look similar. This latter fact gives confidence for complementary analysis. Indeed, the geometric superposition of both states for each complex leads to a strong overlap. This last result allows the inclusion of the vibrational contributions to the electronic transitions for the simulation of the phosphorescence spectrum.

### *2.2.5. Phosphorescence*

On these grounds, it is now possible to simulate the phosphorescence spectrum of the complex. Particular care was taken concerning the progression of the vibronic spectrum where the minimum percentage is, as recommended, at least 90 %.

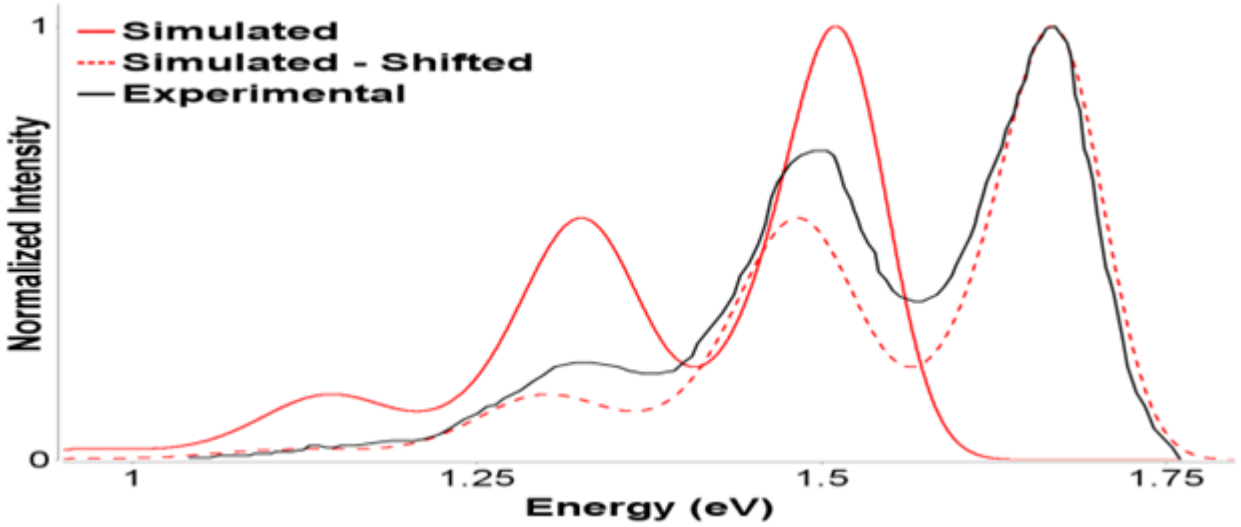



**Figure 2.2.4.** Simulation (red) versus experimental (black) phosphorescence spectrum of the complex. The shifted simulated spectrum (red dashed) is moved to match the maximum of the experimental one.

In Figure 2.2.4 is reported the simulated spectrum of the compound. As one can see, the simulated (red) and experimental spectra (black) are in good agreement, despite a small red shift in our simulations with respect to experiment. This shift is evaluated around 0.15 eV, which is an acceptable error at this level of theory (DFT), and allows us to shift (dashed) our simulated spectrum on the maximum of intensity of the experimental one. It becomes clearer that the simulation allows a direct vis-à-vis with experiment.

The complex exhibits three distinct bands around 1.7, 1.5 and 1.3 eV, and a tail in the IR region. The four distinct data are well reproduced in our simulation, although a small red shift is seen in comparison with experiment in the IR region.

### *2.2.6. Partial conclusion*

The computational protocol (B3PW91/LANL2DZ) seems to be a good compromise if one attempts to reproduce and interpret the phosphorescence spectrum of an iridium complex. This protocol has already been used in previous studies (section 2.2) but here it was proven to be also adequate for the study of a larger system.

# PART II. Gas-phase Formation Routes for Complex Organic Molecules in the Interstellar Medium

## *Computational details*

All DFT calculations have been carried out with a development version of the Gaussian suite of programs.[21] Most of the computations were performed with the B3LYP hybrid density functional[22,23,63], the double hybrid B2PLYP functional[24] and with the second-order Møller-Plesset perturbative many body treatment (MP2).[96–100] These methods have been used in conjunction with the SNSD double-$\xi$ basis set[101] and the m-aug-cc-pVTZ triple-$\xi$ basis set,[26,27] where *d* functions on hydrogens have been removed. Semiempirical dispersion contributions were also included into DFT computations by means of the D3 model of Grimme, leading to B3LYP-D3 and B2PLYP-D3 models.[28,102] Some reported energies were computed using the Complete Basis Set (CBS-QB3) method, which leads to very accurate values by employing a CC ansatz in conjunction with complete basis set extrapolation.[103,104] Full geometry optimizations have been performed for all compounds checking the nature of the obtained structures (minima or transition states) by diagonalizing their Hessians.

Cubic and semidiagonal quartic force constants have been computed by finite differences of analytical Hessians and employed to obtain anharmonic frequencies with the GVPT2 model, taking possible resonances for frequencies,[36] together with IR intensities (including both mechanical and electrical anharmonicities),[37,38] into proper account. For transition states (TS), all the sums exclude the mode corresponding to the imaginary frequency. Those contributions enter in the semiclassical definition of reaction probability proposed by Miller,[105] which includes both tunneling and anharmonicity, and has been used here for the evaluation of reaction rates.

All the spectra have been generated and managed by the VMS-draw graphical user interface.[45]

# *Chapter 1. Cyanomethanimine isomers*

The recent detection of E-cyanomethanimine in Sgr B2(N) in the Green Bank Telescope (GBT) PRIMOS survey by Zaleski et al.[106] has raised the question of how this species is formed under the conditions of the cold interstellar medium (ISM). Before this study, the principal astrochemical databases for molecular formation routes in the ISM (UMIST1265 and KIDA66) did not include any pathway for this species or its isomers.

## 1.1. Re-assessment of the thermodynamic, kinetic, and spectroscopic features of cyanomethanimine derivatives: a full anharmonic perturbative treatment

Imines are able to undergo various transformations and act as reactive intermediates in a large panel of reactions of interest for health and biology in general.[107–110] Furthermore, over the past few years, the C≡N moiety has attracted increasing attention,[46–49] also thanks to its putative role as intermediate in reactions involving purines and proteins.[50] Furthermore, the role of HCN in the interstellar space and prebiotic chemistry is also remarkable.[51–53,111,112] It has been shown that only a very refined model, taking into account anharmonic effects, is able to retrace with accuracy the vibrational signatures of chemical compounds. Following some previous investigations on the thermodynamic and vibrational signatures of organic and inorganic molecules at the anharmonic level,[25,35,37,62,68,69,113] new insights based on the experimental and computed data obtained recently by Osman[46] on the Z- and E-C-cyanomethanimine molecules are presented herein. The vibrational and thermochemical properties of both molecules and the transition state connecting them (Figure 1.1.1) were investigated, taking into account the most refined model available so far.

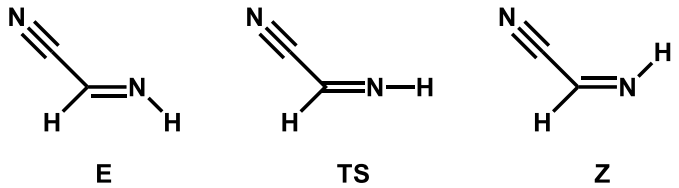



**Figure 1.1.1.** *E*-C-cyanomethanimine (*E*), Transition State (TS), and *Z*-C-cyanomethanimine (*Z*).

Computations of the reaction rate at both harmonic and anharmonic levels are also reported and comparisons with experimental data and harmonic level of theory are shown.

### *1.1.1. Specific computational details*

Most of the computations were performed with the B3LYP-D3 functional,[22,23] in conjunction with the SNSD basis set.[101] Additional computations have been performed with the B2PLYP-D3 functional[24] and with the MP2 treatment,[96–100] in conjunction with the m-aug-cc-pVTZ basis set,[26,27] where *d* functions on hydrogens have been removed.

The anharmonic CN stretching frequency was not in satisfactory agreement with the experimental data in the case of B3LYP(+D3). This problem can be traced back to a bad description at the harmonic level because the Δ (Harm.-Anharm.) is almost identical for all levels of theory (*ca.* 30-40 $cm^{-1}$). To solve this issue, a Scale Factor (SF) on CN stretching has been set up on the HCN molecule and will only be applied to the harmonic CN frequency of both *Z* and *E* isomers, when needed. Then, the anharmonic correction is added to this new value. This SF is defined as below

$$SF = \frac{\nu_{exp} + \Delta(\nu_{Harm.} - \nu_{Anharm.})}{\nu_{Harm.}} \qquad (1)$$

with $\nu_{CN}$ = 2097 $cm^{-1}$ (H-C≡N).[114]

**Densities of States and Partition Functions**

The principal quantities needed for thermodynamic and kinetic calculations are the partition functions (at different temperatures), which can be generally obtained from the microscopic density of states through a Boltzmann integral

$$Q(T) = \int_0^\infty \rho(E) \times e^{-\left(\frac{E-E_0}{k_\beta T}\right)} dE \quad (2)$$

and the density of states ρ(E) is calculated directly from the energy level structure. If the system is composed of *N* independent degrees of freedom, the overall partition function is given by the product of the individual partition functions, *i.e.* $Q(T) = Q_{tr}(T) \times Q_{rot}(T) \times Q_{vib}(T)$, with $Q_{tr}(T)$, $Q_{rot}(T)$ and $Q_{vib}(T)$ representing the translational, rotational and vibrational contributions, respectively. Dealing with isomerization reactions, $Q_{tr}(T)$ is the same for all

stationary points and can be neglected in the evaluation of equilibrium and rate constants. Next, for 3-dimensional rotations, the density of states is approximately given by

$$\rho_{rot}(E) = \frac{2}{\sigma}\sqrt{\frac{E}{ABC}} \qquad (3)$$

where σ is the rotational symmetry number, *i.e.* the number of unique orientations of the rigid molecule that only interchange identical atoms, and A, B and C are the rotational constants. The challenging part is the computation of the vibrational density of states for an anharmonic system, both because of the non-constant energy level spacing and because of the coupling of different vibrations. For the computation of the anharmonic vibrational density of states we have implemented a random walk Wang-Landau algorithm,[115–117] which scales very favorably with the dimension of the system.

**Thermodynamic Functions.**

Through the partition functions and their temperature dependence, the thermodynamic functions (enthalpy, entropy, and Gibbs free energy) of both isomers are obtained according to

$$H = E_0 + k_\beta T + k_\beta T^2 \left(\frac{\partial lnQ}{\partial T}\right) \qquad (4)$$

$$S = k_\beta T \left(\frac{\partial lnQ}{\partial T}\right) + k_\beta lnQ \qquad (5)$$

$$G = H - TS = E_0 + k_\beta T(1 - lnQ) \qquad (6)$$

**Equilibrium and Rate Constants.**

Both microcanonical (fixed E) and canonical (fixed T) equilibrium and rate constants have been calculated. For the microcanonical ensemble (assuming the reaction is written as E↔Z), the equilibrium constant is given by

$$K_{eq}(E) = \frac{\rho_Z(E)}{\rho_E(E)} \qquad (7)$$

where $\rho_Z$ and $\rho_E$ are, respectively, the densities of states of the two species at energy E. On the other hand, in the canonical ensemble, the equilibrium constant is given by

$$K_{eq}(T) = \frac{Q_Z(T)}{Q_E(T)} \times e^{-(\frac{\Delta E}{k_\beta T})} \qquad (8)$$

where $Q_Z$ and $Q_E$ are, respectively, the partition functions of the two species, and $\Delta E = E_0(Z) - E_0(E)$ is the difference of their zero-point energies. As far as the corresponding rate constants are concerned, semiclassical transition state theory (SCST) is used.[105] In this model, the microcanonical rate constant for the forward reaction is given by

$$k(E) = \frac{N(E)}{h\rho_E(E)} \qquad (9)$$

where N(E) is the sum of states for the transition state, corrected for the possibility of tunnelling and quantum reflection

$$N(E) = \sum_{n_1} \sum_{n_2} \cdots \sum_{n_{F-1}} P_n(E) \quad (10)$$

with the semiclassical tunneling probability given by

$$P_n(E) = \frac{1}{1+e^{2\vartheta(n,E)}} \qquad (11)$$

In the above equations, *F*-1 is the total number of vibrational degrees of freedom orthogonal to the reaction path. In the framework of VPT2, $\vartheta(n,E)$ is an explicit function of the harmonic and anharmonic force constants.[105] Using the Wang-Landau algorithm, it is possible to evaluate effectively N(E) separating the energy range in small bins (with width $\delta E$)

$$N(E) = \delta E \sum_{i=1}^{E/\delta E} \rho(E_i)\langle P(E_i)\rangle \qquad (12)$$

where $\langle P(E_i)\rangle$ is the average tunneling probability for the different vibrational quantum number choices leading to $E \in [E_i - \delta E, E_i + \delta E]$.

The corresponding canonical expression for the rate constant is

$$k(T) = \frac{1}{h} \frac{\int_{-\infty}^{+\infty} N(E) e^{-(\frac{E}{k_\beta T})} dE}{Q_{re}(T)} \quad (13)$$

where $Q_{re}(T)$ is the reactants partition function.

In accordance with the principle of detailed balance, both in the canonical and microcanonical ensemble, the equilibrium constant is given by the ratio of the two rate constants (direct and inverse).

### *1.1.2. Infrared investigations*

Table 1.1.1 lists the experimental and computed vibrational wavenumbers of *Z* and *E* isomers. The calculations were performed at the harmonic and anharmonic levels, using the B3LYP-D3 and B2PLYP-D3 models in conjunction with the SNSD and m-aug-cc-pVTZ basis sets. As one can see on *Z* isomer, the computed results are close to experiment. The C–H and N–H stretchings issuing from the VPT2 model are in much better agreement with experiment than the scaled harmonic values of Ref. [46]. The same behavior is observed concerning vibrations below 1500 $cm^{-1}$. However, the C=N and C≡N stretches are not sufficiently well described in this model (B3LYP-D3/SNSD level). To solve this problem, a Scale Factor at the harmonic level has been set up (*Cf.* section 1.1.1). Addition of VPT2 anharmonic corrections to those scaled harmonic values leads to vibrational wavenumbers in remarkable agreement with experiment. It is noteworthy that B2PLYP-D3 results do not require any scaling.

**Table 1.1.1.** Experimental and calculated vibrational wavenumbers ($cm^{-1}$) of Z and E.

| **Z** | | | | | | **E** | | | | | | ***assign.***[a] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***Osman***[a] | | ***B3LYP-D3/ SNSD*** | | ***B2PLYP-D3/ m-aug*** | | ***Osman*** | | ***B3LYP-D3/ SNSD*** | | ***B2PLYP-D3/ m-aug*** | | |
| *exp* | *calc* | *Harm* | *Anharm (+SF)* | *Harm* | *Anharm* | *exp*[b] | *calc*[a] | *Harm* | *Anharm (+SF)* | *Harm* | *Anharm* | |
| 3306 | 3334 | 3456 | 3286 | 3462 | 3295 | 3288 | 3456 | 3472 | 3302 | 3477 | 3310 | $\nu_1$ N–H str. |
| 2943 | 3037 | 3128 | 2955 | 3161 | 2963 | 3050-2930 | 3076 | 3069 | 2906 | 3103 | 2931 | $\nu_2$ C–H str. |
| 2239 | 2251 | 2332 | 2300 (2225) | 2259 | 2220 | 2246 | 2336 | 2345 | 2312 (2237) | 2275 | 2235 | $\nu_3$ C≡N str. |
| 1599 | 1617 | 1672 | 1635 (1582) | 1648 | 1609 | 1609 | 1678 | 1681 | 1644 (1561) | 1663 | 1625 | $\nu_4$ C=N str. |
| 1388 | 1362 | 1410 | 1377 | 1425 | 1391 | 1386 | 1402 | 1407 | 1370 | 1415 | 1379 | $\nu_5$ C–H bend |
| 1218 | 1203 | 1243 | 1213 | 1248 | 1218 | 1218 | 1237 | 1239 | 1212 | 1242 | 1216 | $\nu_6$ CNH bend |
| 904 | 922 | 917 | 904 | 922 | 909 | 908 | 912 | 909 | 898 | 914 | 903 | $\nu_7$ C–C str. |
| | 617 | 619 | 612 | 617 | 610 | | 620 | 620 | 613 | 621 | 614 | $\nu_8$ NCC bend |
| | 213 | 231 | 230 | 232 | 232 | | 242 | 244 | 243 | 244 | 243 | $\nu_9$ CCN bend |
| 1095 | 1093 | 1124 | 1097 | 1138 | 1109 | | 1106 | 1106 | 1079 | 1119 | 1091 | $\nu_{10}$ C=N tor. |
| 815 | 846 | 845 | 833 | 846 | 835 | 800 | 827 | 824 | 812 | 829 | 817 | $\nu_{11}$ C–H O.P. |
| | 324 | 332 | 326 | 333 | 328 | | 322 | 328 | 323 | 329 | 325 | $\nu_{12}$ CNN O.P. |

[a]From Ref. [46], calculations performed at the B3LYP/aug-cc-pVDZ level
[b]From Ref. [118]

A peak with a very low intensity was reported by Evans and co-workers around 3180 $cm^{-1}$, hypothetically assigned to $2\nu_{C=N}$.[118] Thanks to these new computations it is now possible

to confirm this hypothesis. Indeed, this weak band is clearly assigned as an overtone (2 quanta) of the fourth vibration mode (C=N stretching). Using the B2PLYP-D3/m-aug-cc-pvtz level oftheory, this value is computed at ≈3200 $cm^{-1}$ (*Z* isomer), which is in very good agreement with experiment.

According to the available experimental data on the *E* isomer, the agreement is also satisfactory.[118] Concerning the CN stretching, the same SF as before has been used (B3LYP-D3 computations) to obtain even more reliable results for these vibrations.

Moreover, for both isomers, a set of resonances appeared for the C-H stretching, coupled with a combination band involving $\nu_4$ and $\nu_5$. For the sake of consistency, the lowest value of this resonance was always taken in the assignment.

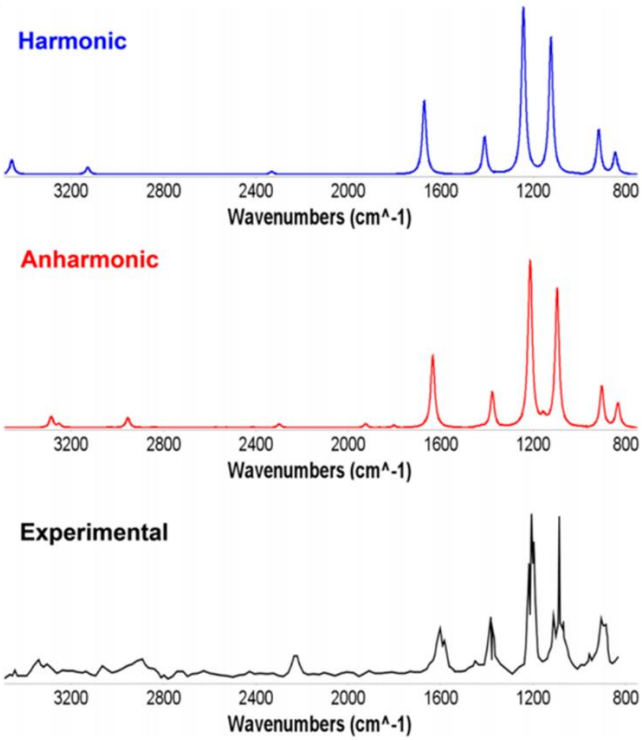



**Figure 1.2.2.** Observed (black, from Ref. [46]) and simulated IR spectra of *Z*-C-cyanomethanimine (B3LYP-D3/SNSD): harmonic (blue) and anharmonic (red). Normalization has been performed with respect to the peak above 1200 $cm^{-1}$.

Finally, the MAE using B2PLYP-D3 and the m-aug-cc-pVTZ basis set is good (≈ 11 $cm^{-1}$). This value is appreciable because the lowest energy modes could not be compared due to the lack of experimental data. If we compare the experimental Infrared spectrum (wavenumbers and intensities) of *Z*-C-cyanomethanimine to the new computations (B3LYP-D3/SNSD) at both harmonic and anharmonic levels of theory with respect to experiment (Figures 1.1.2 and 1.1.3), one should notice that the harmonic spectrum fits quite nicely in the 1000-1500 $cm^{-1}$ region. However, it turns out that experimental and harmonic spectra show a more pronounced disagreement above 1500 $cm^{-1}$. On the other hand, the anharmonic spectrum exhibits a remarkable accuracy with respect to experiment on the investigated region. Furthermore, one can see the presence of simulated small hills around 1700-2000 $cm^{-1}$ which correspond to a set of overtones and combination bands. When zooming in the region 1070-1400 $cm^{-1}$ on the experimental and anharmonic simulated spectra (Figure 1.1.3), it can be observed that all the information given by the experimental spectrum can be retrieved in the simulated one but for one peak around 1310 $cm^{-1}$. This peak, also present in the IR spectrum of *Z*-prop-2-ynylideneamine (structurally close to *Z*-C-cyanomethanimine), was not assigned.[46]

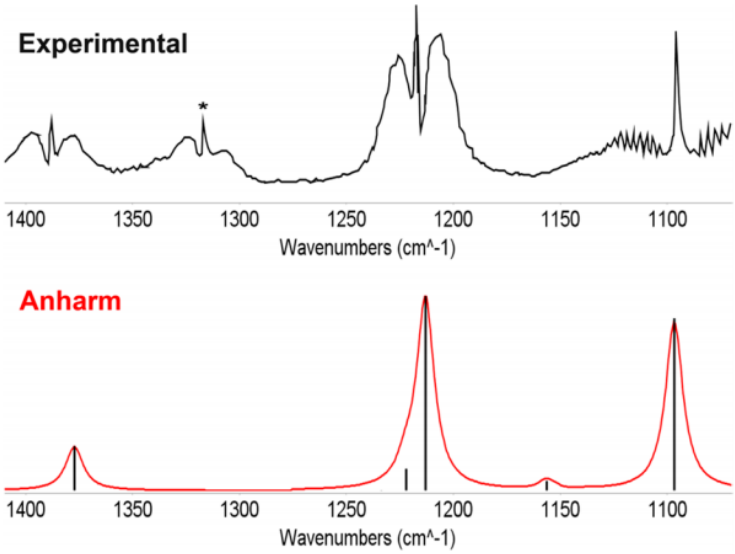



**Figure 1.1.3.** Observed FTIR (black, from Ref. [46]) and simulated IR spectra of *Z*-cyanomethanimine (B3LYP-D3/SNSD) at the anharmonic (red) level in the region 1070-1400 $cm^{-1}$. Normalization has been performed with respect to the peak above 1200 $cm^{-1}$.

Let us now focus attention on the 1070-1400 $cm^{-1}$ region. It turns out that not only the vibrational wavenumber of the C=N torsion is particularly well reproduced in these computations (1097 *vs.* 1095 $cm^{-1}$) but also its intensity. Indeed, the experimental intensity of the CNH bend (1218 $cm^{-1}$) is considerably higher than that of the C=N torsion. This trend is also reproduced in the simulation and with a very accurate wavenumber (1213 $cm^{-1}$). Finally, the low intensity peak at 1388 $cm^{-1}$ is also very well reproduced in our simulation.

### *1.1.3. Thermodynamics and kinetics*

The structures of the E and Z isomers were first optimized at several levels of theory. On those optimized geometries, analytical frequency calculations were performed at the harmonic and anharmonic levels, the latter one by using the generalized second-order vibrational perturbation model.[66] In Table 1.1.2 are listed the energy differences (electronic energy (ΔE), zero-point corrected energy ($\Delta E_0$), and Gibbs free energy (ΔG)) between both compounds at the harmonic level. This table shows a constant trend (the *Z* compound is always favored) and a negligible role of zero point and entropic effects.

This result matches with previous mixed computational and experimental investigations.[46,118] In order to understand the kinetics of the C-cyanomethanimine isomerization, the transition state between *Z* and *E* has been located and found *ca.* 25 kcal/mol higher than both stable isomers (Figure 1.1.4).

This TS exhibits a linear C-N-H angle (≈180°) and sustains a substantial imaginary frequency (-1134 $cm^{-1}$).

**Table 1.1.2.** Energy Differences – Electronic Energy, Zero-Point Energy, and Free Gibbs Energy – between *E* and *Z* Isomers at the harmonic level (kcal/mol).

| Δ: (E) – (Z) | **SNSD** | **m-aug-cc-pVTZ** | |
|---|---|---|---|
| | *B3LYP-D3* | *B2PLYP-D3* | *MP2* |
| ΔE | 0.64 | 0.63 | 0.61 |
| $\Delta E_0$ | 0.54 | 0.54 | 0.54 |
| ΔG | 0.59 | 0.59 | 0.58 |

On these grounds, a full kinetic investigation was undertaken in order to compute the rate constant. The densities of states of both isomers have been computed, as well as the

transition state connecting them, by convolution of the 3D classical rotational DOS and the anharmonic vibrational DOS, obtained by the Wang-Landau algorithm. As sketched in Section 1.1.1, the density of states permits a complete thermodynamic and kinetic characterization of the species. In particular, the enthalpy and entropy changes for the *Z*-to-*E* conversion are calculated as functions of temperature (Figures 1.1.5 and 1.1.6).

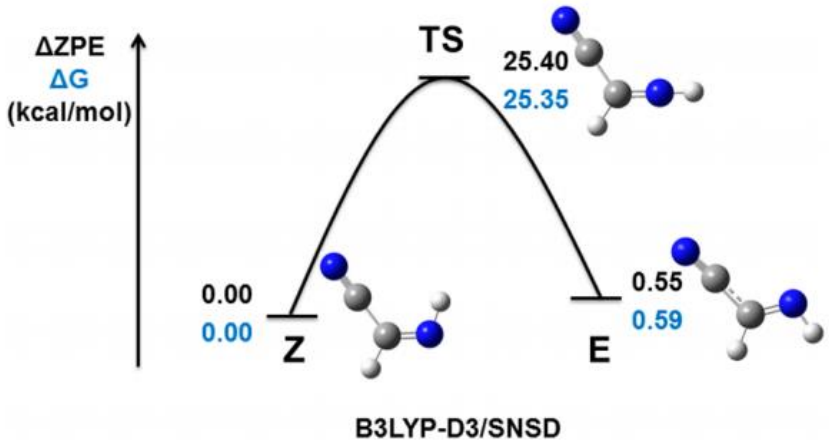



**Figure 1.1.4.** Relative energies (Zero-Point Energy -ZPE- and Gibbs Free Energy -ΔG-) diagram of the *Z*-C-cyanomethanimine ⇔ *E*-C-cyanomethanimine equilibrium calculated at the anharmonic B3LYP-D3/ SNSD level (kcal/mol).

In order to assess the effects of anharmonicity, these quantities are compared with the ones obtained using the harmonic vibrational density of states evaluated by the Stein-Rabinovitch modification of the Beyer-Swinehart algorithm.[119] It must be stressed that, even though the Wang-Landau algorithm is valid for any energy interval, the quadratic approximation for vibrational energy levels issuing from the VPT2 model (see eq. 1) loses reliability at high energies (corresponding to high temperatures) when the various dissociation limits of the vibrational modes are reached since, beyond the quadratic extremum, the energy tends to (unphysically) diminish with the quantum number. Consequently, as a reasonable compromise between reliability of the quadratic energy level formula and temperature range of the data reported, a temperature of 1500 K as our upper limit has been chosen. The trends of the two thermodynamic functions with temperature are very similar. This is to be expected, since the slope of the enthalpy change is the difference in

heat capacities between products and reactants, whereas the slope of the entropy change is the ratio of the same quantity with temperature. In terms of algebraic values, the reaction enthalpy is always positive. The main factor responsible for this is the higher zero-point energy of the *E* isomer. On the other hand, the entropy change is negative for the same reason, since fewer states are available to the *E* isomer. In the harmonic case, the enthalpy increases steadily and almost linearly with temperature, indicating an almost constant difference in heat capacities. This changes at very low temperatures, where the relative sparsity of E states diminishes its heat capacity and ultimately reverses the trend. Inclusion of anharmonicity has the effect of rendering the entropy change more negative, indicating that it increases the *Z* density of states more than the *E* one. Moreover, the difference in heat capacities between the two species becomes much more modest.

The effect of introducing anharmonicity in the density of states can be better evidenced in terms of the so-called anharmonic factor,[120] defined as the ratio of the corresponding densities of states

$$F(E) = \frac{\rho_{Anharm.}(E)}{\rho_{Harm.}(E)} \qquad (15)$$

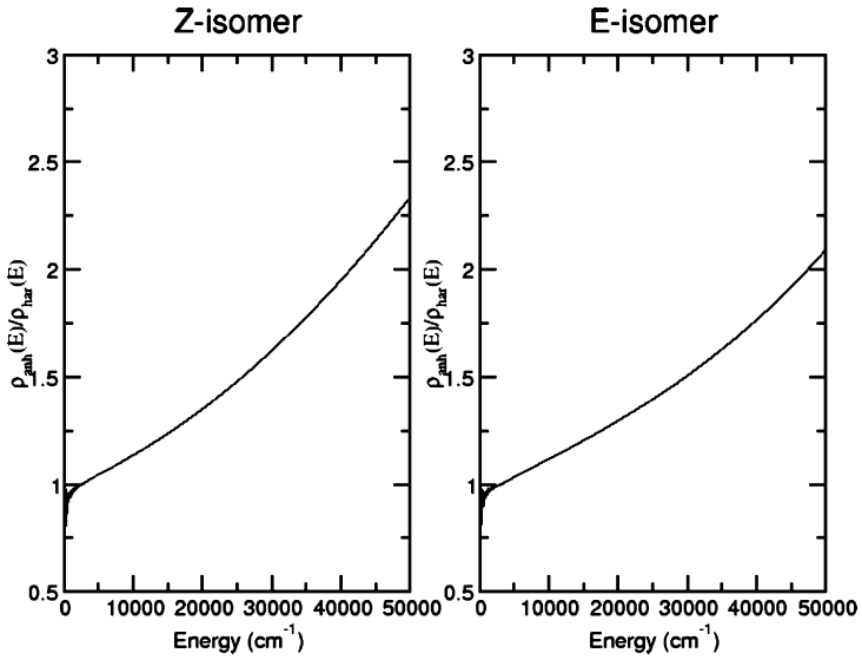



**Figure 1.1.5.** Evolution of the anharmonic factor, F(E), for *Z* and *E* isomers.

As shown in Figure 1.1.5, the anharmonic factor varies almost linearly with the energy. A similar trend was found by Troe *et al.*,[120] who compared harmonic and anharmonic densities of states for various small molecules. The high-frequency deviation in the low-energy region is related to the statistical sampling of the Wang-Landau algorithm. It can be seen that the slope in the case of the *Z* isomer is higher, indicating that anharmonicity increases the density of states more for the *Z* isomer than for its *E* counterpart. This trend is in agreement with the enthalpy and entropy curves shown above. Figure 1.1.6 shows the logarithm of the equilibrium constant for the $Z \rightarrow E$ isomerization. At all temperatures considered, the equilibrium constant is less than one, and this effect is more pronounced at low temperatures where the enthalpy term dominates. As the temperature increases, the difference in zero point energies becomes increasingly less important, and the equilibrium constant approaches unity.

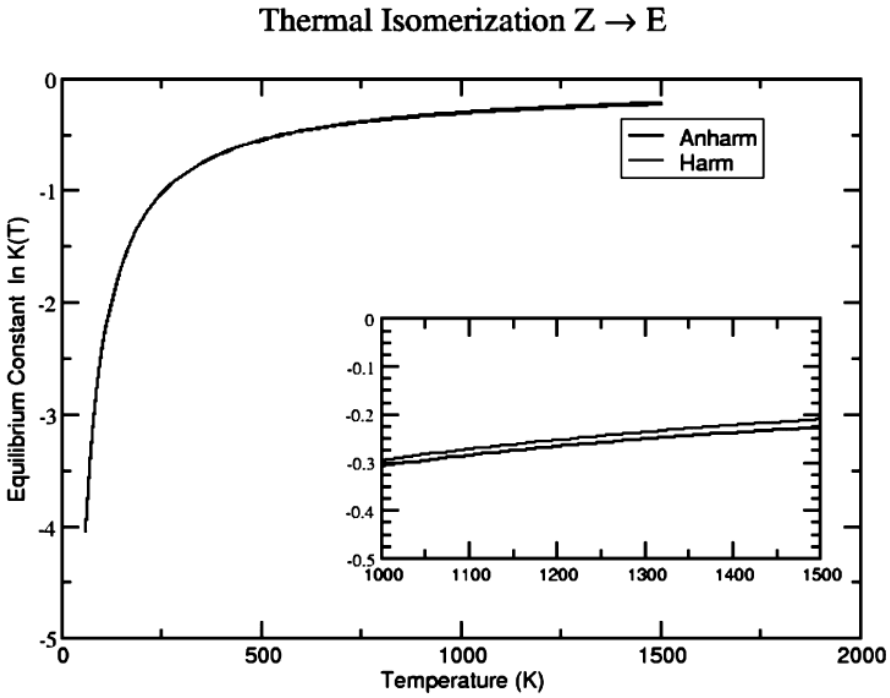



**Figure 1.1.6.** Logarithm of thermal equilibrium constant $K_{eq}(T)$ for the isomerization of cyanomethanimine.

Introducing anharmonicity slightly lowers the equilibrium constant. This appears principally to be an entropy effect since, as shown before, the entropy of the *Z* isomer is substantially larger than that of *E*, while the enthalpy contributions are very close.

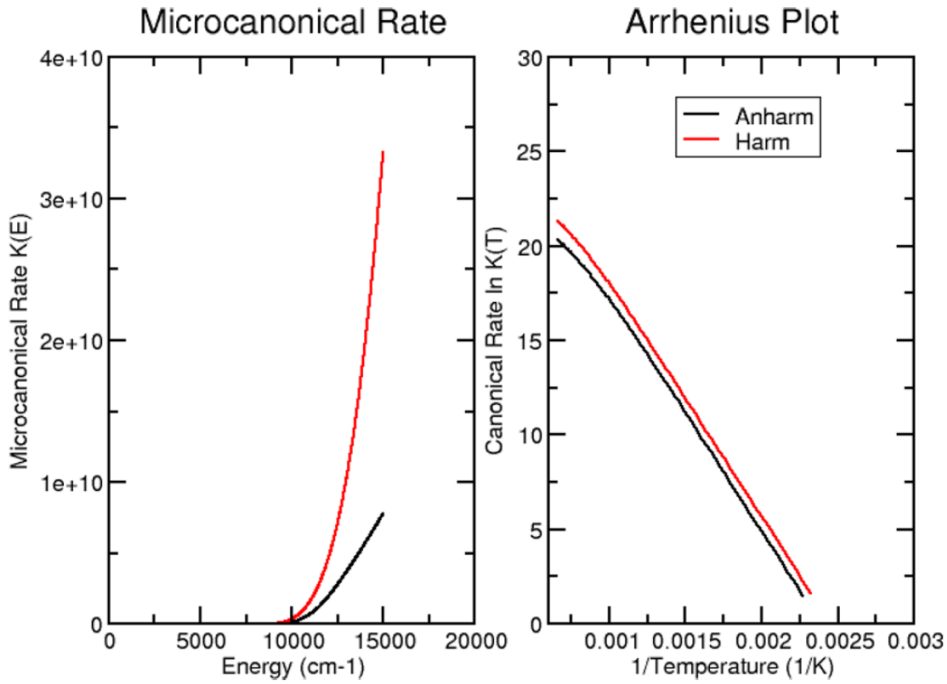



**Figure 1.1.7.** Microcanonical and canonical rate constants for the isomerization of C-cyanomethanimine.

Figure 1.1.7 shows the microcanonical and canonical rate constants for the direct reaction. Tunnelling and nonclassical reflection near the transition state have been included using the semiclassical approach of Miller and co-workers (even though, in this case, quantum effects are not expected to play any significant role, since the imaginary frequency of the transition state is only of the order of 1100 $cm^{-1}$).[105]

It can be seen that the anharmonic rates are lower than the harmonic ones at all energies and temperatures considered. It has already been shown how anharmonicity lowers the Gibbs free energy of the reactants. This effect more than compensates for the corresponding effect on the transition state, thus reducing the rate constant. From the Arrhenius plot, it is seen that the slope of the curve is essentially the same both in the harmonic and anharmonic case. Thus, the activation energy remains invariant on

introduction of anharmonicity, which seems primarily to decrease the pre-exponential factor. The activation energy calculated from the slope of the Arrhenius plot is 23.9 kcal/mol, which is in nice agreement with the potential barrier of the system (25.4 kcal/mol).

### *1.1.4. Partial conclusion*

New insights concerning the behavior of C-cyanomethanimine have been provided. The infrared spectrum of Z-C-cyanomethanimine has been simulated, and the anharmonic spectrum matches nicely the experimental one. Furthermore, vibrational energies of E-C-cyanomethanimine have also been assigned. The reliability and effectiveness of the proposed computational protocol pave the route for systematic investigations of more complex systems. All these results point out once again the non-negligible impact of anharmonicity on the chemical processes involving even small molecules.

## 1.2. Cyanomethanimine isomers in cold interstellar clouds: insights from electronic structure and kinetic calculations

As has been mentioned above, no molecular formation pathway is yet known for these compounds. This situation can be improved by quantum mechanical (QM) calculations of spectroscopic and thermodynamic properties to guide and support observations, line assignments, and data analysis.[121,122] The possible formation of cyanomethanimine in the ISM starting from the widespread HCN and HNC species was amply scrutinized in the past because it was suggested by Chakrabarti & Chakrabarti[123] that the oligomerization of HCN in the gas phase could lead up to adenine in four steps. Adenine is one of the purine RNA nucleobases and, having the gross formula $C_5H_5N_5$, can be viewed as a HCN pentamer. In the same fashion, cyanomethanimine is the dimer of HCN (or HNC, which is also ubiquitous in the ISM). The suggestion by Chakrabarti & Chakrabarti[123] was disproved by Smith *et al.*[124] who demonstrated that the first step of their sequence (that is, the formation of cyanomethanimine starting from two HCN molecules) is already impossible in the low-temperature interstellar environment. Later on, Yim & Choe[125] and Jung & Choe[126] also explored the possibility that the reaction between HCN or HNC with the protonated species $HCNH^+$ could lead to cyanomethanimine. Even though the presence of a proton substantially lowers the reaction barriers, Yim & Choe[125] concluded that those reactions are not viable gas-phase routes to cyanomethanimmine formation in the ISM.

Because of the lack of gas-phase routes, after the detection of E-cyanomethanimine toward Sgr B2(N), Zaleski et al.[106] suggested that a radical chemistry formation route could occur in interstellar ices. Here it is instead shown that the reaction between two widely diffuse species, that is, the cyano radical and methanimine, can easily account for cyanomethanimine formation under the characteristic conditions of interstellar clouds. The interaction of the ·CN radical with methanimine was already explored by Basiuk & Bogillo.[127] Their B3LYP/6-31++G(d,p) calculations demonstrated that the addition of the ·CN radical to the *p* bond of methanimine is barrierless but the reaction exit channels were not explored in that study. Nevertheless, in the low-density conditions of the ISM where no ternary molecular encounters take place, the stabilization of bound reaction intermediates could only take place

via radiative emission, a process with very low probability, which is not competitive with bond fission when exothermic fragmentation channels are available.

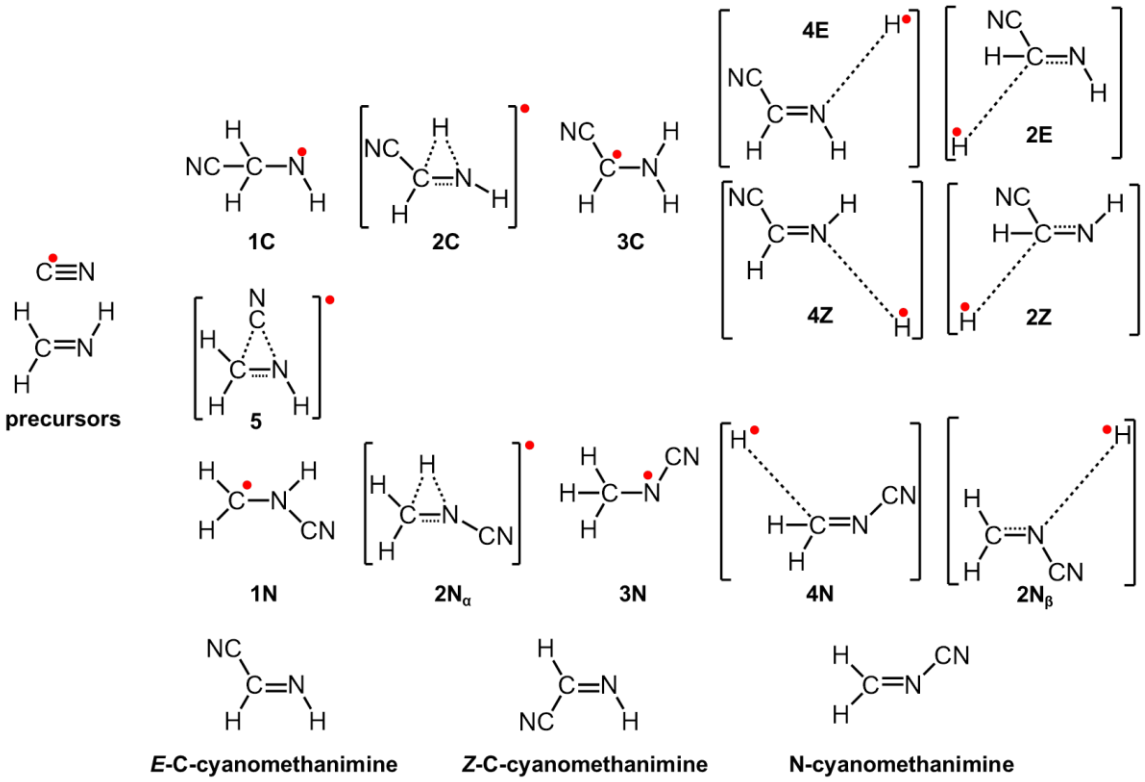



**Figure 1.2.1.** Investigated compounds.

Therefore, a detailed QM investigation has been undertaken, relying on some work already performed on the thermodynamic, kinetic, and spectroscopic features of cyanomethanimine derivatives (*E*-C- and *Z*-C-[70]). The vibrational and thermochemical signatures of all stable molecules herein presented and the transition states connecting them (Figure 1.2.1) were investigated, using one of the most reliable computational models rooted in density functional theory (DFT) coupled to single point energy refinement by a coupled cluster (CC) approach in conjunction with complete basis set extrapolation (CBS-QB3). Computed reaction rate coefficients are also provided using a capture model and the Rice-Ramsperger-Kassel-Marcus (RRKM) method for the open reaction channels, which are:

CN + $CH_2$=NH

→ *Z*-NC-CH=NH + H $\Delta H°_0$ = -18.5 kcal/mol (1)

→ *E*-NC-CH=NH + H $\Delta H°_0$ = -17.8 kcal/mol (2)

→ $CH_2$=N-CN + H $\Delta H°_0$ = -12.5 kcal/mol (3)

→ *Z*-CN-CH=NH + H $\Delta H°_0$ = -2.2 kcal/mol (4)

→ *E*-CN-CH=NH + H $\Delta H°_0$ = -1.5 kcal/mol (5)

(the indicated enthalpies of reactions are those calculated here, see below). As we are going to see, all E-C-, Z-C-, and N-isomers can be formed to some extent, but only E-C-cyanomethanimine has been identified so far in interstellar space. This fact needs to be critically discussed considering that the way used to characterize the compound was only through microwave spectroscopy. Indeed, as this technique strongly depends on the dipole moment of the molecule, this property is a factor which needs to be checked.

### *1.2.1. Specific computational details*

Most of the computations were performed with the B2PLYP-D3 functional[24] in conjunction with the m-aug-cc-pVTZ basis set,[26,27] where the *d* functions on hydrogens have been removed. The energies reported in this part (except where noted) were computed using the Complete Basis Set (CBS-QB3) method.[103,104]

On these grounds, rate constants have been obtained by a combination of capture theory and RRKM calculations using the same technique as described before. The initial bimolecular rate constant of CN + $CH_2NH$ is calculated using capture theory. Here, the assumption is that if the system has enough energy to surmount the centrifugal barrier, then it goes on to form one of the initial adducts. For a long range potential of the form $-C_6/R^6$, the capture cross section is given by the formula

$$\sigma(E) = \pi \times 3 \times 2^{-\frac{2}{3}} \times \left(\frac{C_6}{E}\right)^{1/3} \qquad (1)$$

where $E$ is the translational energy. The corresponding rate constant is given by the cross section multiplied by the collision velocity $(2E/m)^{1/2}$, where $m$ is the reduced mass of the reactants. It must be stressed that it is the translational (collision) rather than the total energy that appears in the capture rate expression. In order to account for the possibility that

the reactants occupy higher internal states (and, consequently, have a lower collision energy), a convolution of the raw capture rate constants with the probability density function for the internal energy (obtained from a simple calculation of the reactant internal density of states) was performed.

As the initial state of the reactants is assumed to be uncorrelated to the initial intermediate formed, the initial populations of the three possible intermediates are simply taken to be proportional to their respective densities of states at the total energy considered. Moreover, the capture model furnishes the maximum value of *J*, the total angular momentum quantum number, as a function of total energy, which is subsequently used in the unimolecular calculations.

In order to calculate the rate of back-dissociation into reactants, the principle of detailed balance was used and, for a reaction of the form A + B ↔ C, would read

$$k(A + B \rightarrow C)\ x\ \rho(A + B) = k(C \rightarrow A + B)\ x\ \rho(C) \qquad (2)$$

where ρ(C) and ρ(A+B) are, respectively, the density of states of intermediate C and the density of states per unit volume of the bimolecular system A+B (including the three-dimensional relative motion). Thus, the rate of back-dissociation into reactants was calculated by multiplying the bimolecular capture rate constant with the total density of states per unit volume of $CN+CH_2NH$ and dividing by the total density of states of the initial adducts (assuming, for each energy, a maximum value for the total angular momentum *J* given by the capture calculations).

After the initial intermediate is formed, the assumption is made that the available energy is statistically distributed among the various degrees of freedom (subject to conservation of total angular momentum) and therefore the RRKM model is suitable for the calculation of unimolecular rate constants. Moreover, in the RRKM model, the existence of a well-defined transition state is postulated for each elementary reaction, i.e., a configuration of "no recrossing" which, once crossed by the reactants, invariably leads to the products. The existence, in all cases, of a well-defined maximum in energy renders this assumption reasonable. The microcanonical rate constant for each elementary step is calculated using the formula

$$k(E) = \frac{N(E)}{h\rho(E)} \qquad (3)$$

where $N(E)$ denotes the sum of states in the transition state at energy $E$, $\rho(E)$ is the reactant density of states at energy $E$, and $h$ is Planck's constant. $N(E)$ is obtained by integrating the relevant density of states up to energy $E$ and the rigid rotor/harmonic oscillator model is assumed. Both densities of states (reactant and transition states) are appropriately symmetrized with respect to the number of identical configurations of the reactants and/or transition state. Where possible, tunneling and quantum reflection have been considered by using the corresponding imaginary frequency of the transition state and calculating the tunneling probability for the corresponding Eckart barrier. After all microcanonical rate constants have been calculated, the master equation is solved for the particular energy in order to take account of the possibility of interconversion between intermediates. A matrix $k$ of the rate constants is set up such that the off-diagonal element $k_{ij}$ represents the rate constant from species $j$ to species $i$ and the diagonal elements are such that the sum of each column is 0. Moreover, a concentration vector $c$ was set up such that the element $c_j$ corresponds to the concentration of channel $j$. Then, all kinetics expressions can be written as a vector differential equation

$$\frac{dc}{dt} = k \times c \qquad (4)$$

This is a linear differential equation and thus can be solved using standard methods. In order to determine the behavior of $c$ in the infinite future, the matrix $k$ was diagonalized and its eigenvectors were determined. These eigenvectors will either correspond to eigenvalues with a negative real part (vanishing in the infinite future) or 0 (stable eigenvectors). The initial concentration vector (as determined by capture theory among the three possible initial intermediates) is written as a linear combination of all eigenvectors and, subsequently, those with a negative eigenvalue are discarded. What remains is the concentration vector in the infinite future, yielding the branching ratios of all channels.

### *1.2.2. Electronic structure calculations*

Starting from the results obtained by Basiuk & Bogillo,[127] the structures of all the stationary points were reanalyzed at higher levels of theory, i.e., B2PLYP-D3/m-augcc- pVTZ

geometry optimizations with subsequent energy determinations also at the CBS-QB3 level. The reliability of the B2PLYP computational model is confirmed by the close correspondence of relative energies with their CBS-QB3 counterparts. All the precursors, intermediates, and products were fully characterized as minima on the potential energy surface and transition states exhibited a single imaginary frequency.

Figure 2.1.2 depicts possible reaction paths of the full ·CN + $CH_2NH$ → *E*-C-, *Z*-C- and N-cyanomethanimine formation reactions. Two pathways have been considered for each case, the first in one step (red) with one transition state (**2Z**, **2E** and **2N$_\beta$**), and the second (blue), which includes two intermediates (**1** and **3**) and two transition states (**2C**, **2N$_\alpha$** and **4**). Let us focus first on the pathway leading to the C-cyanomethanimines (*E*- and *Z*-). As one can see, the first step consists in ·CN attacking the carbon atom of $CH_2NH$, leading to the (NC)$CH_2N$·H radical **1C**, which is about 50 kcal/mol more stable. At this point, three possibilities can be envisaged. On the one hand, one might observe a loss of a hydrogen radical, leading directly to the *E* or *Z* isomer, depending on the original conformation of **1C**. This step has an exit barrier of about 37 kcal/mol. On the other hand, a migration of the CN moiety to the N atom can be realized through transition state **5** with a barrier of 25 kcal/mol. To finish, considering the presence of the stabilizing C≡N moiety on the carbon atom, hydrogen migration can also be observed in order to get the lone electron on this atom. This migration would occur through the transition state **2C** which has a barrier of ca. 32 kcal/mol, and would lead to the most stable compound involved in this reaction: the second intermediate **3C**. Basiuk & Bogillo[127] only studied this reaction up to this compound, which is a reactive radical.

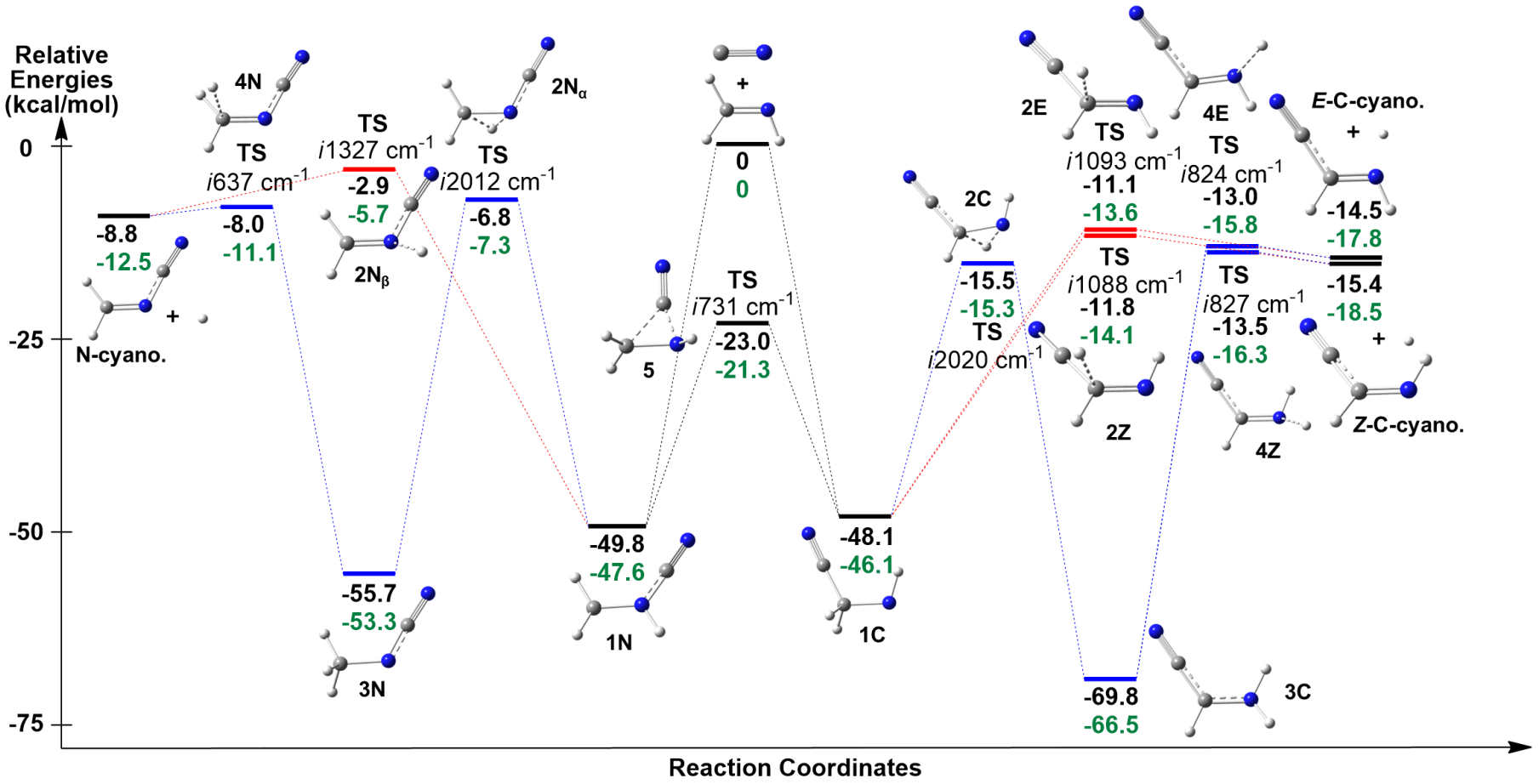



**Figure 1.2.2.** Full reaction paths of the reaction ·CN + $CH_2NH$ → *E*-C-, *Z*-C-, and N-cyanomethanimine with relative electronic energies (black) and zero-point corrected energy (green). The energies were computed using the CBS-QB3 level of theory. Alternative pathways in blue and red.

Therefore, they left room for continuing their study and suggested a deeper analysis including combination with several other radicals. However, they never mentioned the possible loss of a hydrogen radical, which we focus on here. This last step was found to be possible through another transition state with a significant barrier of about 56 kcal/mol, and leads either to *E*-C-cyanomethanimine or its *Z* isomer and a ·H radical. In the first case, the compounds are more stable than the reagents by 14.5 kcal/mol and in the second case by 15.4 kcal/mol. When concentrating on the pathway to N-cyanomethanimine, one can notice a similarity with the previous route. In that case, the ·CN attack occurs toward the nitrogen atom, leading to a first intermediate **1N** that is slightly more stable than the one on the other path (**1C**, of about 2 kcal/mol). From here, three possibilities are again envisaged: a C≡N migration to the C atom that has a 27 kcal/mol barrier, a direct loss of hydrogen through intermediate **2N**$_{\beta}$, which has a 47 kcal/mol barrier; and the presence of the stabilizing C≡N moiety on the nitrogen atom that would lead to a hydrogen migration, 43 kcal/mol.

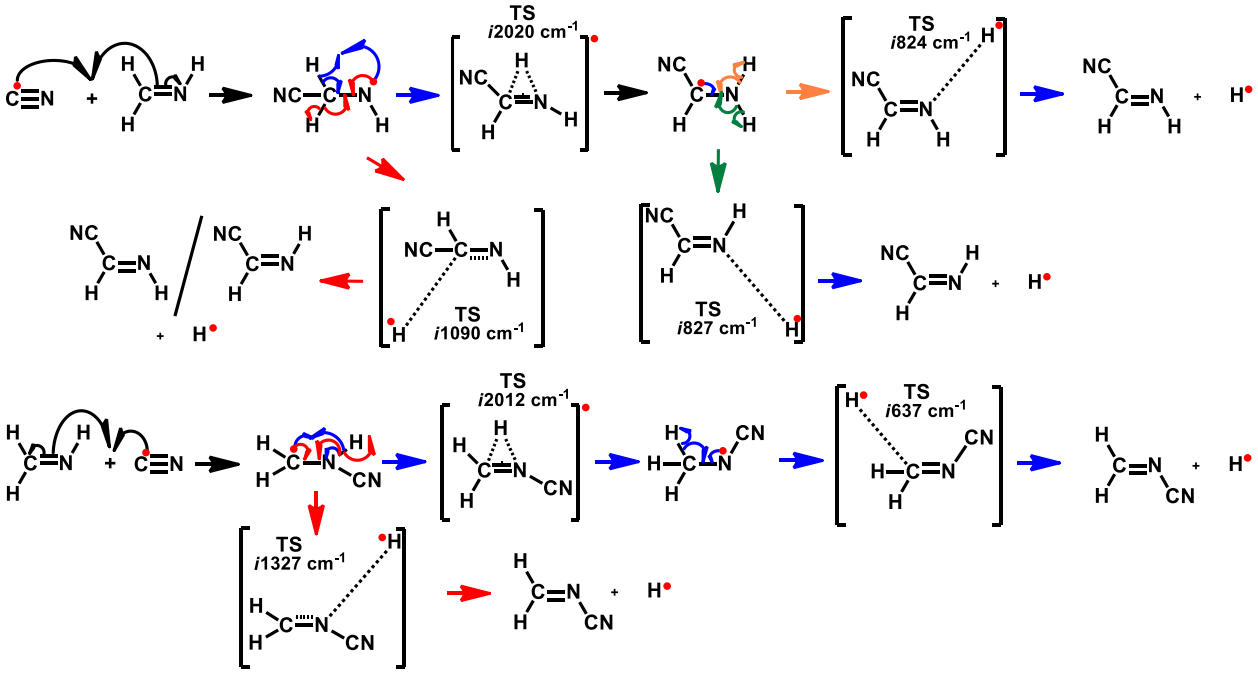



**Figure 2.1.3.** Possible formation mechanisms of E-, Z-C-cyanomethanimine and N-cyanomethanimine.

This second step, which needs to go through a first transition state **2N**$_{\alpha}$, requires more energy than its counterpart on the other path: 43 versus 32 kcal/mol. Moreover, the second intermediate **3N** is less stabilized than its counterpart **3C** due to a shorter delocalization chain. As far as the last step is concerned, one can notice the smaller barrier (48 versus 56 kcal/mol) that leads to a less stable final compound, *i.e.*, N-cyanomethanimine (9 versus 14.5 and 15.5 kcal/mol). These reaction paths are all feasible in the ISM considering that all the steps can be taken without exceeding the allowed energy given by the two precursors. Indeed, we found possible ways to form cyanomethanimine isomers in space. The possible radical mechanisms based on the reaction paths we proposed are given in Figure 2.1.3. The reaction path focusing on an attack of the ·NC radical leading to iso-cyanomethanimines was also considered but was too energetic to occur in the ISM.

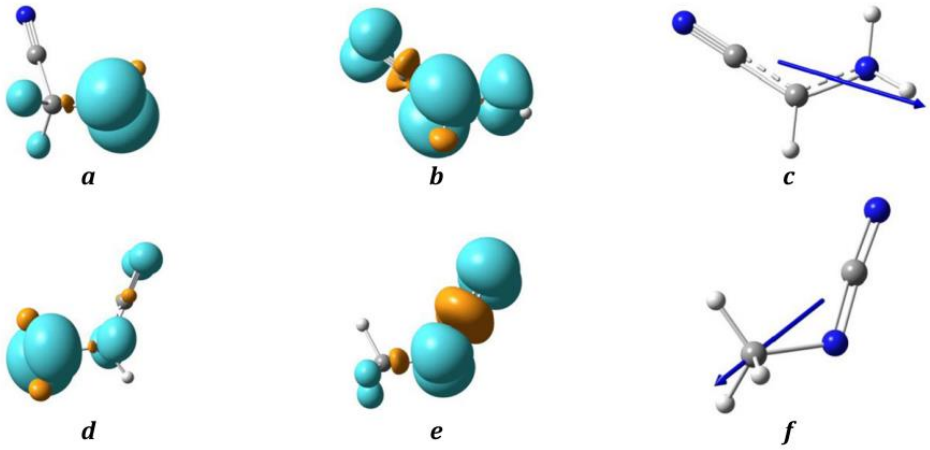



**Figure 1.2.4.** Spin densities of the intermediates **1C**(***a***), **3C**(***b***), **1N**(***d***), and **3N**(***e***) and dipole moments of **3C**(***c***) and **3N**(***f***).

In Figure 2.1.4, one can see the main localization of the spin densities and dipole moments of several compounds. On the first intermediates **1**, a major contribution can be seen on the atom not linked to C≡N, ensuring the presence of the lone electron on this atom. In the case of compound **3C**, one can observe the more delocalized spin density, which is linked to the delocalized free electron. The contribution of the carbon is the most important, however. This phenomenon is also noticeable on **3N**, but constrained to the N-C≡N moiety.

The dipole moments of **3C** and **3N**, as far as they are concerned, start from the averages of both conjugation chains.

After verifying that the formation of the three isomers is feasible in the ISM because of the lack of entrance channels, we consider here the further dissociation that *Z* and N could undergo because they are formed by very exothermic processes and are, therefore, characterized by a non-negligible excess of internal energy. However, these reactions would require too much energy to occur in the ISM. A cyclization has also been envisaged, but the resulting cycle (azirin-2-imine) would also be too energetic.

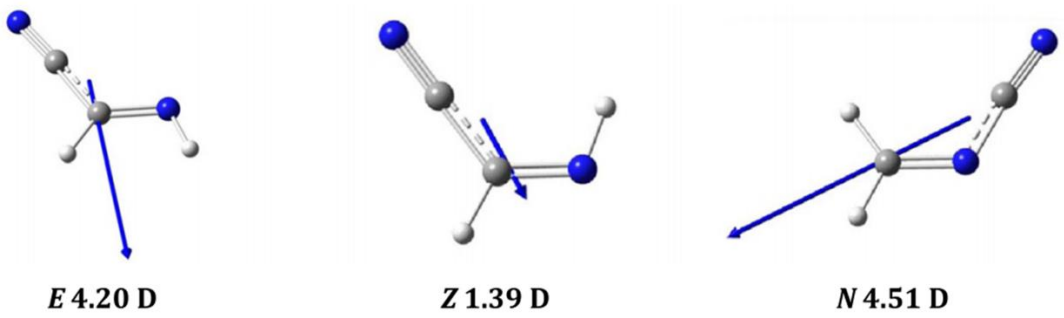



**Figure 1.2.5.** Dipole moments of E-C-, Z-C- and N-cyanomethanimine.

Let us now focus on the detection of the three isomers by microwave (MW) spectroscopy. As mentioned before, line intensities strongly depend on the dipole moments of the compounds of interest. Figure 1.2.5 depicts the dipole moment orientations and values of *E*-C, *Z*-C-, and N-cyanomethanimine. One can see that the *Z* isomer has a very small dipole moment compared to *E* and N, which could explain why it is difficult to detect by MW in space. Also, N-cyanomethanimine has not yet been characterized conclusively in space, although several rotational transitions that could correspond to this compound have been observed by Zaleski et al. (2013). In any case, this compound is not thermodynamically favored compared to its C-isomers (less stable by 6 kcal/mol), which would also explain a reduced presence in the ISM.

### *1.2.3. Kinetics calculations*

Calculations for total energies, up to 1.2 kcal/mol, have been performed (taking as zero the zero-point energy of the reactants CN + $CH_2NH$). As the most probable collision energy at 100 K is 0.2 kcal/mol, this range of energies seems to adequately describe the reaction scheme at the temperatures of the ISM.

As an initial step, the capture rate constant for the initial CN + $CH_2NH$ reaction was calculated. To do so, calculations for various points along the entrance coordinate CN-CH2NH were performed and the energies were subsequently fitted to a $V = V_0 - C_6/R^6$ law to determine the $C_6$ constant which was used in the capture theory calculation. There are three possible initial adducts. Two of them are cyano adducts (*i.e.*, CN adds through its C atom, either on the C or the N atom of the methanimine) and one is an isocyano adduct (CN adds through its N atom on the C atom of the methanimine). The initial overall capture rate constant has been partitioned among the three adducts in proportion to their densities of states at the energy considered. It must be stressed that it is the translational (collision) rather than the total energy that appears in the capture rate expression.

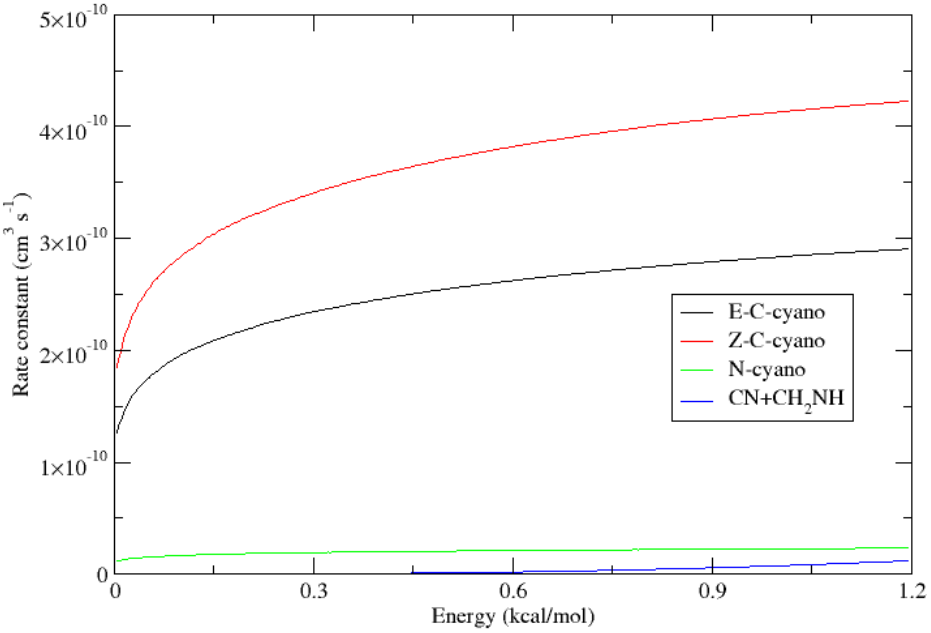



**Figure 1.2.6.** Rate constants for the formation of final products.

The final product-specific rate constants appear in Figure 1.2.6 for all energies considered. At all energies, the rate constant for the production of iso-cyano products was found to be negligible. This fact is easy to understand since, as can be seen from the path, the energies for all transition states from the initial iso-cyano adduct are well above the reactants and hence the only contribution from these is due to tunnelling. Moreover, the density of states of the initial iso-cyano adduct is about two orders of magnitude lower than the others. Regarding the products observed, it is seen that the rate constant is maximum for the *Z* product with the *E* product following (the respective branching ratios are around 57% and 39%). The reason for the prevalence of the *Z* product is the slightly lower energy of the respective transition states compared to the *E* ones. On the other hand, the branching ratio of the N product is only around 4%. This is in spite of the fact that the density of states of the initial N-adduct is higher than that of the initial C-adduct and thus its initial population is higher. As can be seen, the barriers for hydrogen migration and elimination from the N-adduct are much higher than those for the C-adduct. Moreover, the barrier for interconversion between C- and N-adduct is considerably lower.

As a result, the initial populations do not play much of a role. It is rather the subsequent barriers that determine the final outcome. Finally, back-dissociation into reactants is negligibly slow at low energies, whereas it reaches a branching ratio of around 1% to 1.5% at the highest energy. It can be seen from Figure 1.2.6 that the rate constant rises rather steeply towards the high energy end, the reason for this being the high density of states of the reactants (which comprise five rotational degrees of freedom).

### *1.2.4 Partial conclusion*

New insights into the formation of cyanomethanimine isomers in the ISM have been provided. Electronic structure and kinetic calculations demonstrate that the reaction CN + $CH_2$=NH is a facile formation route of *Z*,*E*-C-cyanomethanimine, even under the extreme conditions of density and temperature typical of cold interstellar clouds. Therefore, in all cases where the CN radical and methanimine are relatively abundant, the observation of some cyanomethanimine can be expected as well. It can also be noted that, since the formation of the more stable *Z* isomer is favored by a factor of *ca.* 1.5 and since the detection

of the *E* isomer is favored by a factor of 6,[106] the missing detection of the *Z* isomer can be due to the sensitivity limit of the GBT PRIMOS survey, as already suggested by Zaleski et al.[106] N-cyanomethanimine is formed with a much lower probability and its detection will be extremely challenging.

Finally, since both CN radicals and methanimine are important species in the upper atmosphere of Titan (with methanimine being formed by the reaction $N(^2D)$ + $CH_4$ and $C_2H_6$,[111] we suggest that the CN +$CH_2NH$ reaction could significantly contribute to the build-up of the nitrogen-rich organic aerosols that cover the massive moon of Saturn.

A more general conclusion is that neutral-neutral gas-phase reactions can account for the formation of relatively complex organic molecules even under the extreme conditions of the ISM. Other studies of critical and yet unexplored neutral–neutral gas phase reactions performed with theoretical methods can help to fill the gap of the missing reactions leading to COMs, especially when those reactions cannot be easily investigated in laboratory experiments as in the present case.

# *Chapter 2. The peculiar case of formamide*

Formamide, as the simplest amide, has attracted increasing attention in the field of prebiotic chemistry.[128–131] Indeed, it has the ability to act as a precursor in the abiotic amino acid synthesis and perhaps also in that of nucleic acid bases.[10–12] Formamide is, therefore, a central compound to connect metabolism (conversion of energy), which is ruled by proteins, and genetics (passage of information), ruled by RNA and DNA. Moreover, its key role in the interstellar medium appears to be also remarkable since it has been detected in the galactic center sources Sgr A and Sgr B2,[132] in the Orion-KL region (an active site of high-mass star formation embedded in OMC-1)[133] and more recently in a solar-type protostar.[134] These detections of centimeter- and millimeter-wave emissions due to rotational transitions have led prebiotic chemists to wonder how formamide has been formed in space.

## 2.1. State-of-the-art thermochemical and kinetic computations for formamide formation in cold interstellar clouds

Several formation pathways have already been investigated for formamide in the gas phase, including ionic reactions or radical ones.[135] However, while efficient and successful surface-chemistry mechanisms have been proposed,[136–139] none of these gas-phase reaction channels is likely to be open in space, *i.e.* at low temperatures. Indeed, the first step might exhibit a barrier and would require a non-available amount of energy, a third body collision could be needed which is not likely in the gas phase, or the path could lead mainly to other products.

Therefore, to solve this issue, a comprehensive quantum mechanical investigation has been undertaken, focusing on the addition of the $^{\bullet}OH$ radical to methanimine, as suggested in a study by Ali and Barker.[140] When looking carefully to the system, one can envisage an additional plausible path involving the addition of the $^{\bullet}NH_2$ radical to formaldehyde, already suggested in a study by Kahane and co-workers.[134] Moreover, this channel was also identified by Ali and Barker as a product channel, but they did not analyze it as a reactant channel. In detail, the vibrational signatures and thermochemical properties of all stable molecules involved in these processes and the transition states connecting them (see Figure 2.1.1) were investigated using state-of-the-art computational approaches. Then, reaction rates were

computed using capture and Rice-Ramsperger-Kassel-Marcus (RRKM) theories within a general semiclassical approach including anharmonic contributions together with tunneling and nonclassical reflection effects.

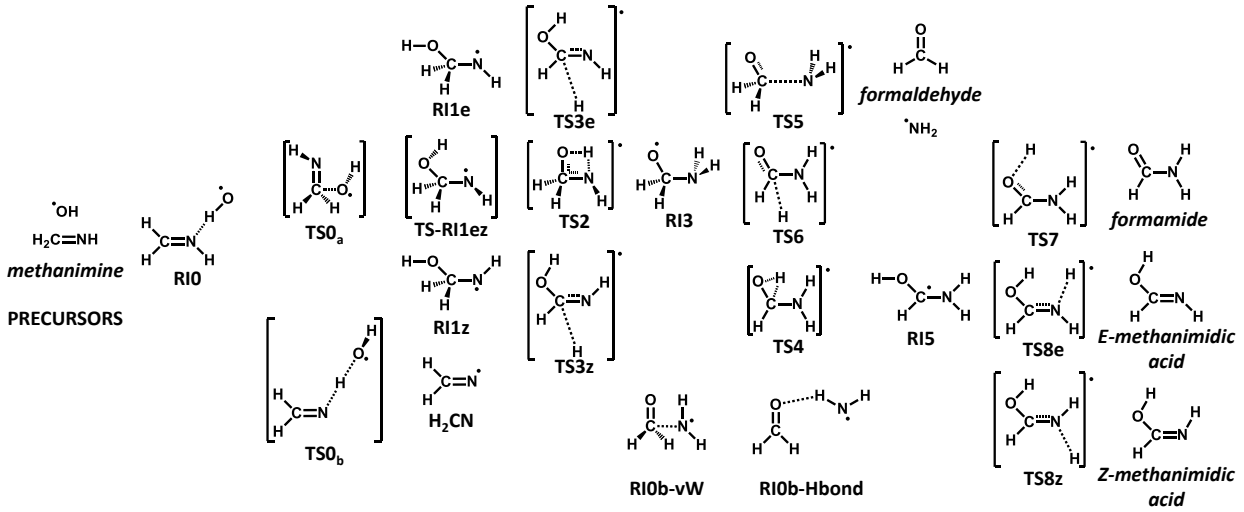



**Figure 2.1.1.** Investigated compounds.

## *2.1.1. Specific computational details*

Computations were mainly performed with the B2PLYP-D3 functional,[24] in conjunction with the m-aug-cc-pVTZ basis set,[26,27] where *d* functions on hydrogens have been removed.

For all stationary points, improved electronic energies were obtained by means of a composite approach based on coupled cluster theory employing the singles and doubles approximation (CCSD) augmented by a perturbative treatment of triple excitations, CCSD(T),[141,142] and implemented in CFOUR.[143,144] The composite approach is based on the additivity approximation and accounts for the following contributions: the Hartree-Fock self-consistent-field (HF-SCF) energy extrapolated to the complete basis-set (CBS) limit, the valence correlation energy at the CCSD(T) level extrapolated to the CBS limit as well, and the core-valence correlation correction.

The correlation-consistent cc-p(C)VnZ (n = T-5) basis sets[27,145] were used throughout. The so-called CCSD(T)/CBS+CV energy is therefore given by

$$E_{CBS+CV} = E^{CBS}(HF-SCF) + \Delta E^{CBS}\big(CCSD(T)\big) + \Delta E(CV) \qquad (1)$$

In this equation, the exponential extrapolation formula of Feller[146] for the HF-SCF energy and the $n^{-3}$ extrapolation scheme for the CCSD(T) correlation contribution[147] were employed. In the formula above, n = T, Q and 5 were chosen for the HF-SCF extrapolation, whereas n = T and Q were used for CCSD(T). Core-valence correlation effects were included by adding the corresponding energy correction, ΔE(CV), obtained as the difference between the all-electron and frozen-core CCSD(T) energies using the core-valence cc-pCVTZ basis set. For selected intermediates and transition states, n = Q and 5 were also used for the extrapolation to the CBS limit of the CCSD(T) correlation energy, and the cc-pCVQZ basis set for evaluating the ΔE(CV) correction. For the **TS5** and **TS6** transition states, to further inspect the small magnitude of the barrier with respect to formaldehyde and $NH_2$, the effects due to a full treatment of triple (fT) and quadruple excitations (fQ) were also accounted for. The corresponding corrections were obtained as differences between CCSDT and CCSD(T) and between CCSDTQ and CCSDT calculations within the frozen-core approximation employing the cc-pVTZ and cc-pVDZ basis sets, respectively. To this purpose, single-point CC singles, doubles, and triples (CCSDT)[148,149] as well as CC singles, doubles, triples, and quadruples (CCSDTQ)[150] energy calculations were carried out using the MRCC package by Kàllay[151] interfaced to CFOUR.

At the VPT2 level, a hindered rotor model was used for those intermediates that involve large amplitude internal rotations. In such cases, the vibrational Hamiltonian has been assumed to be adiabatically separable in a set of coupled anharmonic oscillators and a Large Amplitude Motion (LAM), *i.e.* the hindered rotation. It is well known that this "forced" separation leads to a variation of the effective moment of inertia along the path.[152] On the other hand, recovering this kinetic coupling introduces a coordinate dependent mass that makes the Hamiltonian non Hermitian[153] and this generates some theoretical and numerical ambiguities.[154] In order to include the effective mass variation, a generalized coordinate approach was used, which leads to a unitary reduced mass.[155–157] In practical terms, a relaxed scan along the dihedral angle that defines the internal rotation was performed, starting from an initial geometry, the translational and rotational interactions of the successive geometry was minimized, and the length of the curve was computed as the distance in massweighted

Cartesian coordinates between pairs of successive geometries. Finally, the curve obtained with a cubic spline was fitted to obtain more points, and then the Hamiltonian matrix obtained in a Discrete Variable Representation (DVR) was diagonalized.[158]

The semiclassical definition of reaction probability proposed by Miller,[105] which includes both tunneling and anharmonicity, was used in the present study for the evaluation of reaction rates. Full kinetic calculations were performed by an in-house code described in some papers.[159–161] The initial bimolecular rate constant leading from the reactants to the intermediate has been evaluated using capture theory calculations.[162] In particular, calculations were performed at various long-range distances of the reactants, and the energies obtained were fitted to a $1/R^6$ functional form (both for the London dispersion forces and the rotating dipole ones) which was used to perform the capture calculation. It is assumed that each successful capture leads to the intermediate. It is expected that using, alternatively, variational transition state theory (VTST) would lead to more accurate results. However, judging from the monotonicity of the potential energy curve and the absence of inner transition states, a computationally much less expensive scheme such as capture theory can confidently give reliable results.

As far as dissociation back to reactants is concerned, we have used a detailed balance argument, whereby the unimolecular rate constant for back-dissociation is given by the equation

$$k_{back}(E) = k_{capt}(E)\frac{\rho_R(E)}{\rho_I(E)} \qquad (3)$$

where $k_{capt}$ is the capture rate constant, $\rho(R)$ is the density of states per unit volume for the reactants, and $\rho(I)$ is the density of states for the intermediate. On the other hand, where, as opposed to the reactants, a well-defined transition state exists, a RRKM calculation has been performed. The microcanonical rate constant is calculated using the formula

$$k(E) = \frac{N(E)}{h\rho(E)} \qquad (4)$$

where $N(E)$ denotes the sum of states in the transition state at energy $E$, $\rho(E)$ is the reactant density of states at energy $E$, and $h$ is Planck's constant. At the harmonic level, simple algorithms are available for computing number and densities of states, whereas tunneling and quantum reflection have been treated by a simple Eckart barrier parametrized from the

energies of reactants and products, together with energy and imaginary frequency of the transition state.

More generally, *N(E)* is given by

$$N(E) = \sum_{n_1=0} \sum_{n_2=0} \cdots \sum_{n_{F-1}=0} P_n(E) \qquad (5)$$

with the semiclassical tunneling probability expressed by

$$P_n(E) = \frac{1}{1+e^{2\vartheta(n,E)}} \qquad (6)$$

In the above equations, *F*-1 vibrational degrees of freedom (*F* being the number of internal degrees of freedom, *i.e.* 3*N*-6 where *N* is the number of atoms) are orthogonal to the reaction path. In the framework of VPT2, $\vartheta(n,E)$ is an explicit function of the harmonic and anharmonic force constants.[105] Using the so-called Wang-Landau algorithm,[117] it becomes possible to evaluate effectively both $\rho(E)$ and $N(E)$ by separating the energy range in small bins (with width $\delta E$), as shown explicitly for $N(E)$ in eq. 7

$$N(E) = \delta E \sum_{i=1}^{E/\delta E} \rho(E_i)\langle P(E_i)\rangle \qquad (7)$$

where $\langle P(E_i)\rangle$ is the average tunneling probability for the different vibrational quantum number choices leading to $E \in [E_i - \delta E, E_i + \delta E]$.

The corresponding canonical expression for the rate constant is

$$K(T) = \frac{1}{h} \frac{\int_{-\infty}^{+\infty} N(E) e^{-E/k_B T} dE}{Q_{re}(T)} \qquad (8)$$

where $Q_{re}(T)$ is the reactants partition function.

In agreement with the principle of detailed balance, both in the canonical and microcanonical ensemble, the equilibrium constant is given by the ratio of the two rate constants (direct and inverse). Both densities of states (reactant and transition state) were appropriately symmetrized with respect to the number of identical configurations of the reactants and/or transition state.

After all calculations were performed, for each energy the branching ratio among all possible products and backdissociation was determined, and the corresponding capture rate constant was multiplied by this ratio to give the rate constant for the formation of each

product. Finally, the rate constants were Boltzmann-averaged in order to provide the rate constants as a function of temperature.

### *2.1.2. Validation of structural and vibrational results*

Recently, reliable and accurate semiexperimental equilibrium structures ($r_e^{SE}$) have been obtained for a set of more than 50 molecules containing the most important building blocks of organic and biological systems.[163] From that compilation it is possible to estimate the accuracy of equilibrium structures obtained by means of different computational methods. Table 2.1.1 shows that the B2PLYP double hybrid functional in comparison to the typical global hybrid ones halves the root-mean-square error (RMSE) with respect to $r_e^{SE}$ and approaches the level of accuracy of the most sophisticated post-Hartree-Fock models. Errors comparable to those obtained at the CCSD(T) level are noted, unless extrapolation to the CBS limit and core-correlation corrections are accounted for in conjunction with the latter. These errors on geometries lead to negligible errors on computed energetic and thermodynamic properties, as for instance demonstrated by Refs [164], [165], and [166].

**Table 2.1.1.** Average root-mean-square errors of different geometrical parameters[a] with respect to the semiexperimental equilibrium values of the B3SE set of Ref [163].

| | **B3LYP** | **B2PLYP** |
|---|---|---|
| *C-O* | 0.006 | 0.003 |
| *C-N* | 0.005 | 0.002 |
| *C-H* | 0.007 | 0.001 |
| *N-H* | 0.007 | 0.001 |
| *angles* | 0.41 | 0.26 |

[a]Distances in Å and angles in degrees.

A further check of the reliability of the computational model (*i.e.*, B2PLYP-D3/m-aug-cc-pVTZ) has been performed by comparing the simulated and experimental infrared (IR) spectra of formamide (see Figure 2.1.2).[167]

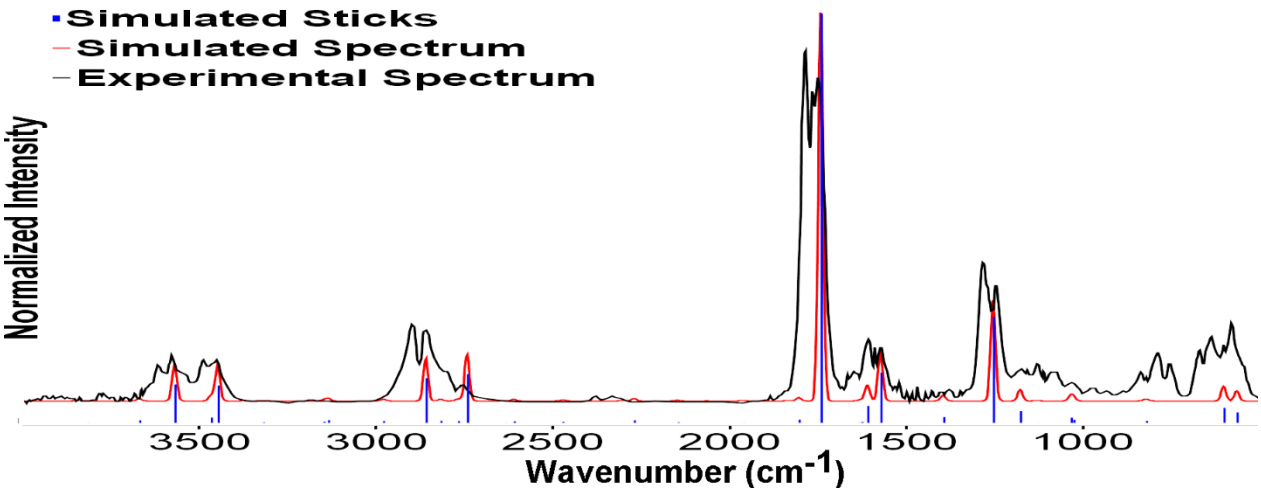



**Figure 2.1.2.** Simulated anharmonic (red, B2PLYP-D3/m-aug-cc-pVTZ) and experimental (black) IR spectra of formamide in the gas phase. Experimental spectrum taken from Ref [167].

As one can see, the simulated spectrum fits nicely with the experimental one: such a good agreement gives further confidence to investigate the present reaction path with the chosen level of theory.

### *2.1.3. Choice of methodology*

In the paper by Ali and Barker,[140] optimized geometries of several compounds at both the BHandHLYP/aug-cc-pVTZ and CCSD(T)/aug-cc-pVTZ levels of theory have been reported. Thanks to these data, a further validation of the B2PLYP-D3/m-aug-cc-pVTZ geometries is possible. Indeed, as we will see later, the optimized geometries of the loose transition states strongly depend on the computational level, whereas those of the tight transition states and intermediates do not exhibit large changes when using different methods.

To further benchmark the accuracy and reliability of the B2PLYP-D3/m-aug-cc-pVTZ energies, the CCSD(T)/CBS+CV approach has been employed, which is known to provide gas-phase thermodynamic properties with a 1 kJ/mol accuracy when used in conjunction with DFT optimized geometries.[164–166] It has also been demonstrated that geometries optimized at the B3LYP/SNS,[164] B2PLYP/aug-cc-pVTZ,[166] and CCSD(T)/ccpVTZ[168] levels are suitable for fulfilling such an accuracy.

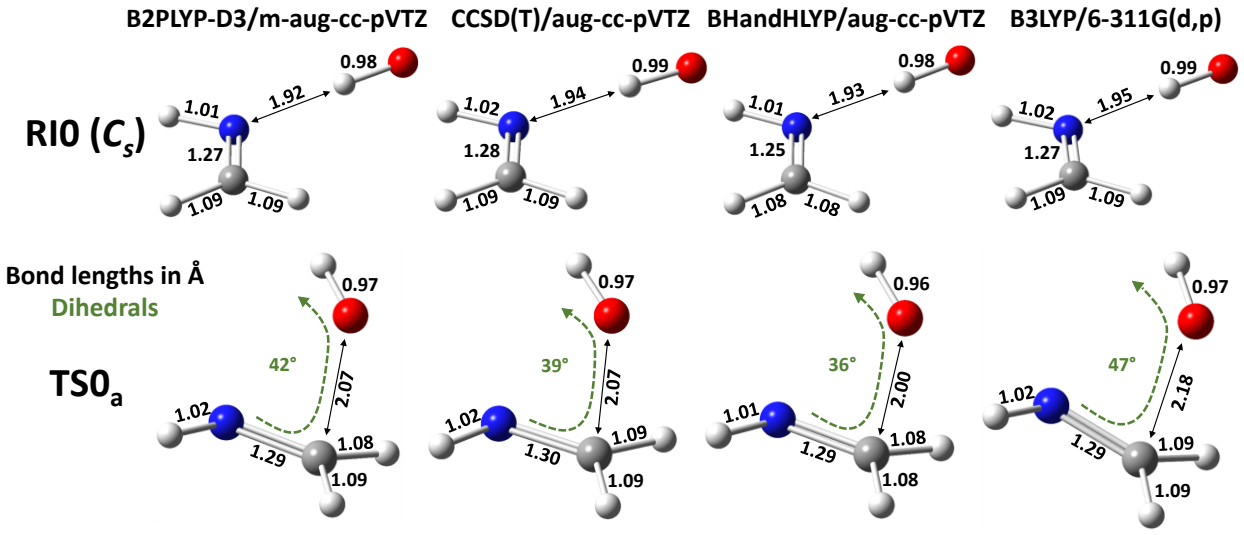



**Figure 2.1.3.** Comparisons between the optimized geometries of **RI0** and **TS0a** obtained at the B2PLYP-D3/m-aug-cc-pVTZ, CCSD(T)/aug-cc-pVTZ, BHandHLYP/aug-cc-pVTZ and B3LYP/6-311G(d,p) levels of theory.

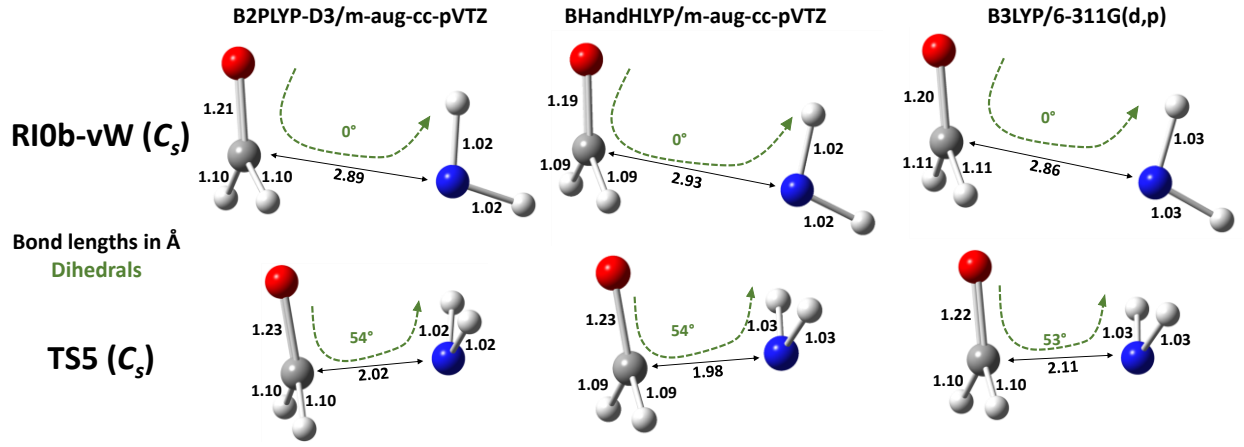



**Figure 2.1.4.** Comparisons between the optimized geometries of **RI0b-vW** and **TS5** obtained at the B2PLYP-D3/m-aug-cc-pVTZ, BHandHLYP/aug-ccpVTZ, and B3LYP/6-311G(d,p) levels of theory.

Figure 2.1.3 shows the comparisons of the optimized geometries at different levels of theory for the **RI0** and **TS0a** species (see Figure 2.1.1). The molecular structures will be discussed in detail later, but at this stage it is noteworthy that **RI0** involves a hydrogen bond

between •OH and methanimine, while **TS0$_a$** is a weak van der Waals complex. In the case of **RI0**, one can see that the optimized geometries obtained with the different levels of theory are quite similar to one another. However, when looking at **TS0$_a$**, the BHandHLYP functional underestimates the distance between •OH and methanimine, while B3LYP overestimates it compared to CCSD(T). Inclusion of semiempirical dispersion (D3) and of a fraction of MP2 contributions (B2PLYP-D3 functional) restores a remarkable agreement with CCSD(T).

In Figure 2.1.4, the comparison between the optimized geometries of **RI0b-vW** and **TS5**, which are both van der Waals complexes, obtained with the BHandHLYP, B2PLYP-D3, and B3LYP functionals is depicted. Once again, it is apparent that the geometries of the intermediate optimized at the different levels of theory considered are quite similar, while the BHandHLYP geometry of the transition state underestimates the distance between the approaching partners (in this case $NH_2$ and formaldehyde). In this connection, it is noteworthy that also B3LYP provides inaccurate geometries for **TS5**, but in this case the distances between $NH_2$ and formaldehyde are overestimated. This trend is possibly related to the percentage of HF exchange included in the different global hybrid functionals.

With these two verifications, together with the results discussed in the preceding section and several recent studies,[25,169,170] it is possible to be fully confident about geometrical structures and nonpotential energy contributions (*i.e.*, zero point energies, entropies, etc.) obtained at the B2PLYP-D3/m-aug-cc-pVTZ level. Then, as mentioned above, electronic energies have been improved by means of CCSD(T) computations (CCSD(T)/CBS+CV).

### *2.1.4. Mechanistic study*

**Approach •OH and methanimine**

In Figure 2.1.5, one can see plausible paths concerning the approach of •OH to methanimine, with the selected energies collected in Table 2.1.2. Starting from the separated reactants, a complex stabilized by the formation of a hydrogen bond (**RI0**), more stable by 30.4 kJ/mol than the reactants, is formed. If we add zero-point corrections, this complex is found to be more stable than the precursors by 22.2 kJ/mol.

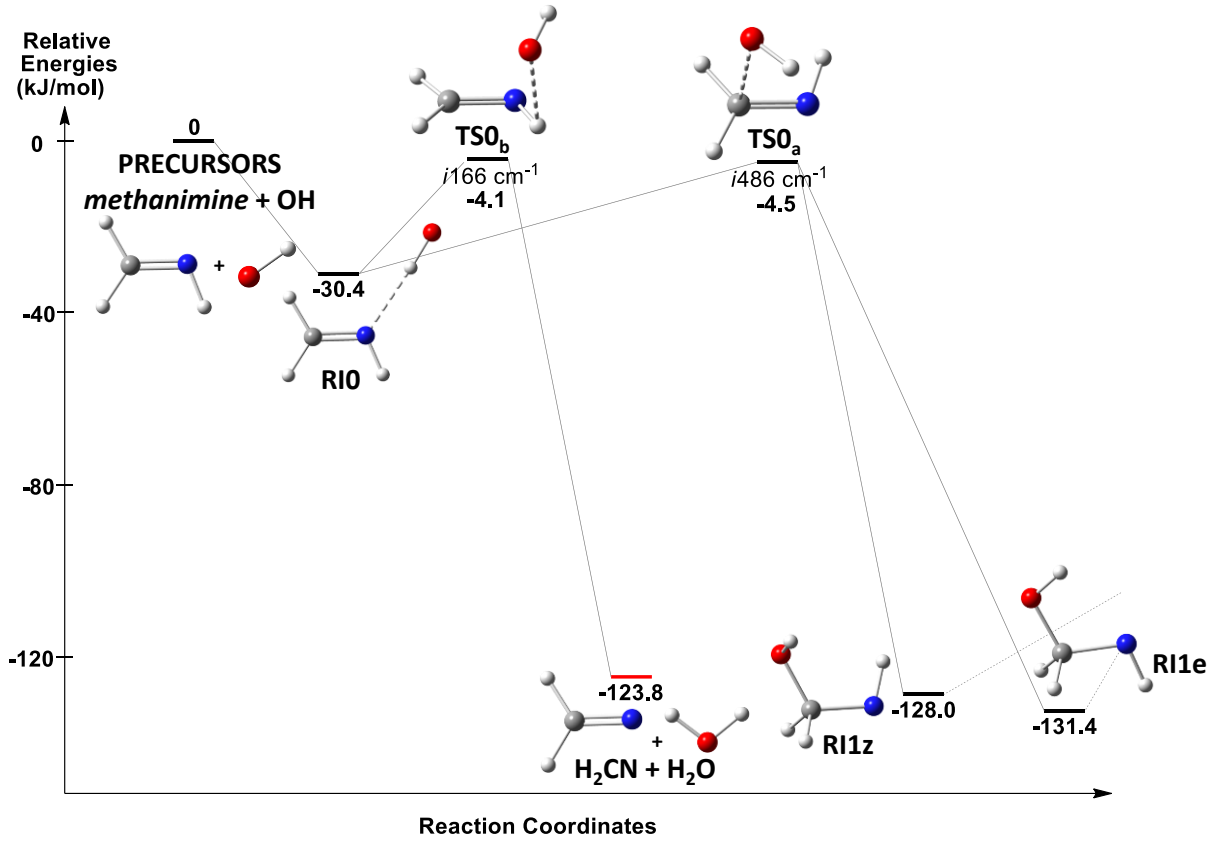



**Figure 2.1.5.** Proposed path for the approach of OH· and methanimine. Relative electronic CCSD(T)/CBS+CV energies (in kJ/mol) in black.

**Table 2.1.2.** Relative energies[a] (in kJ/mol, with respect to methanimine + OH) of **RI0**, **TS0a** and **TS0b** obtained at different levels of theory.

| | *B2PLYP-D3/ m-aug-cc-pVTZ* | *CCSD(T)/ m-aug-cc-pVTZ* | *CCSD(T)/ CBS(TQ)+CV(CT)*[b] | *CCSD(T)/ CBS(Q5)+CV(CQ)*[c] |
|---|---|---|---|---|
| ***∑precursors*** | 0 | 0 | 0 | 0 |
| ***RI0 (Cs)*** | -32.09 | -29.37 | -30.38 | -30.42 |
| ***TS0a*** | -7.53 | -0.33 | -4.52 | -5.19 |
| ***TS0b*** | -8.28 | -2.34 | -4.14 | -4.02 |
| ***RI1z*** | -125.99 | - | -128.03 | - |
| ***RI1e*** | -128.87 | - | -131.38 | - |

[a]Electronic energies computed at the corresponding B2PLYP/m-aug-cc-VTZ optimized geometries.
[b]The cc-pVTZ and cc-pVQZ basis sets were used for the extrapolation to the CBS limit and the cc-pCVTZ set for the core-correlation correction.
[c]The cc-pVQZ and cc-pV5Z basis sets were used for the extrapolation to the CBS limit and the cc-pCVQZ set for the core-correlation correction.

The comparison of our result with others obtained at the CCSD(T) level deserves to be briefly addressed. In the paper by Ali and Barker,[140] CCSD(T)/aug-ccpVTZ calculations (carried out at geometries optimized at the same level) gave a ZPE-corrected relative energy of -19.8 kJ/mol for this compound, while Bunkan et al.[171] report a CCSD(T)/CBS(T,Q) ZPE-corrected value of -23.8 kJ/mol (at CCSD(T)/cc-pVTZ optimized geometries). In particular, we note a difference of 1.6 kJ/mol between our CCSD(T)/CBS+CV value and the Bunkan et al. CCSD(T)/CBS result. Such a small discrepancy can be ascribed to the missing consideration of core correlation in Ref [171] as well as the use of geometries not entirely suitable (*i.e.*, optimized using a basis set -cc-pVTZ- without diffuse functions).

Two transition states, **TS0$_a$** or **TS0$_b$**, with relative electronic energies of -4.5 and -4.1 kJ/mol (5.0 and -3.3 with ZPE corrections here, 2.8 and -2.8 in Ref [140]), respectively, can follow this first step and govern hydrogen abstraction and addition of OH to the π bond of methanimine, leading to the intermediates **RI1** (**RI1z** or **RI1e**) and to $H_2CN$ + $H_2O$, respectively.

Under such circumstances, capture theory seems to be appropriate for describing this step, considering the monotonicity of the entrance potential energy curve and the absence of inner transition states.

**Proposed full OH + $CH_2NH$ reaction path**

Figure 2.1.6 depicts the possible paths of the $CH_2NH$ + •OH reaction (skipping details about the approach step) and reports the relative electronic (at the CCSD(T)/CBS+CV level) and zero-point corrected energies (electronic CCSD(T)/CBS+CV energies + ZPE issuing from B2PLYP-D3/m-aug-cc-pVTZ anharmonic vibrational calculations) of all minima and transition states. Together with the channels shown in Figure 2.1.6, the addition of OH to the N atom of methanimine was also investigated, with other possible abstractions of hydrogen atoms (linked to C atom). However, the energies of the transition states governing the first steps were so high (of the order of 15 kJ/mol) that it was decided to skip those channels.

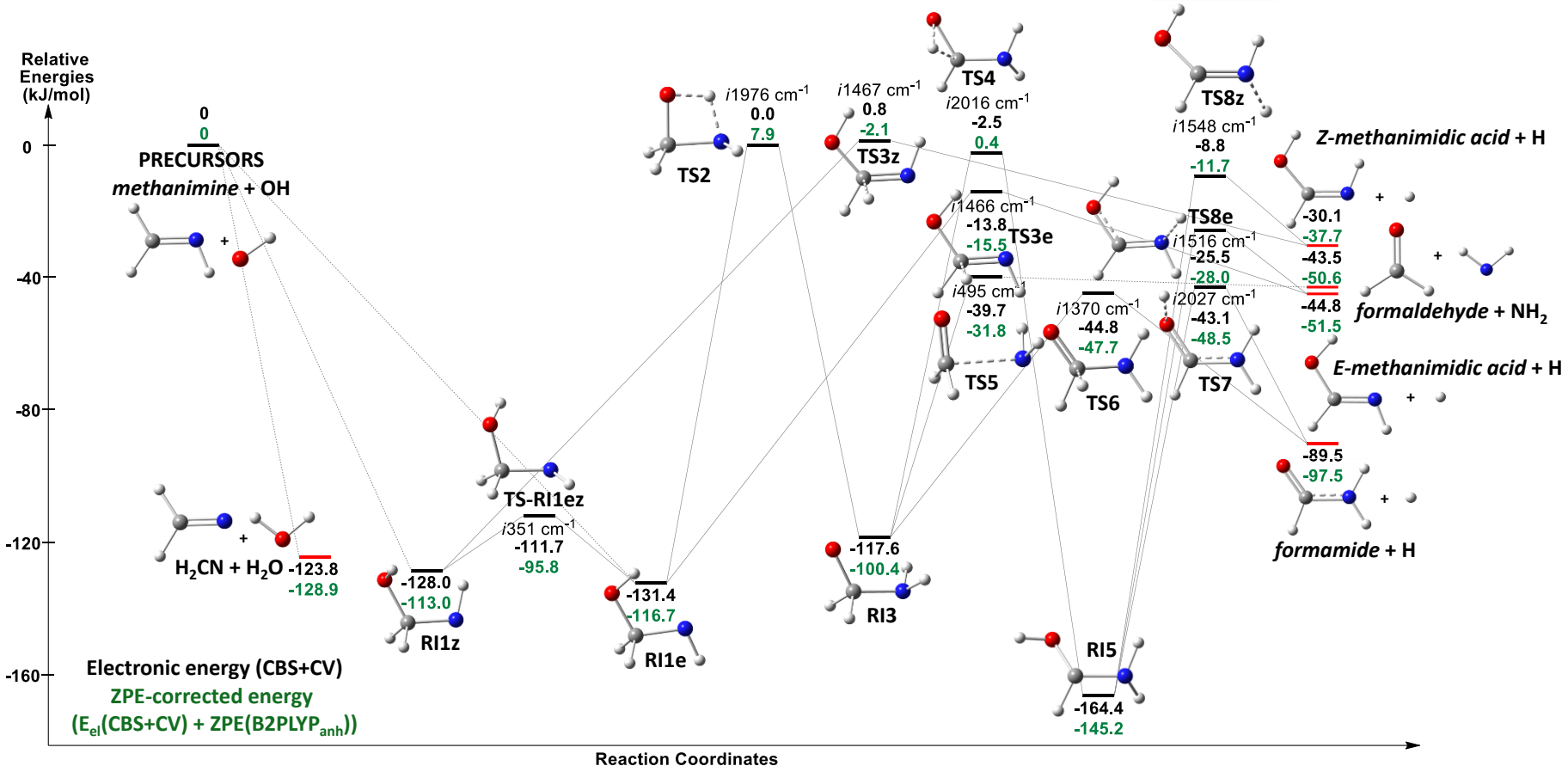



**Figure 2.1.6.** Proposed full reaction path. Electronic energies (black) are at the CCSD(T)/CBS+CV level and the ZPE corrected energies (green) are obtained by including the ZPE issuing from B2PLYP-D3/m-aug-cc-pVTZ anharmonic vibrational calculations. All energies are given in kJ/mol.

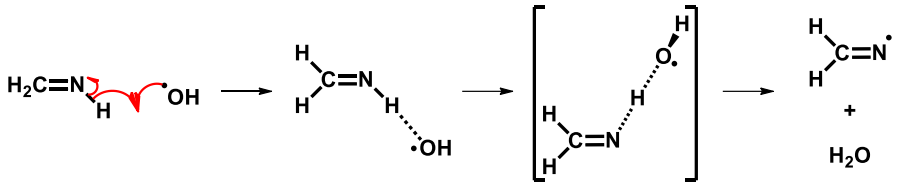


**Figure 2.1.7.** Possible radical mechanism for H abstraction.

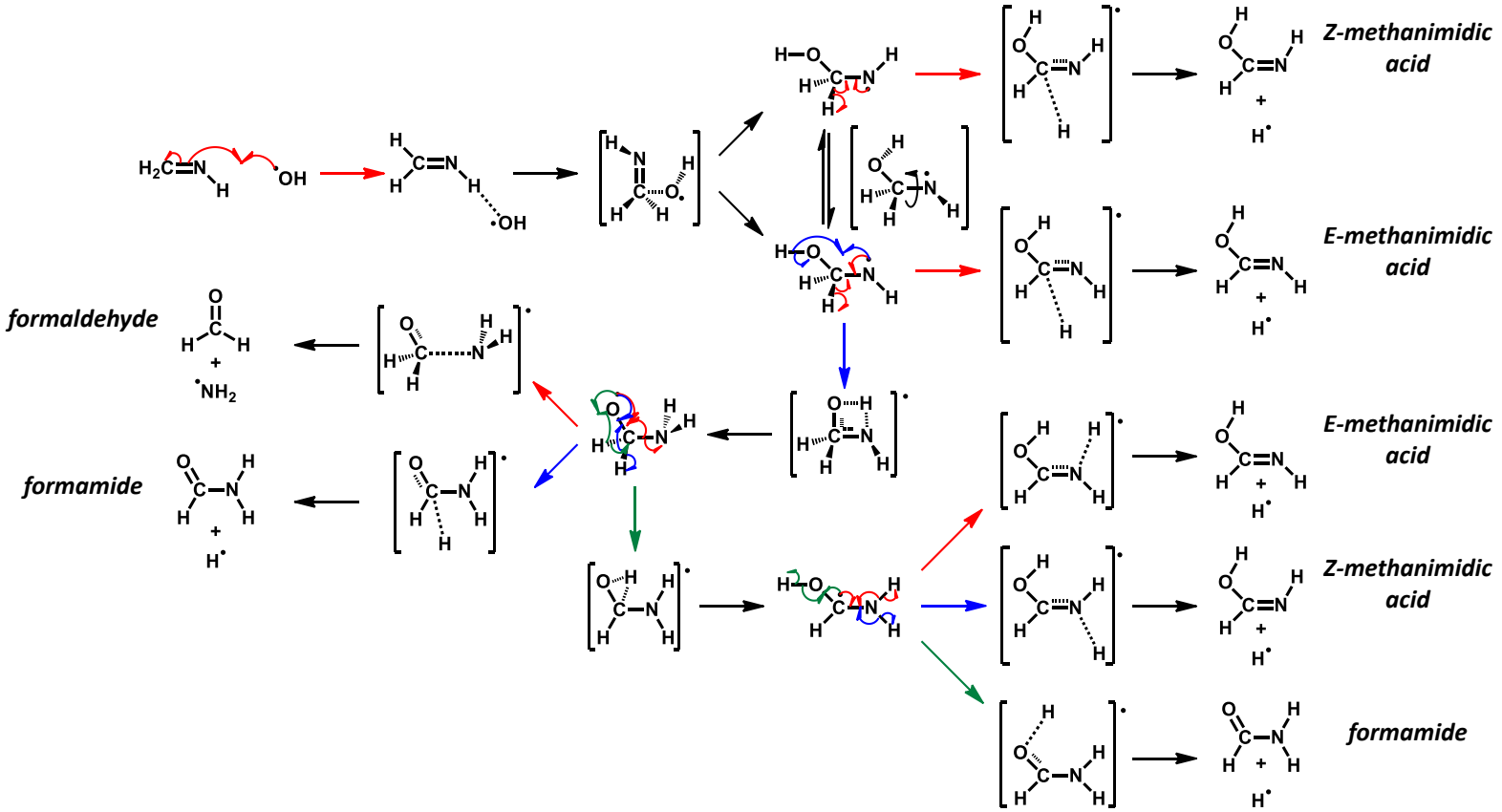


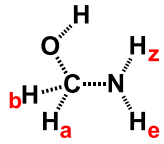

**Figure 2.1.8.** Possible radical mechanism for the addition of OH· to methanimine.

Starting from $CH_2NH$ and $^{\bullet}OH$, a hydrogen abstraction from $CH_2NH$ by $^{\bullet}OH$ can be observed, thus leading to $H_2CN^{\bullet}$ and $H_2O$ that are around 120 kJ/mol more stable than the reactants. An addition of $^{\bullet}OH$ to the π bond of $CH_2NH$ can also occur, thus giving two potential intermediates, **RI1e** or **RI1z**, which are respectively 131.4 and 128.0 kJ/mol more stable than the precursors and connected by a weak barrier (*ca.* 5 kJ/mol, through **TS-RI1ez**). Then, both **RI1** intermediates can lose a hydrogen atom, thus leading directly to the methanimidic acid isomers, through *ca.* 125 and 120 kJ/mol barriers for the *Z*- and *E*-isomers, respectively. In the case of **RI1e**, the hydrogen atom can migrate from oxygen to nitrogen, to give the intermediate **RI3**. This step exhibits a ~125 kJ/mol barrier. At this point, three reaction channels can be followed. The first one is a lengthening of the C-N bond that would lead to the formation of $^{\bullet}NH_2$ and formaldehyde through an ~80 kJ/mol barrier. The second one involves the loss of the hydrogen atom linked to carbon, thus leading directly to formamide through a slightly lower barrier (by ~4 kJ/mol). As far as the third reaction path is concerned, the hydrogen atom linked to carbon migrates to the oxygen atom. This step exhibits a *ca.* 113 kJ/mol barrier and leads to the intermediate **RI5**. The latter is then able to lose the hydrogen atom bonded either to nitrogen or to the oxygen atom. The first possibility would give *E*- or *Z*-methanimidic acid, through *ca.* 138 and 155 kJ/mol barriers, respectively, and the second would lead to formamide, through a ~120 kJ/mol barrier. If we focus on the possible products, formamide + H is the most stable one, with a relative energy of -89.5 kJ/mol. The following ones are *E*-methanimidic acid + H and formaldehyde + $NH_2$, at -44.8 and -43.5 kJ/mol, respectively, and then the less stable product is *Z*-methanimidic acid + H that exhibits a relative energy of -30.1 kJ/mol. It is noteworthy that all compounds involved in the reaction scheme of Figure 2.1.6 are lower in energy than the reactants, which makes the entire path viable in the interstellar medium.

The mechanisms corresponding to these paths are depicted in Figures 2.1.7 and 2.1.8.

Let us now focus on the structural variations along the reaction paths. Table 2.1.3 reports the bond lengths and symmetry point group of all investigated compounds obtained from B2PLYPD3/m-aug-cc-pVTZ geometry optimizations. One can first notice that the C-O and C-N bonds fluctuate significantly, while the O-H, N-H, and C-H bonds remain close in distance (except for the transition states corresponding to H migrations and loss).

**Table 2.1.3.** Bond lengths (Å) and symmetry point groups of the investigated compounds[a]




| | C-O | C-N | N-$H_e$ | N-$H_z$ | O-H | C-$H_a$ | C-$H_b$ | Sym. |
|---|---|---|---|---|---|---|---|---|
| *methanimine* | - | 1.27 | 1.02 | - | - | 1.09 | 1.09 | $C_S$ |
| *OH* | - | - | - | - | 0.97 | - | - | $C_{\infty V}$ |
| **RI1z** | 1.42 | 1.41 | - | 1.02 | 0.96 | 1.10 | 1.10 | $C_S$ |
| **TS-RI1ez** | 1.42 | 1.45 | 1.02 | | 0.96 | 1.09 | 1.09 | $C_1$ |
| **RI1e** | 1.41 | 1.42 | 1.02 | - | 0.96 | 1.10 | 1.10 | $C_S$ |
| **TS2** | 1.42 | 1.47 | 1.02 | 1.27 | 1.24 | 1.09 | 1.09 | $C_1$ |
| **RI3** | **1.37** | **1.44** | **1.01** | **1.01** | **-** | **1.10** | **1.10** | ***$C_S$*** |
| **TS3z** | 1.36 | 1.29 | - | 1.02 | 0.97 | 1.71 | 1.08 | $C_1$ |
| **TS3e** | 1.35 | 1.29 | 1.02 | - | 0.97 | 1.71 | 1.09 | $C_1$ |
| **TS4** | 1.38 | 1.38 | 1.01 | 1.01 | 1.22 | 1.29 | 1.09 | $C_1$ |
| **TS5** | 1.23 | 2.02 | 1.02 | 1.02 | - | 1.10 | 1.10 | $C_S$ |
| **TS6** | **1.24** | **1.38** | **1.01** | **1.01** | **-** | **1.56** | **1.10** | ***$C_1$*** |
| **RI5** | 1.39 | 1.39 | 1.01 | 1.01 | 0.96 | 1.08 | - | $C_1$ |
| **TS7** | 1.25 | 1.35 | 1.00 | 1.01 | 1.38 | 1.09 | - | $C_1$ |
| **TS8z** | 1.36 | 1.27 | 1.64 | 1.02 | 0.96 | 1.08 | - | $C_1$ |
| **TS8e** | 1.34 | 1.28 | 1.01 | 1.62 | 0.97 | 1.08 | - | $C_1$ |
| *$H_2CN$* | - | 1.24 | - | - | 1.09 | 1.09 | - | $C_{2V}$ |
| *H2O* | - | - | - | - | 0.96 | - | - | $C_{2V}$ |
| *Z-methan. ac.* | 1.36 | 1.26 | - | 1.02 | 0.97 | 1.08 | - | $C_S$ |
| *E-methan. ac.* | 1.35 | 1.26 | 1.01 | - | 0.97 | 1.09 | - | $C_S$ |
| *formaldehyde* | **1.21** | **-** | **-** | **-** | **-** | **1.10** | **1.10** | ***$C_{2V}$*** |
| *$NH_2$* | **-** | **-** | **1.03** | **1.03** | **-** | **-** | **-** | ***$C_{2V}$*** |
| *formamide* | **1.21** | **1.36** | **1.00** | **1.00** | **-** | | **-** | ***$C_S$*** |

[a]Optimizations are at the B2PLYP-D3/m-aug-cc-pVTZ level. The values related to the $NH_2$ + formaldehyde reaction are in bold.

First, let us discuss the C-O bonds. In formaldehyde and formamide, the C-O distances, about 1.21 Å, are those typical of a double bond. In the methanimidic acid isomers, C-O bonds are single bonds, linked to a C-N double bond, thus exhibiting a length of 1.35 Å. If we analyze the intermediates and transition states, we can classify C-O bond lengths into three types. The first type, with C-O bond lengths around 1.25 Å, includes **TS5**, **TS6**, and **TS7**. All three transition states lead directly to formaldehyde and formamide and exhibit therefore a formal C-O double bond, with a weak influence of the hydrogen atom. The second type, with C-O bond distances of about 1.35 Å, is present in **RI3**, both **TS3** transition states, **TS4**, and both **TS8** transition states. The four **TS3** and **TS8** compounds evolve in methanimidic acid isomers and exhibit therefore similar C-O bond lengths. The C-O bond in **RI3** consists of a $CO^{\bullet}$ radical

moiety, and in **TS4** the hydrogen atom is migrating from C to the O atom. This explains why in these compounds the C-O distance is in between a single and a double bond. The last class, with C-O bond lengths around 1.40 Å, includes both **RI1** isomers, the transition state linking them (*i.e.*, **TS-RI1ez**), **TS2**, and **RI5**. In all these five compounds, an OH moiety is linked to a C-N single bond; therefore, the C-O bond is a formal single bond as well.

Let us now consider how the C-N bond varies along the reaction. Once again, three types of bonds are observed. The first one, with the C-N bond lengths around 1.25 Å, is present in compounds that exhibit a formal double C-N bond: methanimine, both **TS3** transition states, both **TS8** transition states, $H_2CN$, and both methanimidic acid isomers. The second type is characterized by C-N bond lengths of about 1.37 Å. It is found in **TS4**, **TS6**, **TS7**, and formamide, that is to say in compounds with a delocalization that makes the C-N bond lying somewhat in between a single and a double bond. The last class of the C-N bond, which includes both **RI1** isomers, the connecting transition state (*i.e.*, **TS-RI1ez**), **TS2**, **RI3**, and **RI5**, exhibits a formal single bond of about 1.45 Å.

As far as symmetry is concerned, it is noticeable that out of 14 intermediates and transition states, only 3 compounds have a $C_S$ group symmetry. However, all the precursors and products have relatively high symmetries ($C_S$, $C_{\infty v}$, $C_{2v}$).

**Approach $^{\bullet}NH_2$ and formaldehyde**

Let us now consider an alternative path involving $NH_2$ and formaldehyde as precursors for formamide.

As in the case of the previous approach, CCSD(T)/CBS+CV relative energies are considered together with those at the B2PLYP-D3/m-aug-cc-pVTZ and CCSD(T)/m-aug-cc-pVTZ levels. The most relevant results are collected in Table 2.1.4. As one can see in Figure 2.1.9, starting from the reactants either a weak van der Waals complex (**RI0b-vW**) or an intermediate stabilized by a hydrogen bond (**RI0b-Hbond**) can be formed. Both of them are followed by a low lying, first order saddle point (**TS5**) having an energy very close to that of reactants, which then leads to an intermediate (**RI3**). Once again, one can notice that this first step, leading to **RI3**, is nearly barrierless, our best estimate for the transition state height with respect to reactants being 3.8 kJ/mol. Inclusion of the full-T (CBS+CV+fT) and full-Q (CBS+CV+fT+fQ) corrections further lowers this barrier to 2.05 and 1.67 kJ/mol, respectively.

Again, capture theory is used to describe this step.

**Table 2.1.4.** Relative energies[a] (in kJ/mol, with respect to formaldehyde+$NH_2$) of **RI0b-vW**, **RI0b-Hbond, TS5** and **RI3** obtained at different levels of theory.

| | *B2PLYP-D3/* m-aug-cc-pVTZ | *CCSD(T)/* m-aug-cc-pVTZ | *CCSD(T)/ CBS(TQ)+CV(CT)*[b] | *CCSD(T)/ CBS(Q5)+CV(CQ)*[c] |
|---|---|---|---|---|
| ***∑precursors*** | 0 | 0 | 0 | 0 |
| ***RI0b-vW (Cs)*** | -12.72 | -11.63 | -12.18 | -12.21 |
| ***RI0b-Hbond (Cs)*** | -15.69 | -15.15 | -15.90 | -15.86 |
| ***TS5 (Cs)*** | 0.67 | 7.45 | 3.64 | 3.77 |
| ***RI3 (Cs)*** | -71.55 | - | -74.48 | - |

[a]Electronic energies computed at the corresponding B2PLYP/m-aug-cc-VTZ optimized geometries.

[b]The cc-pVTZ and cc-pVQZ basis sets were used for the extrapolation to the CBS limit and the cc-pCVTZ set for the core-correlation correction.

[c]The cc-pVQZ and cc-pV5Z basis sets were used for the extrapolation to the CBS limit and the cc-pCVQZ set for the core-correlation correction.

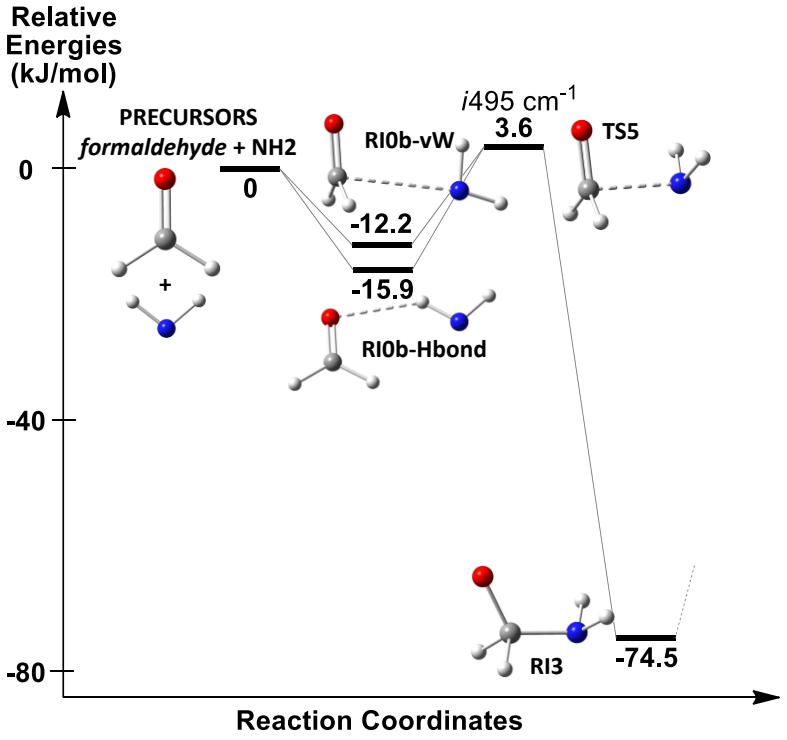



**Figure 2.1.9.** Proposed path for the approach of $NH_2^{\bullet}$ and formaldehyde. Relative electronic CCSD(T)/CBS+CV (in kJ/mol) energies in black.

**Proposed full $NH_2$+$H_2CO$ reaction path.**

Figure 2.1.10 displays the possible path of the $CH_2O$ + $^{\bullet}NH_2$ → $NH_2CHO$ + H formation reaction (skipping again details about the approach step) and the relative electronic (CCSD(T)/CSB+CV) and zero-point corrected energies (ZPE at the B2PLYP-D3/m-aug-cc-pVTZ level) of all minima and transition states. In the case of the intermediate **RI3**, the hindered $NH_2$ rotation needs to be carefully managed. To this end, all normal modes except that corresponding to the $NH_2$ torsion have been included in a VPT2 treatment, whereas a one-dimensional numerical approach has been used for the $NH_2$ hindered rotation (HR). The overall zero point energy is then obtained by summing $ZPE_{GVPT2}$ and $ZPE_{HR}$.

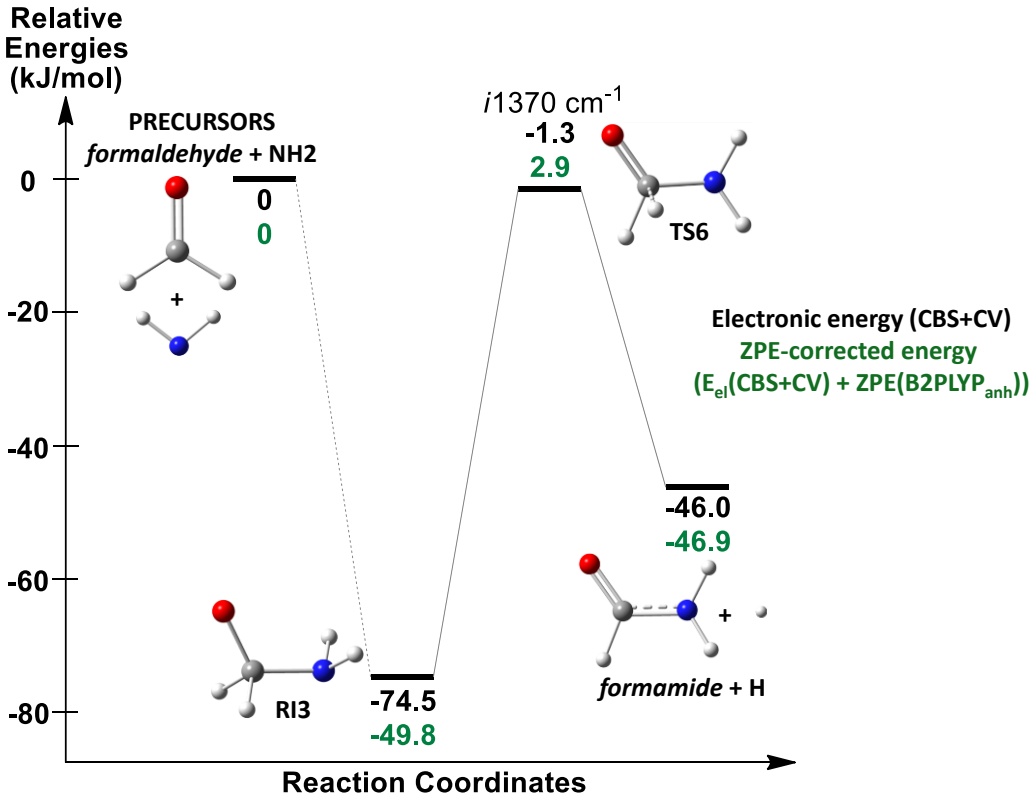



**Figure 2.1.10.** Proposed reaction path for formamide formation. Electronic energies (black) are at the CCSD(T)/CBS+CV level and the ZPE-corrected energies (green) are obtained by including the ZPE issuing from B2PLYP-D3/m-aug-cc-pVTZ anharmonic vibrational calculations. All energies are given in kJ/mol.

After the formation of the intermediate **RI3**, hydrogen loss can be observed, leading to formamide and the •H radical through the transition state **TS6** that has a barrier of 73.2 kJ/mol. The rotation around the C-N bond in **RI3** can lead to another energy minimum, which forms again formamide through the same transition state **TS6**. The products were found to be 46 kJ/mol more stable than the reactants. The possible radical mechanism based on the proposed reaction path is given in Figure 2.1.11.

If we take a look at the structural variations along this reaction path (see Table 2.1.3, values in bold), it is evident that the C-O and C-N bonds fluctuate. Indeed, the C-O bonds in formaldehyde and formamide are formally double bonds, while in the first intermediate **RI3**, where the $NH_2$ moiety is linked to the carbon atom, the CO bond becomes almost a single one. In the case of the transition state **TS6**, the presence of the H atom influences the C-O distance because it is somewhat longer than in formamide. As far as C-N is concerned, it is a single bond in the intermediate **RI3** and then shortens once the H atom is removed, to reach in formamide a length between the typical single and double bond. This observation can be explained by the delocalization that the amide moiety undergoes. It is also noted that this reaction occurs keeping a relatively high symmetry during the process. Indeed, after an addition of two $C_{2v}$ moieties, a compound of $C_S$ symmetry is obtained, which proceeds toward the product molecule, of $C_S$ symmetry as well, after the last transition state.

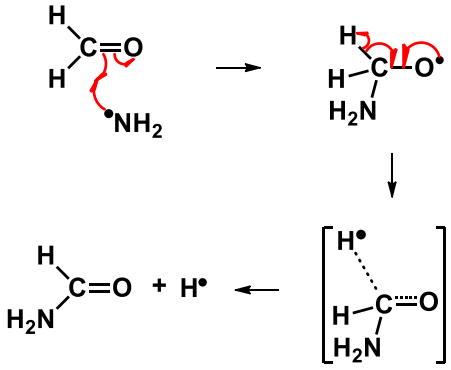



**Figure 2.1.11.** Possible radical mechanism for formamide formation through the $NH_2$ + $CH_2O$ reaction.

Let us now have a deeper insight into the hindered rotation occurring in the intermediate **RI3** mentioned above. Figure 2.1.12 shows the potential energy profile along the $\varphi$ dihedral angle, which describes the $NH_2$ rotation. In addition to the absolute energy minimum, the potential energy surface shows two additional equivalent local minima. All these minima correspond to staggered conformations and are connected by three transition states, corresponding to eclipsed conformations. The equivalent conformers ***a*** and ***c*** are separated by ***b***, which is the highest transition state, with a barrier of ca. 3.8 kJ/mol. The most stable conformer (***e***) is found to be about 10.5 kJ/mol more stable than the other minima. The ***d*** and ***f*** equivalent transition states govern the ***c*** to ***e*** and ***e*** to ***a*** transformations and have a relative energy about 14.6 kJ/mol higher than the most stable conformer ***e***.

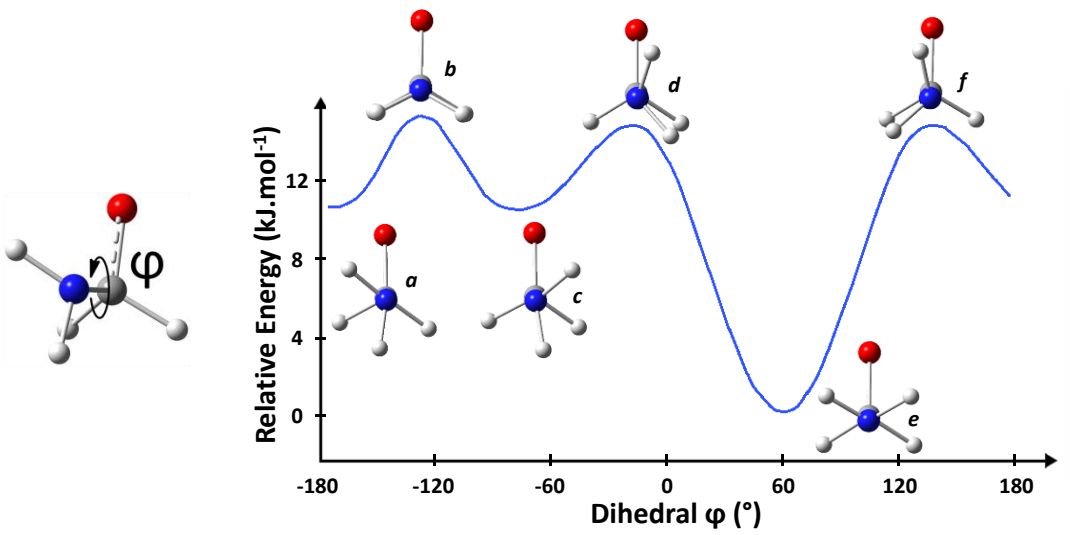



**Figure 2.1.12.** Potential energy profile along the $\varphi$ dihedral angle ($NH_2$ rotation) for **RI3** obtained at the B2PLYP-D3/maug-cc-pVTZ level.

## *2.1.5. Kinetics study*

**OH + $CH_2NH$**

The rate of back dissociation of the initial complex (assuming the validity of capture theory and using the principle of detailed balance) at a particular energy is proportional to the density of states (per unit volume) of the reactants. As the translational component of this density of states is proportional to the square root of the translational energy, at low

energies, back dissociation of the initial complex is suppressed due to the reactant density of states tending to zero. As a result, at low energies the initial van der Waals complex is present with the chance to either rearrange to $HOCH_2NH$ (with subsequent possible formation of methanimidic acid) or eliminate a water molecule forming the $H_2CN$ radical. The other two possible products (formamide + H or formaldehyde + $NH_2$) have negligible rates due to the fact that more than one step is required to reach them from the original complex and that one of these steps consists in overpassing the high energy **TS2**.

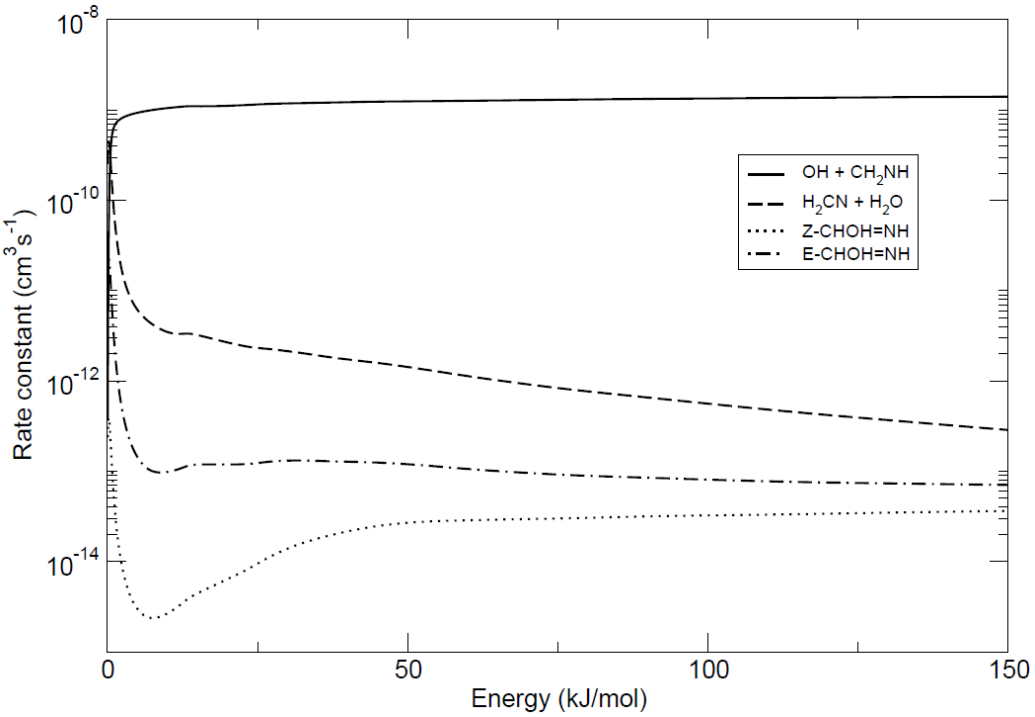



**Figure 2.1.13.** Rate constants for the formation of some products starting from OH + methanimine as a function of energy.

Both transition states corresponding to these channels lie ca. 4.5 kJ/mol below the reactant zero-point energy. However, the transition state leading to water elimination presents a higher sum of states than the second one at low energies, leading to a low-energy dominance of $H_2O$ elimination. This can be seen in Figures 2.1.13 and 2.1.14 as a constant ratio of around 8 between the rate constants of water elimination and methanimidic acid formation. It is to be noted that *Z*-methanimidic acid has a much lower rate of formation than

its *E*-isomer. This is essentially due to the fact that the corresponding transition state is rather lower in energy. Moreover, as the two isomers of $HOCH_2NH$ (which dissociate to the corresponding isomers of methanimidic acid) equilibrate rapidly between themselves and the *E*-isomer is more stable, this is another factor contributing to this effect. As the energy increases, the reactant density of states increases rapidly, and this brings about a rapid increase in the rate of redissociation of the complex. Indeed, the two next highest rate constants diminish with increasing energy because of this effect (on the other hand, the rate constant of *Z*-methanimidic acid continues to increase slightly).

The same effects are also seen in the canonical rate constants, albeit smeared out due to Boltzmann averaging. Despite the predominance of redissociation, $H_2CN$ and *E*-methanimidic acid formation remain important processes up to 300 K (with rate constants above $10^{-11}$ $cm^3.s^{-1}$).

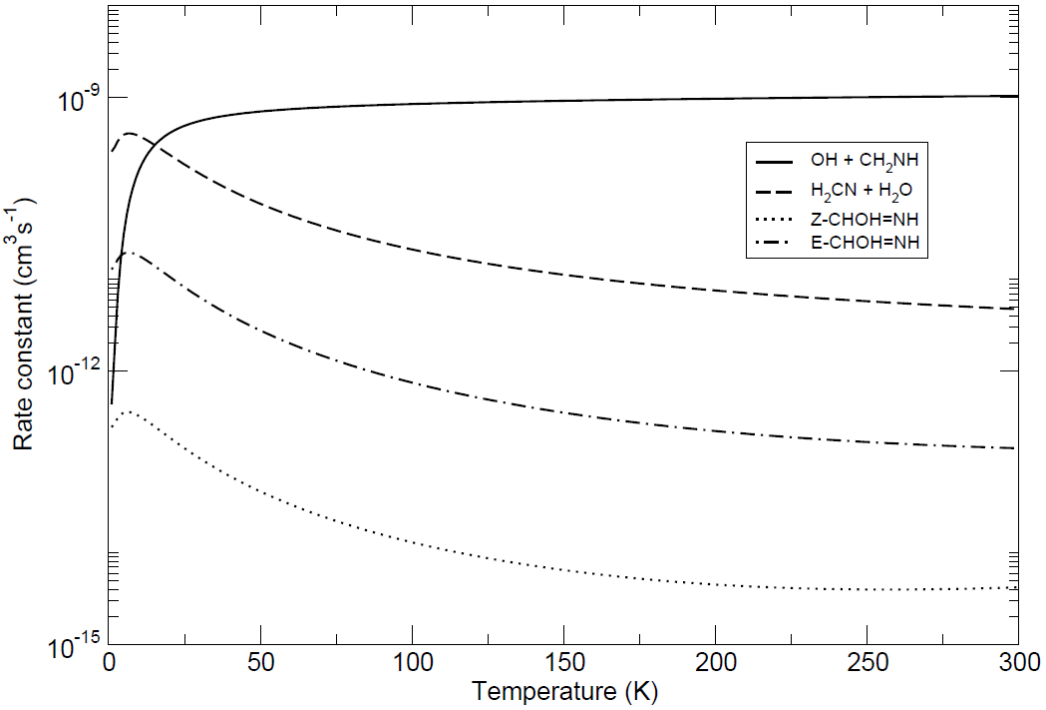



**Figure 2.1.14.** Rate constants for the formation of some products starting from OH + methanimine as a function of temperature.

**$NH_2$ + $H_2CO$**

A similar RRKM study was performed on the $NH_2$ + $H_2CO$ reaction, based on calculations at several levels of theory, including CBS-QB3.[131] This enabled us to ensure that, even at low temperatures, the formation of formamide through that reaction was largely predominant over the back-dissociation into $NH_2$ + $H_2CO$. This result was confirmed by new RRKM calculations on this system using the CCSD(T)/CBS+CV energies, as one can see in Figure 2.1.15. The rate constant is slightly lower than in the CBS-QB3 case (it drops from 3.1 × $10^{-10}$ to 7.6 × $10^{-11}$ $cm^3.s^{-1}$ at 5 K, with an error around 3-4%), which is explained by a higher barrier for the transformation of **RI3** into formamide + H (represented by **TS6**), but the formamide formation still predominates at low temperatures since the reactants density of states tends to zero.

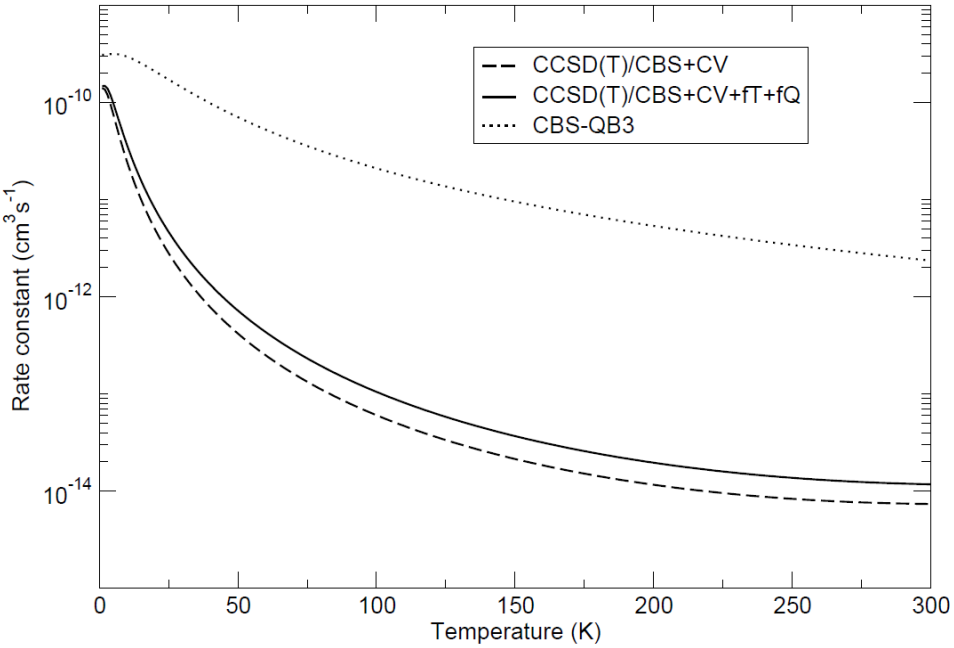



**Figure 2.1.15.** Rate constants for the formation of formamide starting from $NH_2$ + formaldehyde as a function of temperature.

Some information regarding the lifetime of the **RI3** complex is in order. At the lowest energies studied, the unimolecular rate constant for formamide formation was around 7 × $10^7$ $s^{-1}$, increasing with energy up to 3 × $10^{12}$ $s^{-1}$ (at around 150 kJ/mol). On the other hand,

back-dissociation starts from much smaller values (around $10^7$ $s^{-1}$) but rapidly gains on formamide formation, reaching a rate of around $7 \times 10^{16}$ $s^{-1}$. We have also considered the possibility of deactivation of the intermediate through spontaneous IR emission. In order to assess the order of magnitude, we have considered the vibrational mode with maximum transition dipole moment (corresponding to out-of-plane motion of H atoms) at the harmonic approximation. The A Einstein coefficient for the $v = 1 \rightarrow 0$ process was found to be $1.33 \times 10^4$ $s^{-1}$. This value is more than 3 orders of magnitude smaller than our lowest first order rate constants. Even considering excitation up to v = 10, this would still remain too low to be of any importance, and hence it has not been considered further.

Moreover, when including the full-T (CBS+CV+fT) and full-Q (CBS+CV+fT+fQ) corrections, the **TS6** relative energy that was found at −1.3 kJ/mol drops to −2.3 and −2.8 kJ/mol, respectively, which leads to slightly higher rate constants ($9.9 \times 10^{-11}$ $cm^3.s^{-1}$ at 5 K with CCSD(T)/CBS+CV+fT+fQ energies).

### *2.1.6. Partial conclusion*

New insights concerning the formation of formamide in the interstellar medium have been provided. Our computations allowed us to suggest two reaction paths combined to probable mechanisms concerning this formation. In both cases, $OH+CH_2NH$ and $NH_2+H_2CO$, the addition step does not involve significant barriers and can therefore occur in space. While RRKM calculations confirmed the effectiveness of the $NH_2+H_2CO$ reaction (since, once this first addition has been done, formamide formation largely predominates over back-dissociation at low energies), the $OH+CH_2NH$ reaction leads preferentially to other products of prebiotic interest such as methanimidic acid and the $H_2CN$ radical.

## 2.2. Quantum chemical computations of formamide deuteration

One particular aspect of the solar type star forming regions is the largely enhanced deuterium fractionation observed in H-bearing molecules.[172] Indeed, the deuteration of molecules formed in the first cold prestellar phase can be several orders of magnitude larger than the statistical value based on the elemental D/H ratio ($1.5 \times 10^{-5}$).[173] Emblematic examples are those of formaldehyde and methanol, where the doubly (and even triply, for methanol) deuterated forms have been detected with D-bearing/H-bearing ratios up to 30 (and 5) per cent.[174,175] Interestingly, the deuteration ratio in different molecules might provide the temporal sequence of their formation and, possibly, information on their synthetic route.[172,176]

Very recently, Coutens *et al*.[177] have reported the detection of deuterated formamide in the solar-type protostar IRAS16293.[178,179] In their work, they detected the three forms of deuterated formamide: *trans*- and *cis*- HCONHD, and $DCONH_2$. They found that they have approximately the same deuteration ratio, 2-5 %. These observations have been obtained with the ALMA interferometer and they refer to the hot corino region only. They noticed that the fact that the three forms have a similar deuteration ratio provides constraints on the synthesis of formamide. Different routes have been suggested in the literature: the gas-phase reaction $NH_2 + H_2CO$,[131] grain-surface radical recombination, synthesis dominated by UV or ion irradiation of ices containing various species, like methanol.[138,139] Coutens *et al*. noted that the relatively small percentage of deuterated formamide with respect to formaldehyde might suggest a preference for synthesis on ice, as gas-phase routes should mostly retain the deuteration ratio of parent molecules.

However, this is not necessarily true. First of all, even though all the isotopologues are characterized by the same electronic energies, the reaction rate coefficients for the isotopic variants can be influenced by factors such as zero point energies (ZPE) and densities of states. In addition, the reaction mechanism controls whether D or H displacement is favored in gas phase reactions involving partially deuterated species. In particular, the release of D over H can be favored in direct abstraction processes, while the release of H over D is favored in indirect reactions involving partially deuterated species.

Since the substitution of an H atom by a D one affects the density of states of all intermediates and transition states, as well as the tunnelling efficiency, a thorough investigation of the NHD + $H_2CO$ and $NH_2$ + HDCO reactions is presented here to derive the specific rate coefficients. They will allow to assess whether they are compatible or not with a formamide formation in the gas-phase through the above reactions.

In particular, the following isotopologues have been considered:

$NH_2 + HDCO \rightarrow DCONH_2 + H$ $\Delta H°_0 = -47.29$ kJ/mol (1)

$NH_2 + HDCO \rightarrow HCONH_2 + D$ $\Delta H°_0 = -39.07$ kJ/mol (2)

$NHD + H_2CO \rightarrow$ *trans*-HCONHD + H $\Delta H°_0 = -48.49$ kJ/mol (3)

$NHD + H_2CO \rightarrow$ *cis*-HCONHD + H $\Delta H°_0 = -48.50$ kJ/mol (4)

The exothermicities given above have been computed as described in the computational details. It can be seen that, for each of the two monodeuterated versions of the reactants, there are two possible products to consider. In the case of monodeuterated formaldehyde, both regular and C-deuterated formamide is formed. In the case of monodeuterated amide radical, the HCONHD can be formed in the *cis*- or *trans*-conformations.

### *2.2.1. Specific computational details*

The computations have been carried out using the B2PLYP-D3 functional,[24] in conjunction with the m-aug-cc-pVTZ basis set,[26,27] where *d* functions on hydrogens have been removed. Full geometry optimizations have been performed for all minima and transition states –deuterated or not-, also checking the nature of the obtained structures by diagonalizing their Hessians that have been used also to evaluate ZPEs at the harmonic level. Additional calculations based on coupled cluster methods have been performed in order to reevaluate the energies of all species. The coupled-cluster singles and doubles approximation augmented by a perturbative treatment of triple excitations (CCSD(T))[141] has been employed in conjunction with extrapolation to the complete basis set limit and inclusion of core-correlation effects (CCSD(T)/CBS+CV).

As previously, the results of the electronic structure calculations have been used to derive the capture theory rate coefficient for the formation of the initial bound intermediate

(the intermolecular potential remains invariant as in the case of formamide, but the different reduced masses slightly change the capture rate constants). Subsequently, using the RRKM scheme described before,[159,180] the rate constants (as a function of energy) have been calculated for the elimination of a hydrogen atom (or a D atom) leading to formation of formamide. Using a detailed balance argument along with the capture rate constants, the energy-dependent rate constant for back-dissociation of the intermediate to the reactants have also been calculated, as previously done. The master equation has been solved as a function of energy to determine the bimolecular rate constant for formamide formation and its isotopologues. Finally, a Boltzmann averaging has been carried out to determine rate constants as a function of temperature. The rate constants have subsequently been fitted to the appropriate rate law of the form

$$k(T) = A \times ({}^{T}/_{300K})^{\beta} \times e^{-\gamma/T} \qquad (1)$$

### *2.2.2. Electronic structure and kinetics calculations*

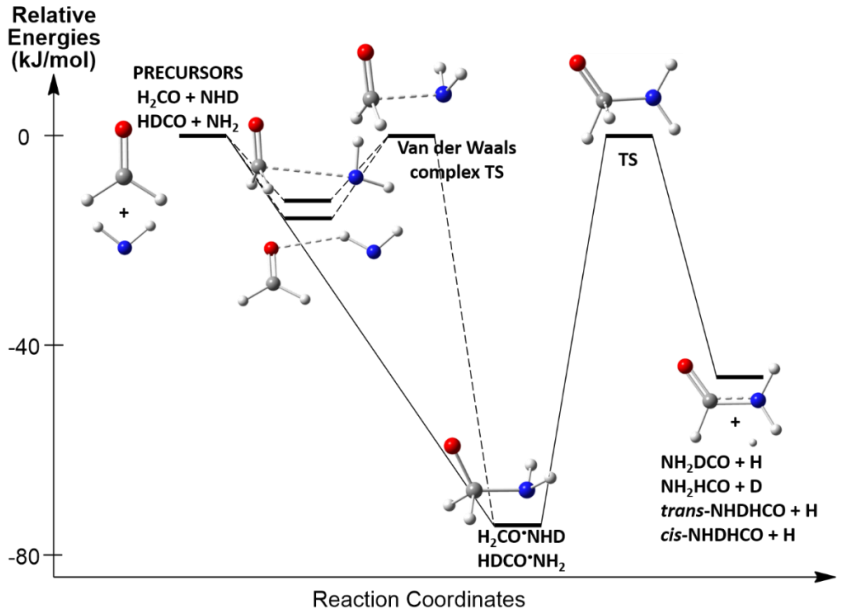



**Figure 2.2.1.** Complete path for the $H_2CO + NHD \rightarrow NHDHCO + H$ and $HDCO + NH_2 \rightarrow NH_2DCO + H$ processes.

As one can see on Figure 2.2.1, the interaction between the reactants starts with the formation of a shallow van der Waals complex and the subsequent formation of a bound

intermediate. This one can then eliminate an H atom to form formamide or back-dissociate to the reactants. The electronic potential energy surface is the same for all isotopomers, but the energy levels, vibrational frequencies, rotational constants and zero point energies depend on the atomic masses involved, as it can be seen in Table 2.2.1. This can change the relative densities of states and the resulting reaction rate constants.

Song & Kastner[181] questioned the gas-phase formation of formamide through the $NH_2$ + $H_2CO$ reaction, arguing that the transition state connecting the van der Waals complex to the intermediate is too high when including the ZPE (from 2.7 kJ/mol to 17.8 kJ/mol with ZPE at the UCCSD(T)-F12/cc-pVTZ-F12 level of calculations), and the reaction rate coefficient very small (of the order of $10^{-22}$ $cm^3s^{-1}$). Against this claim, it must be stressed that: *i)* Even though the calculated energy of the transition state leading from the van der Waals complex to the bound intermediate is higher than the reactant asymptote according to coupled cluster calculations, its energy substantially drops by including higher excitation orders in the cluster operator (when a full treatment of triple –fT- and quadruple excitations –fQ- is included, thus leading to the CCSD(T)/CBS+CV+fT and CCSD(T)/CBS+CV+fT+fQ approaches, the barrier issuing from this transition state significantly decreases -3.6 kJ/mol drops to 2.05 kJ/mol with full-T and to 1.67 kJ/mol with full-T and full-Q). Extrapolating to the full configuration interaction limit, the electronic energy of this transition state drops slightly below the reactant level[182] in line with the variational principle of approximating from above. *ii)* The use of the ZPE correction for the van der Waals complex and its transition state is not warranted. In an electronic calculation the ZPE is derived from local considerations on the potential minimum (harmonic or anharmonic). However, the three new vibrational modes in the van der Waals complex consist of a very loose stretching mode (where even a perturbative approach including anharmonicity would tend to grossly overestimate the frequency) and two loose bending modes that constitute almost free rotations. Under these conditions, in order to derive the exact ground state level one would have to exactly solve the Schrödinger equation on the overall potential. It is assumed that it is a much more realistic approach to neglect what emerges as "the zero point energy" as its full inclusion is unwarranted. Taking into account the two points above, it is much more reasonable to omit the van der Waals complex from the reaction scheme.

**Table 2.2.1.** Summary of the energy of each species involved in the deuterated versions of the $NH_2$ + $H_2CO$ reaction (CCSD(T)/CBS+CV electronic energies corrected with ZPE obtained with B2PLYP-D3/m-aug-cc-pVTZ). Each energy is given in hartree and in parentheses is given its relative energy with respect to the reactants in kJ/mol.

| Species | Electronic energies | ZPE corrected (X=H, Y=D) | | ZPE corrected (X=D, Y=H) | |
|---|---|---|---|---|---|
| *NHX + HYCO* | -170.367866 (0.00) | -170.3248386 (0.00) | | -170.3245346 (0.00) | |
| *HYCONHX* | -170.396176 (-74.33) | -170.3440276 (-50.38) | | -170.3442026 (-48.98) | |
| *TS* | -170.368405 (-1.41) | -170.323737 (2.89) $DCONH_2$ | -170.321363 (9.13) $HCONH_2$ | -170.323873 (1.74) cis-HCONHD | -170.323877 (1.73) trans-HCONHD |
| *Products* | -170.385371 (-45.96) | -170.342852 (-47.29) $DCONH_2$ | -170.339721 (-39.07) $HCONH_2$ | -170.343008 (-48.50) cis-HCONHD | -170.343005 (48.49) trans-HCONHD |

**Table 2.2.2.** Summary of the rate coefficients of the reactions involving deuterated formamide. The undeuterated case is also reported.[182] The last three columns report the values of the reaction rate at 10, 60 and 100 K (x $10^{-13}$ $cm^3s^{-1}$), respectively the temperature of a cold molecular cloud, of a shocked gas,[183] and of the IRAS16293-2422 B hot corino where formamide has been detected.[177]

| Reaction | A (x $10^{-16}cm^3s^{-1}$) | β | γ (K) | k (T=10K) | k(T=60K) | k(T=100K) |
|---|---|---|---|---|---|---|
| (1) $NH_2$ + HDCO → $DCONH_2$ + H | 20.7 | -2.75 | 4.34 | 153 | 1.60 | 0.41 |
| (2) $NH_2$ + HDCO → $HCONH_2$ + D | 1.08 | -2.15 | 0.96 | 1.49 | 0.03 | 0.01 |
| (3) NHD + $H_2CO$ → *trans*-HCONHD + H | 30.7 | -2.63 | 5.05 | 141 | 1.93 | 0.52 |
| (4) NHD + $H_2CO$ → *cis*-HCONHD + H | 26.2 | -2.64 | 5.04 | 126 | 1.69 | 0.45 |
| (5) $NH_2$ + $H_2CO$ → $HCONH_2$ + H | 77.9 | -2.56 | 4.88 | 287 | 4.41 | 1.23 |

Once formed, the bound addition intermediate has two possible fates: either *a)* react, eliminating an H/D atom and forming a formamide molecule or *b)* reverse the association step, returning to the original reactants. The relative rates of the two competing steps are important in determining overall rate constants. In this respect, it is very important to stress that back-dissociation of the addition intermediate is always in competition with the reactive event and its consideration in the calculations is essential to obtain meaningful rate coefficients because it reduces by much their real values. The practice of giving the capture rate coefficient as the global rate coefficient is therefore wrong and should be avoided.

Figure 2.2.2 illustrates this point, showing the canonical rate constants for all deuteration reactions compared to the capture rate constant. It can be seen that all reaction rate constants decrease at high enough energies while the overall capture rate constant increases. This is because of the rapidly increasing back-dissociation rate which renders simple capture rate constants completely unrealistic as estimates of reaction rate coefficients.

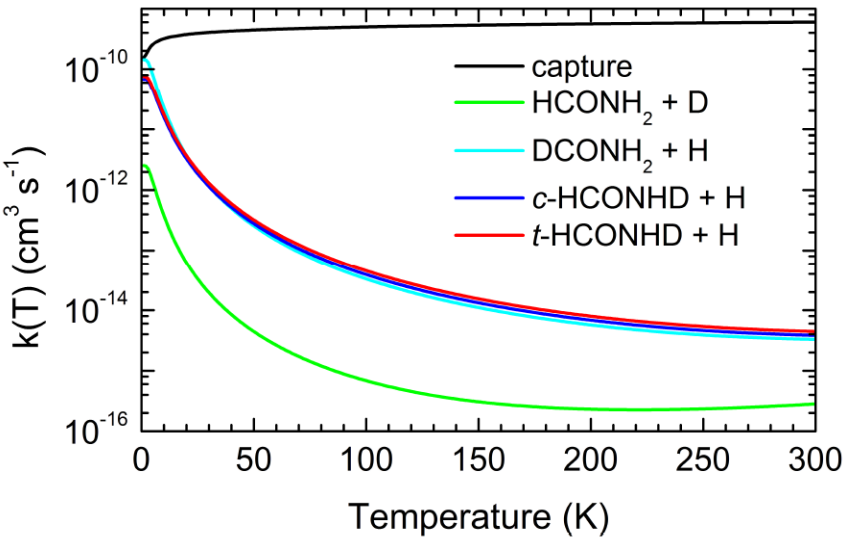



**Figure 2.2.2.** Canonical rate constants for the four different deuteration reactions (1)-(4) of Table 2.2.2 as a function of temperature. The capture rate constant is also shown for comparison.

Concerning the trend with the temperature, we can note that at low energies (temperatures), the reaction is promoted by tunnelling while the rate of back-dissociation is negligible (as a result of the relative translation density of states tending to zero, in accordance with the detailed balance principle). As a result, reaction is going to predominate. On the contrary, at higher temperatures back-dissociation predominates because the relative translation density of states increases more rapidly than the one of the transition state and tunnelling is not sufficient anymore.

The main focus of these new calculations concerns the rate coefficients and product branching ratios of the partially deuterated species, exhibited in Table 2.2.2. The most relevant results can so be summarized: *i)* the reactions involving the exit of an H atom have essentially equal rates, which are lower by a factor of 2-4 than the rate of the undeuterated reaction. For the NHD reaction, the constant capture incoming flux is split between two distinct product channels accounting for a factor of 2 (that is the observed difference at 10 K). The larger factor observed at higher temperatures and the reduced k(T) for the $NH_2$ + HDCO → $DCONH_2$ + H rate are, instead, direct consequences of the prevailing back-dissociation; *ii)* the reaction $NH_2$ + HDCO mainly produces $DCONH_2$, with only 1-2% going into $HCONH_2$. The reason for this is twofold: the lighter H atom tunnels more easily through the barrier towards the products by a factor of around 50 and the vibrational ZPE of the transition state is lower when an H rather than a D atom has started dissociating.

In conclusion, even though the partially deuterated formamide products retain the degree of deuteration of the parent molecules, the three isotopic variants of the $NH_2$ + $H_2CO$ reactions are not characterized by the same rate coefficient. In particular, since the fastest reaction is the one with undeuterated reactants, the expected value of the degree of formamide deuteration at 100 K is a factor of 3 lower than that of the parent molecules $H_2CO$ or $NH_2$.

### *2.2.3. Comparison with astronomical observations*

So far, deuterated formamide has only be detected towards the hot corino of IRAS16293-2422 B by Coutens et al.[177] They found that the $DCONH_2/HCONH_2$ and $HCONHD/HCONH_2$ abundance ratios are similar and around 0.02-0.05. Based on this and on

the fact that these ratios are smaller than the measured HDCO/$H_2CO$ abundance ratio, they favoured a grain-surface formation of $NH_2CHO$. With the new computations reported in Table 2.2.2, it is now possible to quantitatively discuss this issue. For that, let us consider the rates computed for a gas temperature of 100 K, namely the temperature in the IRAS16293-2422 B hot corino.

*(i) $DCONH_2/HCONH_2$ and $HCONHD/HCONH_2$ abundance ratios smaller than $HDCO/H_2CO$:* The values in Table 2.2.2 show that there is a factor of three difference in the rates of formation of H- against D- formamide. Therefore, the deuteration of formamide synthesised in the gas-phase will be three times smaller than that of the mother molecules $NH_2$ and $H_2CO$. No observations of $NHD/NH_2$ exist so far, so no constraints can be obtained from amidogen. Similarly, no observations exist of the HDCO/$H_2CO$ abundance ratio in the hot corino of IRAS16293-2422 B. However, single-dish observations of deuterated formaldehyde towards IRAS16293 measured HDCO/$H_2CO$ ~0.15,[174] namely a factor 3-8 larger than $DCONH_2/HCONH_2$. Therefore, within the uncertainty of the available observations, the hypothesis of formamide gas formation is fully consistent with them. In addition, as correctly pointed out by Coutens et al., single-dish observations encompass both the hot corino and the extended envelope. Ceccarelli et *al*. mapped the $D_2CO$ line emission towards IRAS16293 and showed that the $D_2CO/H_2CO$ abundance ratio is about 0.03. Likewise, HDCO/$H_2CO$ is also high in the extended envelope and, thus, the single-dish measurements are likely largely contaminated by the extended envelope.

*(ii) Similar $DCONH_2/HCONH_2$ and $HCONHD/HCONH_2$ abundance ratios:* The values in Table 2.2.2 show that indeed the two abundance ratios are similar (0.8-0.9) if the gaseous $NHD/NH_2$ and HDCO/$H_2CO$ are similar. In this respect, then, the gas-phase formation of formamide is perfectly consistent with the observed values.

### *2.2.4. Partial conclusion*

It is now possible to conclude that the available observations do not allow to rule out a gas-phase formation route for formamide but, on the contrary, support it. New high-spatial resolution observations of both NHD and HDCO towards IRAS16293 will be necessary to challenge this conclusion.

## 2.3. New observations as a support for gas-phase formation

In order to further test the gas-phase formation theory, high spatial resolution observations towards the nearby (250 pc) star forming region L1157 have been obtained. Figure 2.3.1 shows the overall structure of the L1157 system. It consists of a protostar, L1157-mm, having a luminosity three times that of the Sun. A precessing jet of matter at supersonic velocity emerges from this protostar. The jet has excavated two main cavities, called B1 and B2,[184] impacting them along the trajectory (projected on the sky) indicated on the figure. It needs to be noted that the further from L1157-mm is the cavity, the older is the impact. The B1 structure is located at about 16000 A. U. south of the protostar. Its bow morphology is the result of different impacts of the jet on the B1 cavity. The apex is the oldest shocked region within B1, with a kinematical age of ~1100 years,[185] while the B1a shock (Figure 2.3.2) is a few hundreds of years younger.[185]

Previous observations revealed that in B1a the jet impact caused the partial erosion of the grain cores and ices, injecting into the gas phase large quantities of SiO ($\sim 10^{-7}$ with respect to H nuclei), $H_2O$ ($\sim 10^{-4}$) and $CH_3CN$ ($\sim 10^{-8}$) among several other species.[184–188] Hence, L1157-B1 provides us with a perfect place where to study the reactions occurring when previously frozen species are injected into the gas, as their relative abundance evolution depends on the relative efficiency of the various reactions.

Besides, given the density and temperature of the B1 shocked gas (~105 $cm^{-3}$ and 60 K),[187] the timescale for this evolution is a few thousands of years, which is also the dynamical age of L1157-B1.[185] This allows us to follow the chemical evolution with time by measuring the abundances throughout the shocked gas since the northern regions, closer to the impacting jet, are also the younger ones. Therefore, with the reported high spatial resolution observations, we can put constraints on the reactions at work.

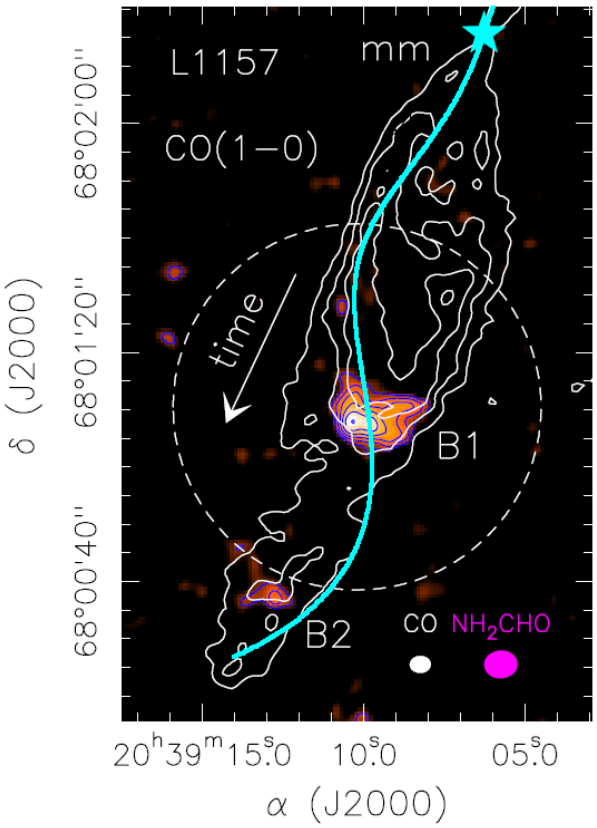



**Figure 2.3.1.** The L1157 south outflow lobe: the overall structure of the L1157 south lobe is illustrated by the CO (1-0) line emission (white contours), previously observed with the IRAM interferometer.[189,190] The jet ejected by the central object L1157-mm (blue star) excavated two cavities, of which the apexes are labelled B1 and B2 respectively. B1 is at 0.08 pc (16000 A. U.) from L1157-mm and its emission is blue-shifted by +2.6 km.s$^{-1}$ with respect to L1157-mm. The emission map of the $NH_2CHO$ ($4_{1,4}$-$3_{1,3}$) line (integrated over the velocity range -5 to +5 km.s$^{-1}$) is shown in the color image (and blue contours). For the CO image, first contour and step are 6σ (1σ = 0.5 Jy.beam$^{-1}$.km.s$^{-1}$) and 4σ, respectively. First contours and steps of the $NH_2CHO$ map correspond to 3σ (1σ = 5 mJy.beam$^{-1}$.km.s$^{-1}$) and 1σ, respectively. The dashed circle shows the primary beam of the $NH_2CHO$ ($4_{1,4}$-$3_{1,3}$) line observations (64”). The magenta and white ellipses depict the synthesized beams of the $NH_2CHO$ (5.”79 x 4.”81, PA = -94°) and CO (3.”65 x 2.”96, PA = 88°) observations, respectively. The cyan solid curve represents the loci of the impact of the precessing jet (projected on the plane of the sky) driven by L1157-mm against the B1 and B2 cavities.[185] The white arrow indicates the direction of increasing time, namely increasing kinematical age of the outflow.

### *2.3.1. Details of the observations*

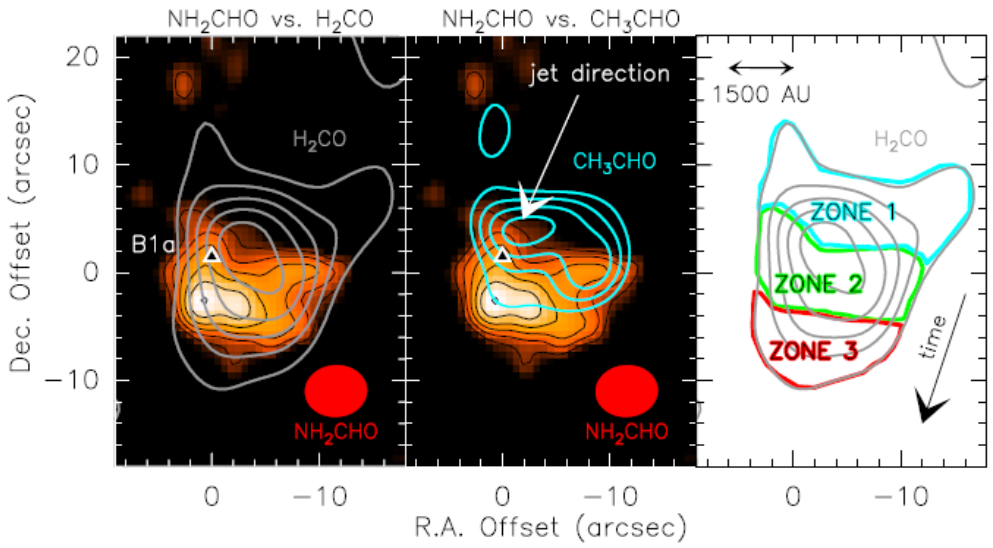



**Figure 2.3.2.** Chemical segregation in the L1157-B1 protostellar shock: the maps are centred at α(J2000) = $20^h$ $39^m$ $10.^s2$, δ(J2000) = +68° 01' 10."5, at *i.e.* Δα = +22" and Δδ = -65" from the driving protostar L1157-mm (see Figure 2.3.1). All the images are smoothed to the same angular resolution (red ellipse, see Figure 2.3.1) of the $NH_2CHO$ ($4_{1,4}$-$3_{1,3}$) line data. *Left panel*: p-H2CO ($2_{0,2}$-$1_{0,1}$) integrated emission (grey contours)[191] on top of the $NH_2CHO$ line emission map (colour image, black contours). For the $H_2CO$ image, first contour and step are 3σ (1σ = 3.3 mJy.beam$^{-1}$.km.s$^{-1}$) and 6σ, respectively. The formaldehyde line emission traces the overall structure of the B1 shock, while that of formamide covers only the southern part of B1. The triangle identifies the B1 youngest position, where the dynamical age has been estimated to be 1100 yr,[185] and where the precessing jet driven by L1157-mm impacts the B1 cavity.[184] *Middle panel*: Same as in the left panel for the $CH_3CHO$ ($7_{0,7}$-$6_{0,6}$) E + A velocity-integrated emission (cyano contours; first contour and steps correspond to 3σ, 4 mJy.beam$^{-1}$.km.s$^{-1}$, and 1σ, respectively).[191] The acetaldehyde line emission covers the northern part of B1. *Right panel*: Sketch of the three zones that compose B1: zone 1 (blue), $CH_3CHO$ (and not $NH_2CHO$) is detected, zone 2 (green), both $CH_3CHO$ and $NH_2CHO$ are detected and zone 3 (red), $NH_2CHO$ (and not $CH_3CHO$) is detected. It has to be noted that time increases and chemistry evolves going from north (Zone 1) to south (Zone 3). The black arrow indicates the direction of increasing time, namely increasing kinematical age of the outflow.

The observations of the formamide line ($4_{1,4}$-$3_{1,3}$), at 81.7 GHz, have been obtained in the framework of the large program SOLIS (Seeds Of Life In Space) at IRAM NOEMA (Northern Extended Millimeter Array). Figure 2.3.1 shows the large-scale line intensity map. The brightest line intensity coincides with the B1 apex, where the line is detected with a S/N ≥ 8. Weaker (S/N ~ 4) emission is also detected towards the older B2 apex. However, since it is outside the telescope primary-beam it has been excluded from further analysis. The region emitting formamide in B1 has a beam deconvolved size of ~9" (2300 A. U.). The intensity peak is situated at 5" (1250 A. U.) south of B1a, namely in an older region. The line averaged over the whole emitting region has a linewidth with a Full Width Half Maximum (FWHM) of 4.6 ±0.6 km.s$^{-1}$ and is blueshifted by -2.6 km.s$^{-1}$ with respect to the L1157-mm velocity,[187] in agreement with previous observations.[187] Using the emitting size and assuming an excitation temperature of 10 K (as derived by several formamide lines observed with the IRAM 30-m antenna[183]), the average formamide column density results in 8 ±4 x $10^{12}$ cm$^{-2}$, corresponding to an estimated average abundance (with respect to H nuclei) of around 4 x $10^{-9}$ (assuming a H column density of 2 x $10^{21}$ cm$^{-2}$)[187].

In Figure 2.3.2, a zoom-in of the B1 structure is reported. The formamide emission traces only the southern section of the B1 structure, while formaldehyde is more extended and it traces the whole shocked region. This figure also shows the spatial distribution of formaldehyde emission,[191] which depicts the overall bow-structure of L1157-B1. A first important result is that formamide emission does not completely coincide with that from formaldehyde, which traces the whole B1 structure, but only covers the southern portion of it. Figure 2.3.2 reports as well the emission from another COM, acetaldehyde ($CH_3CHO$).[192] Surprisingly, acetaldehyde behaves in an almost opposite way with respect to formamide, that is to say, it is mostly associated with the northern portion of B1. When the difference between these two species is considered, one can identify three zones. Zone 1: the northern and youngest one, where only acetaldehyde emits; Zone 2: an intermediate zone, where both formamide and acetaldehyde are present; Zone 3: the southern and oldest region, where only formamide emits. It can be noticed that the difference in the three zones cannot be attributed to excitation effects, as the mapped formamide and acetaldehyde lines have similar upper level energies (from 11 K to 26 K), similar Einstein coefficients ($A_{ij}$ ~ $10^{-5}$ s$^{-1}$), and the derived

excitation temperatures are also similar (10 K against 15 K, for formamide and acetaldehyde, respectively).[183,192] Therefore the difference between the three zones must be due to a difference in the chemical composition, thus showing a clear evolutionary effect.

### *2.3.2. Astrochemical model: assumptions and results*

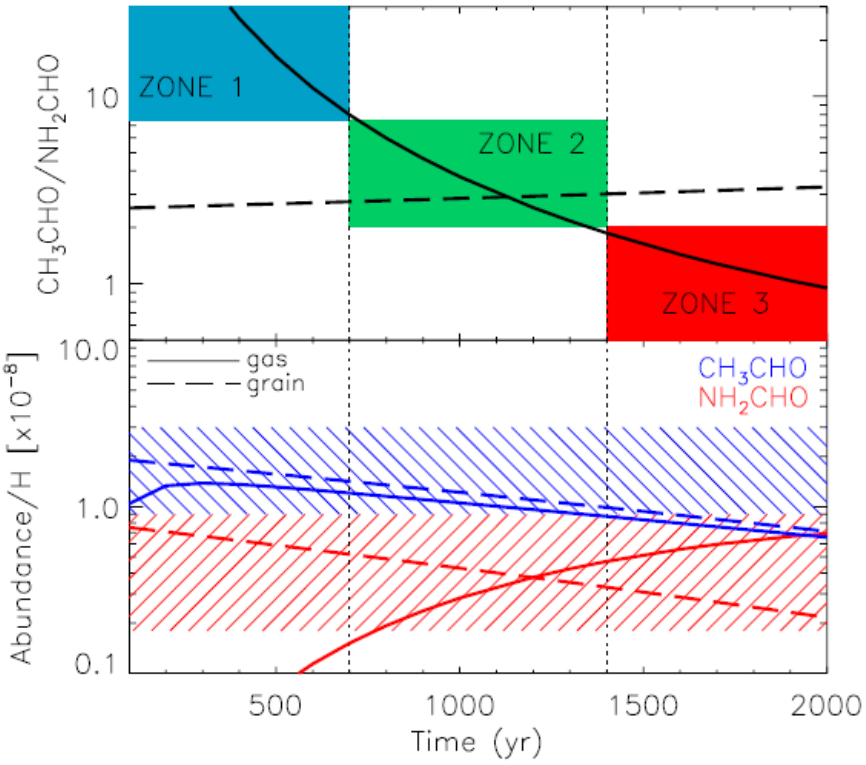



**Figure 2.3.3.** Model predictions of acetaldehyde and formamide after a shock occurrence: *Lower panel*: Acetaldehyde ($CH_3CHO$, blue) and formamide ($NH_2CHO$, red) abundances, with respect to H nuclei, as a function of time from the passage of the shock. The dashed blue and red regions show the maximum and minimum $CH_3CHO$ and $NH_2CHO$ measured abundances. Solid lines refer to a model where acetaldehyde and formamide are both synthesized in the gas phase whereas dashed lines refer to predictions assuming that they are injected into the gas phase directly from the grain mantles. The two dotted vertical lines define the time when the formamide and acetaldehyde abundances (as derived by the gas phase model) fall below the minimum measured values. Upper panel: $CH_3CHO/NH_2CHO$ calculated abundance ratio as a function of time for pure-gas (solid) and pure-grain (dashed) chemistry. Zones 1, 2 and 3 are defined in Figure 2.3.2.

To understand what the observed chemical difference implies, an astrochemical model has been run considering three possibilities: *(i)* formamide and acetaldehyde are grain-surface chemistry products, *(ii)* formamide and acetaldehyde are gas-phase chemistry products, *(iii)* one of the two species is a grain-surface and the other one a gas-phase chemistry product. Briefly, the model consists in a time-dependent gas-phase code that follows the chemical evolution of the gas. It starts with the chemical composition of a molecular cloud and then simulates the passage of the shock by suddenly increasing the gas density and temperature (to $10^5$ $cm^{-3}$ and 60 K, respectively),[187] and the gaseous abundance of grain mantle molecules. The abundances of the mantle molecules are assumed to be similar to those measured by IR observations of the dust ices[193] or specifically constrained by previous studies on L1157-B1.[192,194] A chemical network of 511 species connected by almost 8000 reactions was adopted.

*(i) Acetaldehyde and formamide are grain-surface chemistry products:*

First, both formamide and acetaldehyde are assumed to be synthesized on the grain surfaces and that the passage of the shock injects these two species into the gas phase in quantities such that the measured abundances are roughly reproduced. The predicted abundances as a function of the time are shown in Figure 3.2.3. They decrease by approximately the same factor in a 2000 yr interval. Actually, the predicted [$CH_3CHO$]/[$NH_2CHO$] abundance ratio slightly increases with time, which is in contrast with the observations that show exactly the opposite trend. Therefore, the pure grain-surface hypothesis cannot explain the observed formamide/acetaldehyde segregation. In other words, the observations rule out the hypothesis that acetaldehyde and formamide are directly injected from the grain mantles into the gas phase.

*(ii) Acetaldehyde and formamide are gas-phase chemistry products:*

Then, it is assumed that both acetaldehyde and formamide are formed in the gas phase from species previously on the grain mantles and injected into the gas-phase at the shock passage. Acetaldehyde is assumed to be formed by the reaction of ethyl radical ($CH_3CH_2$) with atomic oxygen, following several previous studies:[192] $CH_3CH_2 + O \rightarrow CH_3CHO + H$. Formamide is assumed to be formed by the reaction of amidogen with formaldehyde:[131,182] $NH_2 + H_2CO \rightarrow NH_2CHO + H$. Various models have been run with different values of ethyl radical, ammonia

(mother of $NH_2$), and formaldehyde, to reproduce the observed abundances. The best agreement with observations is obtained by injecting into the gas phase 4 x $10^{-8}$ of ethyl radical, 2 x $10^{-5}$ of ammonia and 1 x $10^{-6}$ of formaldehyde. The predicted acetaldehyde and formamide abundances as a function of time are shown in Figure 2.3.3. This model not only reproduces fairly well the observed abundances, it also fits the behavior of the [$CH_3CHO$]/[$NH_2CHO$] abundance ratio, with acetaldehyde more abundant in the younger northern Zone 1 and formamide more abundant in the older southern Zone 3. It needs to be noted that the evolution timescale is sensitive to the cosmic ray ionization rate ζ. Thus, it was varied to verify how it affects the results. The best agreement with the observations is obtained when ζ = 6 x $10^{-16}$ $s^{-1}$, which is very close to that previously found,[194] based on the analysis of the ions in L1157-B1.

*(iii) Either acetaldehyde or formamide is a grain-surface and the other a gas-phase chemistry product:*

The possibility that acetaldehyde is synthesized on the grain surfaces and formamide in the gas was investigated and similar results to the case (ii) have been obtained. Hence, the gaseous $CH_3CHO$ abundance evolution is rather independent on the formation route (surface chemistry or gas-phase chemistry). Finally, the possibility that the gas-phase reaction $NH_2$ + $H_2CO$ is not efficient was checked.[181] In this case, no model can reproduce the observations (both the abundance and the evolution).

In summary, the new SOLIS observations indicate that the formation of formamide in L1157-B1 is dominated by gas-phase reactions involving species previously hydrogenated on the grain surfaces. The formamide abundance needs to peak when the acetaldehyde abundance has already started to decrease. This is only possible if formamide is formed in the gas phase and the reaction between amidogen and formaldehyde[131,182] successfully reproduces the observations. In addition, it is extremely encouraging that independent dynamical studies of L1157-B1[185] show that Zone 1 is, within the uncertainty affecting the estimates derived from observations, a few hundred years younger than Zone 3, as predicted by the astrochemical model. Although unable to place constraints on the acetaldehyde formation route, it can be noted that quantum chemistry computations have shown that the simple combination of the

methyl radical ($CH_3$) and formyl radical (HCO) is an inefficient channel on water ice surfaces,[195] so that it is likely that acetaldehyde is also a gas-phase product.

### *2.3.3. Partial conclusion*

In conclusion, these new observations demonstrate that formamide is a gas-phase product and that the reaction $NH_2 + H_2CO \rightarrow NH_2CHO + H$ explains well the observations. The detection of COMs in cold objects in the last few years[196] has already challenged a pure grain-surface chemistry paradigm for their formation.[18,197] These new observations add evidence that gas-phase chemistry plays an important role in the game of the COMs formation.

# *Chapter 3. Astrochemical study of the formation of glycolaldehyde and acetic acid starting from ethanol*

Glycolaldehyde $HOCH_2CHO$, as the simplest monosaccharide, and its isomer acetic acid $CH_3COOH$ are of large importance in the field of prebiotic chemistry, which aims at figuring out how life appeared on an originally inorganic Earth. Indeed, while the first is itself very important in biochemistry as a sugar, the second is able to act as a precursor for the synthesis of glycine, one of the most important amino acid.[198]

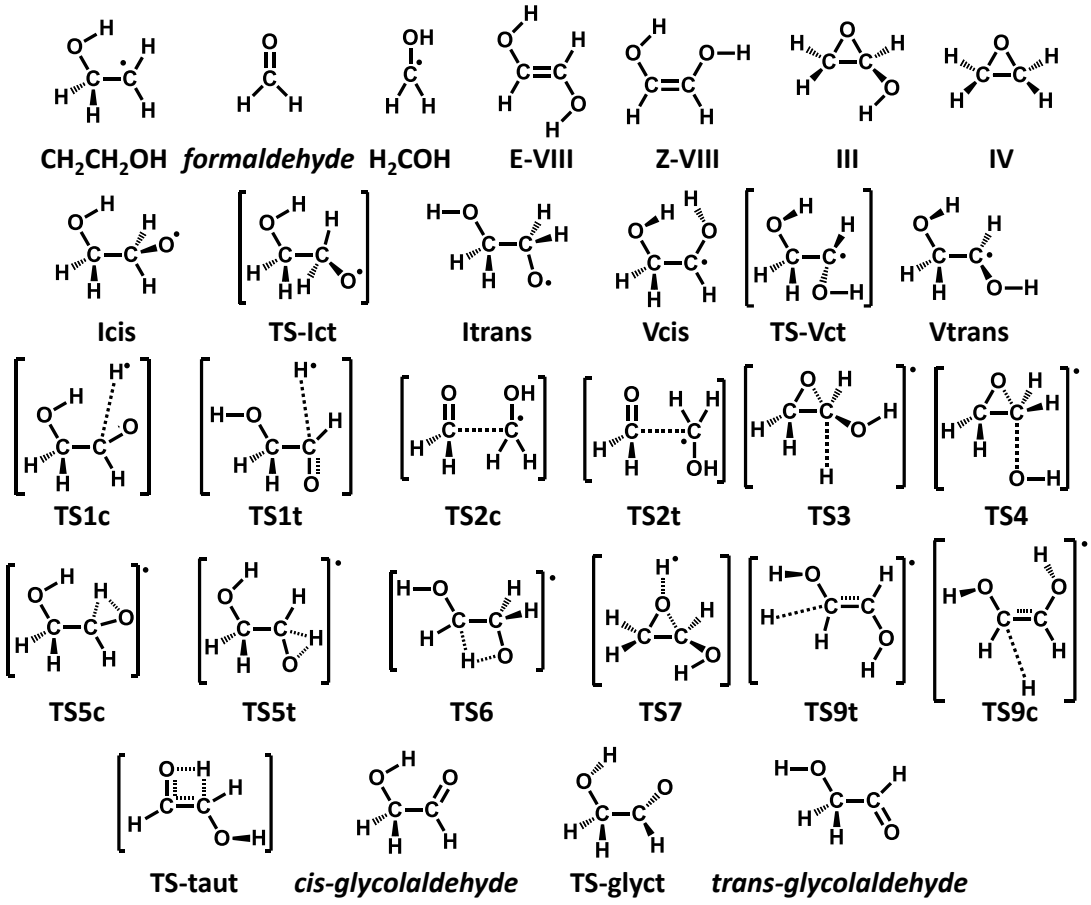



**Figure 3.1.** Investigated compounds related to glycolaldehyde formation.

Both isomers have been detected in the ISM by means of micro-wave spectroscopy. The first detection of acetic acid in 1997 towards the galactic center source Sgr B2 was presented by Mehringer *et al.*.[199] After that, it has also been observed towards several hot

molecular cores.[200,201] Glycolaldehyde, as far as it is concerned, has been also originally detected towards Sgr B2[202] and then towards hot cores outside the Galactic Center.[203,204] More recently, it has been observed towards a low-mass binary protostellar system IRAS 16293–2422.[13] So far, glycolaldehyde has been detected in a lower abundance than acetic acid.

The gas-phase formation routes for glycolaldehyde and acetic acid proposed up to now have not been able to successfully reproduce their estimated abundances. Therefore, they have been suspected to be formed through grain surface chemistry.[202,205–208] In order to verify if the formation of glycolaldehyde and acetic acid could not be also explained thanks to gas-phase chemistry, new radical formation paths in the gas-phase involving the atomic oxygen and hydroxyethyl radicals were computationally studied. All the investigated compounds present in the schemes are shown in Figures 3.1 and 3.2.

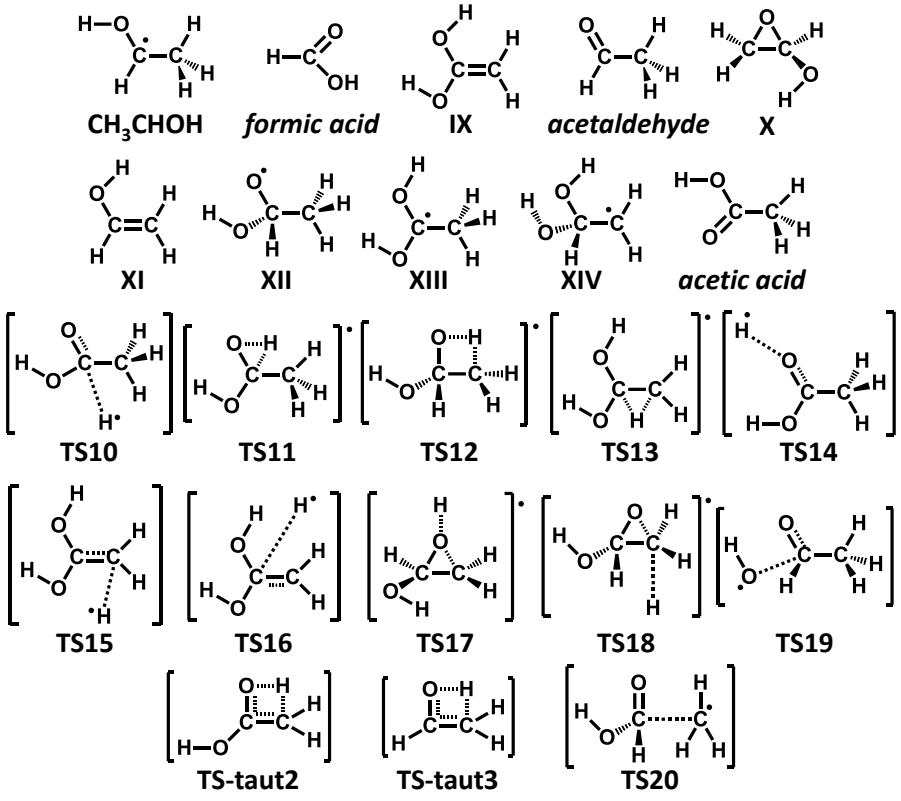



**Figure 3.2.** Investigated compounds related to acetic acid formation.

## *3.1. Specific computational details*

All the computations performed in this study follow the same scheme as presented previously for the formamide formation, in Section 2.1.1.

## *3.2. Vibrational study of glycolaldehyde and acetic acid*

In order to further check the reliability of our computational model (*i.e.*, B2PLYP-D3/m-aug-cc-pVTZ) we compared simulated and experimental infrared (IR) spectra of *cis*-glycolaldehyde and acetic acid (see Figures 3.3 and 3.4).[209,210]

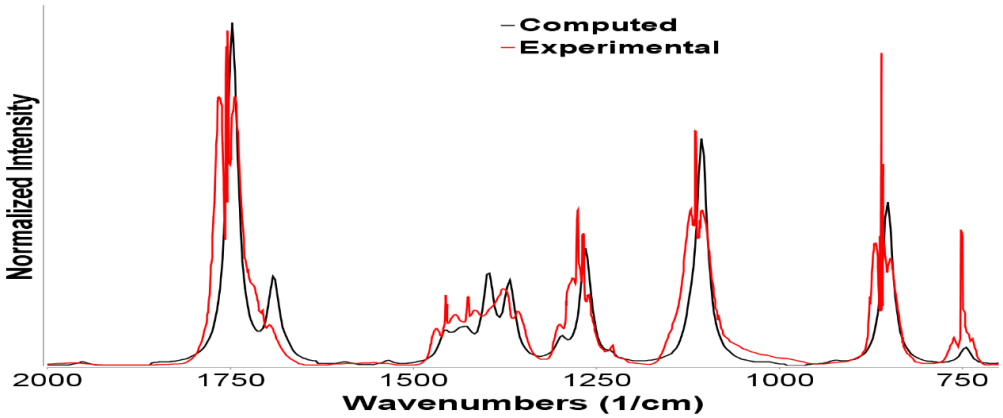



**Figure 3.3.** Computed (black, B2PLYP-D3/m-aug-cc-pVTZ anharmonic treatment) vs. experimental (red, taken from Ref. [209]) infrared spectrum of *cis*-glycolaldehyde.

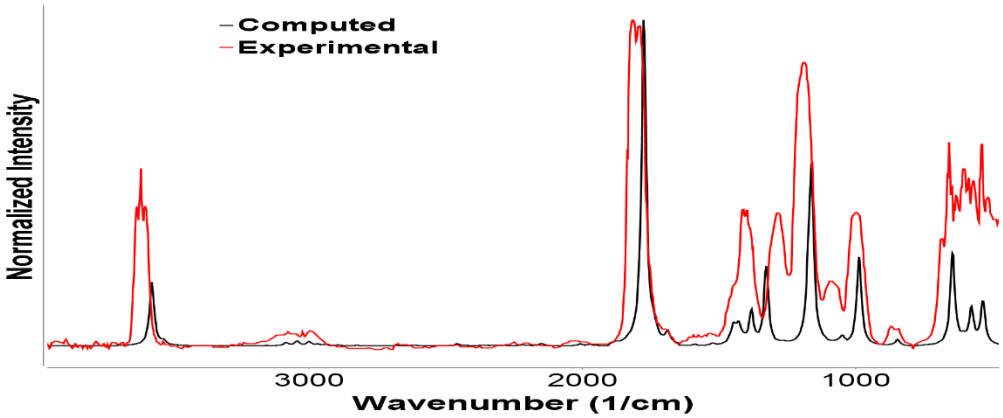



**Figure 3.4.** Computed (black, B2PLYP-D3/m-aug-cc-pVTZ anharmonic treatment) vs. experimental (red, taken from Ref. [210]) infrared spectrum of acetic acid.

It is noticeable that the simulated spectra fit nicely the experimental ones. Such a good agreement gives us further confidence to investigate the reaction path concerning the formation of glycolaldehyde and acetic acid with this method.

### *3.3. Hydroxyethyl radicals*

Both reactions involve an addition of oxygen O($^3$P) on a hydroxyethyl radical, easily obtained from ethanol and radical •OH. In the case of the formation of glycolaldehyde, the oxygen atom is added on the 2-hydroxyethyl radical •$CH_2CH_2OH$, while in the case of the acetic acid formation, the oxygen atom is added on the 1-hydroxyethyl radical $CH_3$•CHOH. It is therefore interesting to figure out how accessible the transformation from one into the other is.

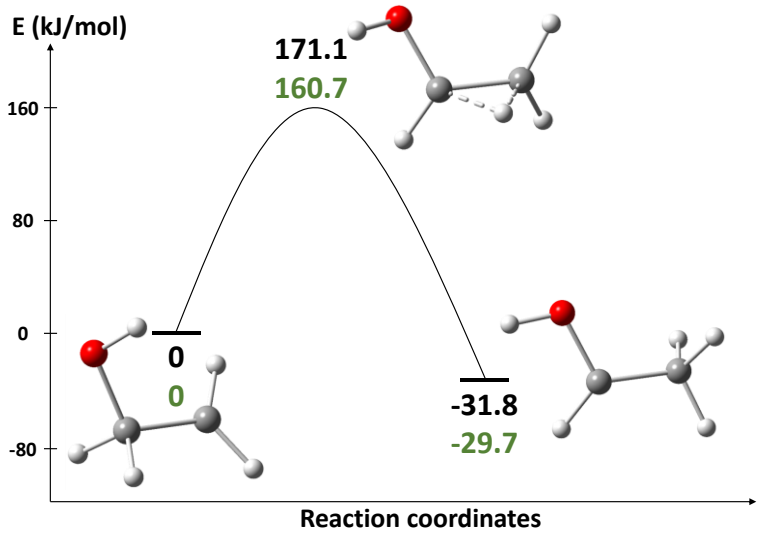



**Figure 3.5.** Relative energies (electronic in black and zero-point corrected in green, in kJ/mol) diagram of the •$CH_2CH_2OH$ ↔ $CH_3$•CHOH equilibrium.

On Figure 3.5, one can see the relative energies of $CH_3$•CHOH and of the transition state linking both hydroxyethyl radicals, with respect to •$CH_2CH_2OH$. It is noteworthy first that the $CH_3$•CHOH radical is more stable than its •$CH_2CH_2OH$ isomer, by around 30 kJ/mol. From another side, the required energy to go from one to the other is very high in the ISM (*ca.* 170

kJ/mol) and makes this reaction not likely to occur in this medium. Furthermore, any transition state was found to link both paths, which made us suppose them to be isolated.

## *3.4. Electronic calculations*

### *3.4.1. Glycolaldehyde*

Figure 3.6 exhibits the full reaction path following the addition of atomic oxygen O($^3$P) on the 2-hydroxyethyl radical $^{\bullet}CH_2CH_2OH$. While both fragments are approaching, the H atom belonging to the OH moiety on the $^{\bullet}CH_2CH_2OH$ radical stabilizes the oxygen atom. Indeed, the barrier-less addition of oxygen leads to the **Icis** compound, which is *ca.* 400 kJ/mol more stable than the reactants. Its *trans* counterpart, the slightly less stable (by 8 kJ/mol) intermediate **Itrans**, can easily be reached from the *cis*, through a 20 kJ/mol barrier (**TS-Ict**). Both **Ix** species are then able to undergo a dissociation into formaldehyde and the $^{\bullet}CH_2OH$ radical, thanks to transition states **TS2c** and **TS2t**. These dissociations exhibit around 55 kJ/mol barriers. Other dissociations can also be observed from both **Ix**, leading this time to *cis*- or *trans*-glycolaldehyde, and $^{\bullet}H$, respectively through **TS1c** and **TS1t** that are about 115 kJ/mol higher in energy than their **Ix** corresponding species. *Cis*- and *trans*-glycolaldehyde are connected by **TS-glyct** and require a 30 kJ/mol energy to go from one to the other. The *cis* conformer is the marginally most stable one, by 15 kJ/mol. Two epoxidations can also be envisaged. The first one, starting from the compound **Icis**, leads to the epoxide **III** and $H^{\bullet}$ and exhibits a barrier (represented by **TS3**) of *ca.* 240 kJ/mol. The second one, as far as it is concerned, starts from **Itrans**, leads to the epoxide **IV** and $^{\bullet}OH$ and has to go through a *ca.* 110 kJ/mol barrier, represented by **TS4**. Again starting from both **Ix** species, one can see hydrogen migrations from carbon atoms to the oxygen atom bearing the lone electron. Both **TS5x** correspond to hydrogen migrations from the carbon atom linked to the oxygen atom bearing the lone electron and are 128 kJ/mol more energetic than their corresponding **Ix** intermediates. The **TS6** transition state can be reached from both **Ix** and consists in a hydrogen migration from the carbon that is not linked to the oxygen atom bearing the lone electron. It is 122 and 115 kJ/mol energetically higher than **Icis** and **Itrans**, respectively. **TS5c** and **TS5t** repectively lead to **Vcis** and **Vtrans** while **TS6** is able to lead to both.

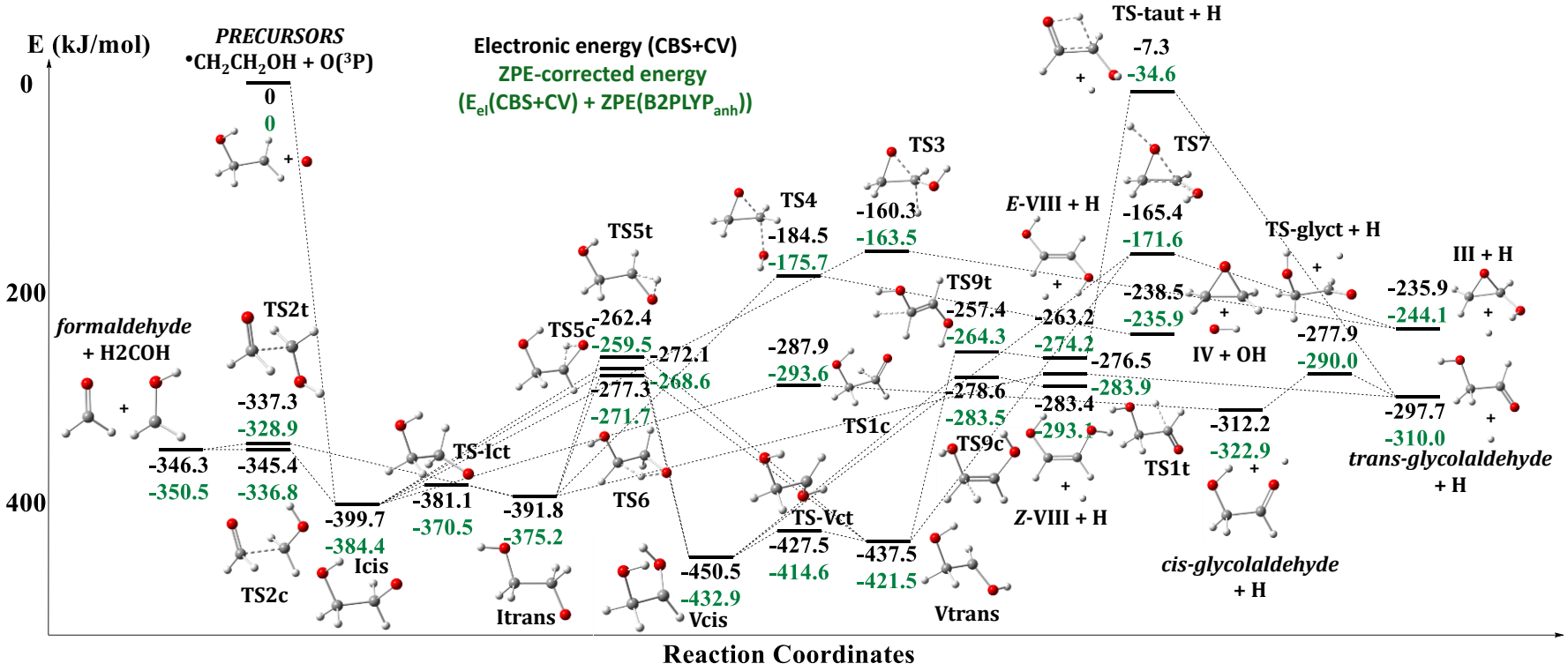



**Figure 3.6.** Proposed path for the addition of O(³P) on •CH2CH2OH reaction. Relative electronic energies (black) are issuing from CCSD(T)/CBS(T,Q)+CV(CT) calculations while the applied ZPE corrections (green) come from anharmonic vibrational calculations at B2PLYP-D3/m-aug-cc-pVTZ level. All the energies are given in kJ/mol.

These **Vx** intermediates are found to be more than 430 kJ/mol more stable than the precursors and can be linked to each other thanks to the **TS-Vct** transition state with a small barrier of around 25 kJ/mol. Starting from both **Vx** compounds, an epoxidation leading to the epoxide **III** and •H can be envisaged and exhibits a *ca.* 280 kJ/mol barrier (**TS7**).

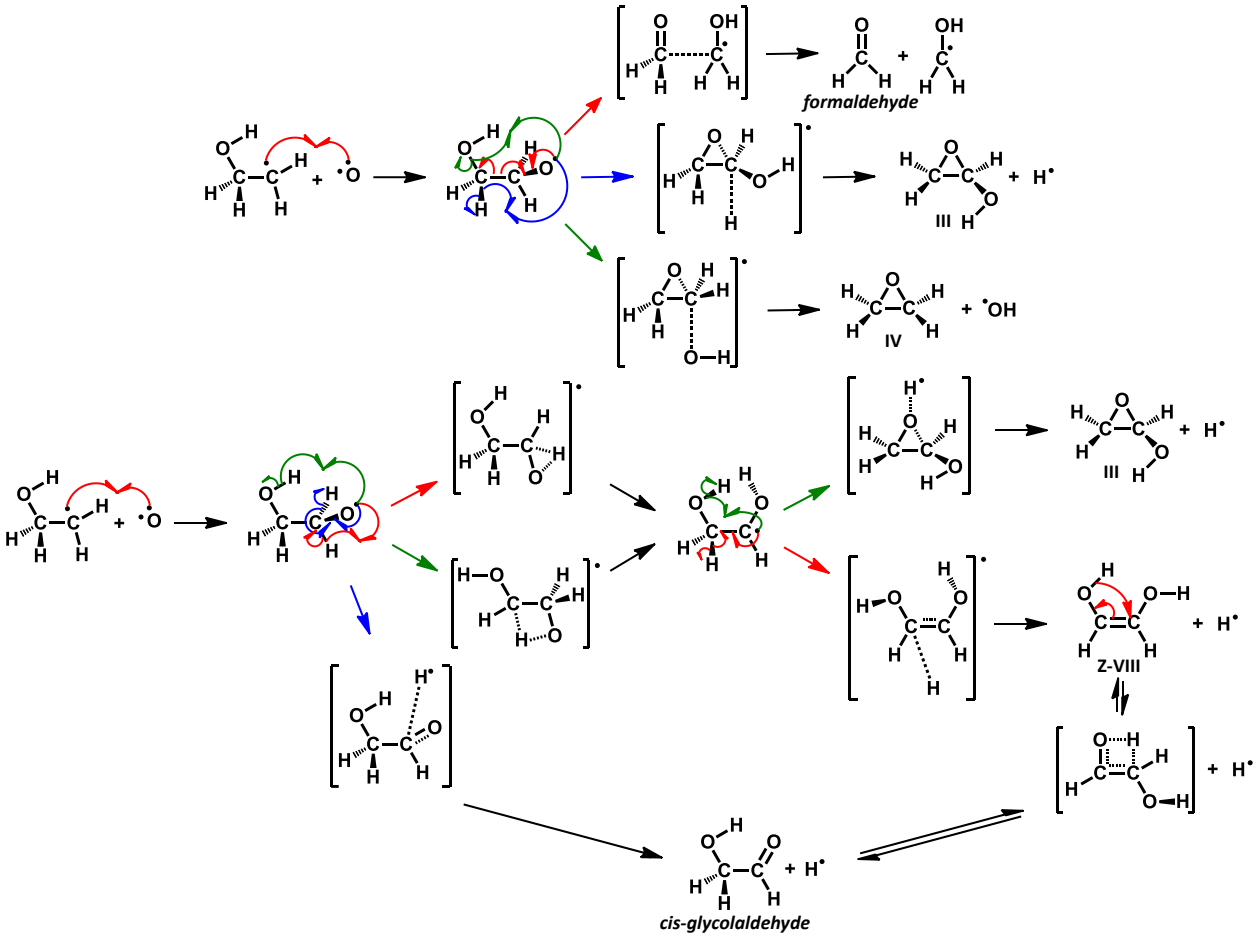



**Figure 3.7.** Possible radical mechanism for glycolaldehyde formation, focused on *cis* species.

The other possible reaction is a dissociation from both **Vx** into *Z*- and *E*-ethene-1,2-diol isomers (**Z-VIII** and **E-VIII**) and •H. The step from **Vcis** to **Z-VIII** + •H and the one from **Vtrans** to **E-VIII** + •H exhibit barriers of about 175 kJ/mol (through **TS9c** and **TS9t** transition states, respectively), both isomers being found around 270 kJ/mol more stable than the precursors and the *Z* one slightly more stable (by 20 kJ/mol) than its *E* counterpart. This **E-VIII** species can then undergo a tautomerization leading to *trans*-glycolaldehyde that exhibits a barrier of

255 kJ/mol. The same tautomerization from **Z-VIII** to *cis*-glycolaldehyde was also considered but would require a non-available energy in interstellar medium (the transition state linking them being *ca.* 40 kJ/mol higher in energy than the precursors).

If we look more carefully at the products that can be obtained with this path, one can see that the most stable ones are formaldehyde + $H_2COH$, with a relative energy of around -346 kJ/mol. The following ones, in decreasing order of stability, are glycolaldehyde + H, **Z-VIII** + H, **IV** + OH and **III** + H with relative energies of about -310, -280, -338 and -335 kJ/mol, respectively.

The proposed radical mechanism based on this path is given in Figure 3.7.

### *3.4.2. Acetic acid*

Figure 3.8 exhibits the full reaction path following the addition of atomic oxygen $O(^3P)$ on the 1-hydroxyethyl radical $CH_3{}^{\bullet}CHOH$. The barrier-less addition of oxygen leads to the **XII** compound, which is *ca.* 410 kJ/mol more stable than the reactants. This **XII** species is then able to undergo a dissociation into formic acid and the $^{\bullet}CH_3$ radical, through the transition state **TS20**. This dissociation exhibits a ca. 45 kJ/mol barrier. Other dissociations can also be observed, leading this time to acetaldehyde and $^{\bullet}OH$ or to acetic acid and $^{\bullet}H$, respectively through **TS19** and **TS10** that are about 100 and 60 kJ/mol higher in energy than **XII**. Thanks to a keto-enol tautomerism, acetaldehyde can turn into its **XI** enol form, through a 285 kJ/mol barrier (**TS-taut3**). Starting again from **XII**, a ca. 250 kJ/mol barrier (**TS18**) epoxidation is possible, leading to epoxide **X** and $^{\bullet}H$. One can envisage also two types of hydrogen migration. The first one would be a migration from the carbon atom in the carboxylic acid moiety and the second one from the carbon atom in the $CH_3$ moiety, both to the oxygen atom bearing the lone electron. These two migrations respectively lead to compounds **XIII** and **XIV** and exhibit 105 (**TS11**) and 135 (**TS12**) kJ/mol barriers. The intermediates **XIII** and **XIV** are found to be respectively *ca.* 460 and 410 kJ/mol more stable than the precursors and are connected to each other thanks to **TS13** which is 215 kJ/mol less stable than **XIII**. This compound **XIII** can also undergo two types of hydrogen loss. The first one concerns the hydrogen linked to an oxygen atom and results in acetic acid and $^{\bullet}H$ through a 117 kJ/mol barrier (**TS14**) while the second one concerns a hydrogen atom on the $CH_3$ moiety. This second possibility leads to the enol **IX** and $^{\bullet}H$ and requires an energy of 178 kJ/mol to overcome the barrier (**TS15**).

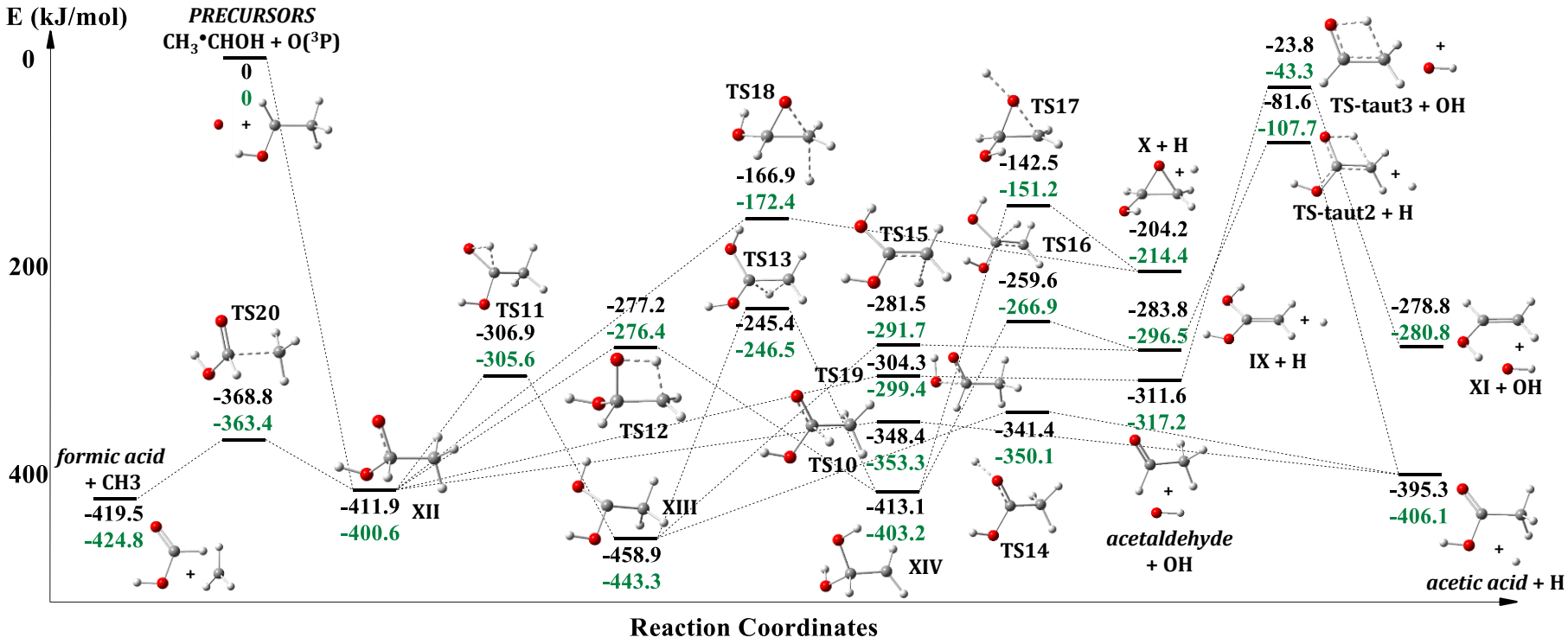



**Figure 3.8.** Proposed path for the addition of O(³P) on CH₃•CHOH reaction. Relative electronic energies (black) are issuing from CCSD(T)/CBS(T,Q)+CV(CT) calculations while the applied ZPE corrections (green) come from anharmonic vibrational calculations at B2PLYP-D3/m-aug-cc-pVTZ level. All the energies are given in kJ/mol.

Starting from intermediate **XIV**, it is feasible to obtain again the epoxide **X** and •H through a 271 kJ/mol barrier (**TS17**) or the enol **IV** and •H through a 154 kJ/mol barrier (**TS16**). Through a *ca.* 200 kJ/mol barrier (**TS-taut2**), this enol **IX** can give acetic acid thanks to a keto-enol tautomerism.

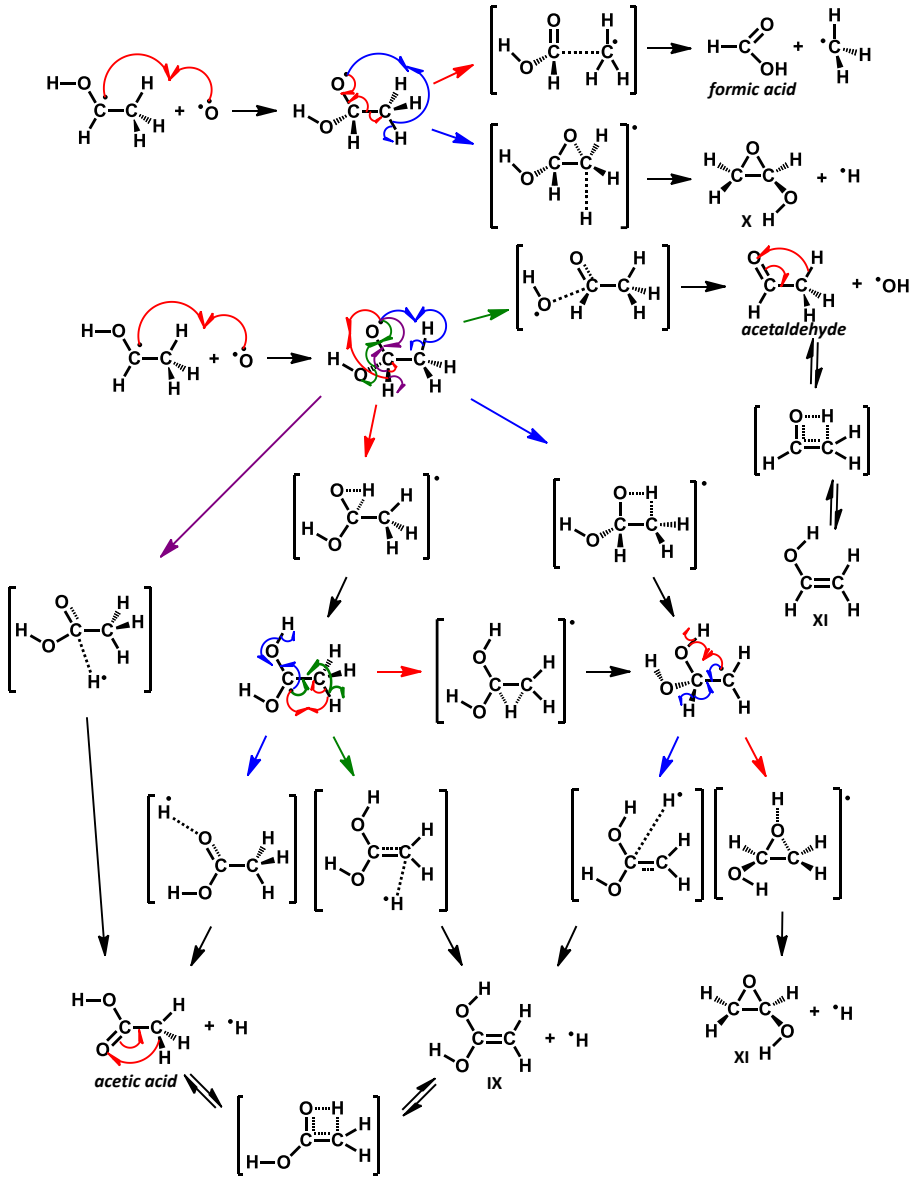



**Figure 3.9.** Possible radical mechanism for acetic acid formation.

If we take a closer look to the possible reaction products, it is noticeable that the most stable ones are formic acid and •$CH_3$, with a relative energy of -419.5 kJ/mol with respect to the

precursors. Following in decreasing order of stability are acetic acid + •H, acetaldehyde + •OH, and their enol forms (**IX** + •H, **XI** + •OH) that exhibit relative energies of -395.3, -311.6, -283.8 and 278.8 kJ/mol, respectively.

The proposed radical mechanism based on this path is given in Figure 3.9.

## *3.5. Kinetics study*

Figures 3.10 and 3.11 exhibit the rate constants for the formation of some final products issuing from the reactions leading to glycolaldehyde and acetic acid, respectively. In both cases, the formation of some products was found negligible and therefore not shown on the figures.

First, when focusing on the reaction leading to glycolaldehyde, one can see that the rate constant is maximum for the formaldehyde. It can be explained by the low energy of the transition states leading to it and the fact that only one step is required to reach it. Then, the formations of *cis*- and *trans*-glycolaldehyde exhibit the second and third highest rate constant.

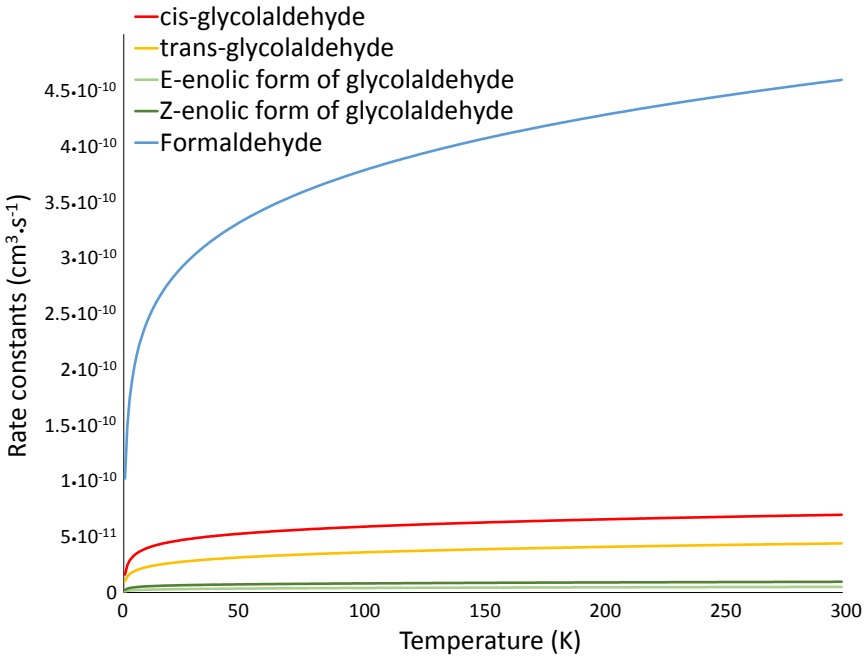



**Figure 3.10.** Rate constants for the formation of some final products from the $O(^3P)$ + $^{\bullet}CH_2CH_2OH$ reaction.

As far as the reaction leading to acetic acid is concerned, the rate constant is again not maximum for acetic acid, but for formic acid. This is for the same reasons as before. However, the rate constant for the formation of acetic acid comes second and is far from being negligible. One can also notice that this reaction leads also to acetaldehyde.

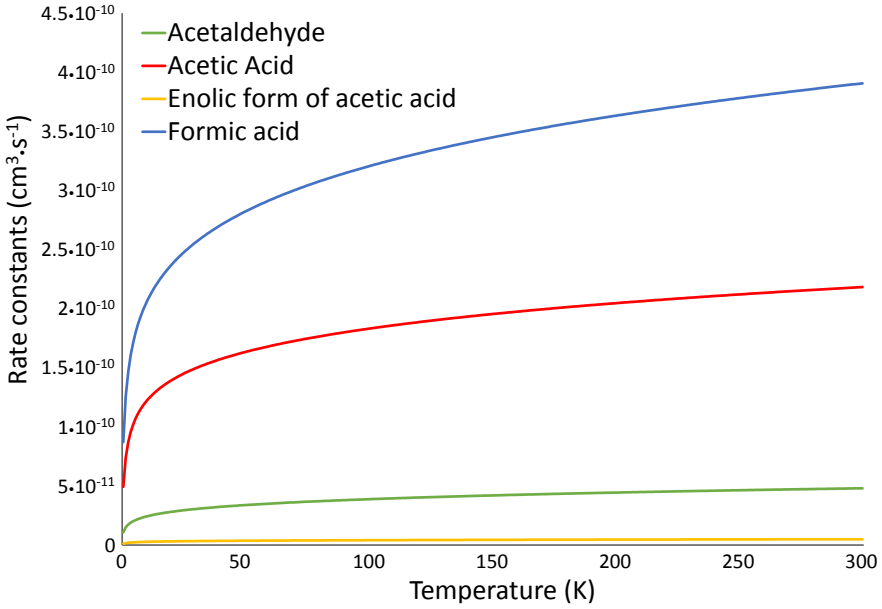



**Figure 3.11.** Rate constants for the formation of some final products from the $O(^3P)$ + $CH_3$•CHOH reaction.

In both reactions, even if glycolaldehyde and acetic acid are not the major products, their formation is still efficient. Moreover, all the by-products can be considered of interest in prebiotic chemistry and have been detected in ISM.

## *3.6. Astrochemical modelling*

In order to understand whether the proposed new reaction schemes and rates can explain the observations towards the hot corinos, an astrochemical model that simulates the conditions of the hot cores/corinos was used.

Figure 3.12 shows the abundance of glycolaldehyde, acetic acid, formic acid and ethanol as a function of time. In the conditions assumed by the model, the injected ethanol is all consumed in about 2000 yr. Formic acid is the one that benefits most, followed by acetic

acid and, finally, glycolaldehyde. Before ethanol is fully burned, the abundance ratios are $HCOOH/CH_3COOH \sim 1.5$ and $CH_3COOH/HCOCH_2OH \sim 10$, and are mostly governed by the branching ratios of the first two steps of the proposed reaction paths (Figures 3.6 and 3.8).

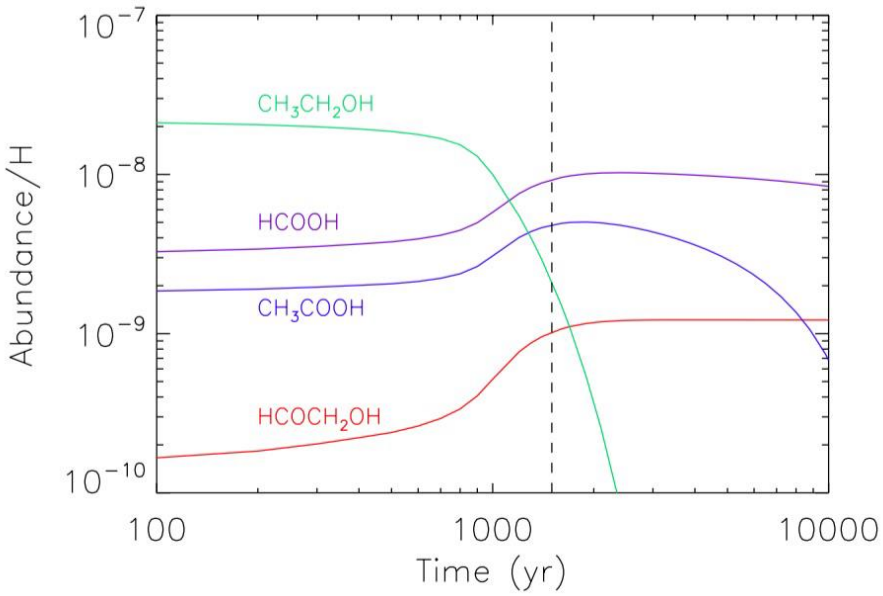



**Figure 3.12.** Abundance of glycolaldehyde (red), acetic acid (purple), formic acid (blue) and ethanol (green) as a function of time, at 100 K. The black dashed line shows the time of 1.5 x $10^3$ yr, time of interest for the observations.

The comparison of the model predictions with measurements of the glycolaldehyde abundance for generic hot corino conditions gives very encouraging agreement. In addition, the proposed schemes to synthesise glycolaldehyde and acetic acid from ethanol naturally explains the correlation seen by Lefloch et al.[211] between the abundance of these two species. On the contrary, the abundance of acetic acid and formic acid is predicted to be about one order of magnitude larger than the ones measured. However, the measurement of the acetic acid abundance and an upper limit to that of formic acid is L1157-B1, which is not a hot corino, so a more specific modelling is necessary before firmly conclude that there is a problem.

## *3.7. Partial conclusion*

New insights into the formation of glycolaldehyde and acetic acid in the ISM were provided. Electronic structure and kinetics calculations demonstrated that the reactions $O(^3P)$ + $^{\bullet}CH_2CH_2OH$ and $O(^3P)$ + $CH_3{}^{\bullet}CHOH$ are viable formations route of glycolaldehyde and acetic acid, respectively, even under the extreme conditions of density and temperature typical of ISM. The difficult interconversion between both hydroxyethyl radicals and the absence of transition states leading to one path to the other assures that both of them are separate. The obtained rate constants were inserted in an updated chemical network to produce predicted abundances of glycolaldehyde, acetic acid and formic acid as a function of that of ethanol. The predictions match extremely well with the measured abundance of glycolaldehyde in solar type hot corinos and shock sites. On the contrary, acetic acid and formic acid are predicted to be about ten times more abundant than the extremely sparse detections so far available towards hot corinos and shock sites (only one in each case). This might point to a lack of important routes of destruction of these two molecules in our network, possibly via reactions involving radicals and leading to larger molecules. However, since observations are published towards only two sources and the model presented here contain a very generic description of hot corinos conditions, more observations and source-dedicated modelling are necessary to confirm this discrepancy.

# PART III. Miscellaneous Investigations

## *Computational details*

All calculations have been carried out with a development version of the Gaussian suite of programs.[21] The computations were performed with the B3LYP[22,23] hybrid and B2PLYP[24] double hybrid density functionals, in conjunction with the m-aug-cc-pVTZ basis set, where *d* functions on hydrogens have been removed.[26,27] Semiempirical dispersion contributions were also included into DFT computations by means of the D3 model of Grimme, leading to B3LYP-D3 and B2PLYP-D3 models.[28,102] Full geometry optimizations have been performed for all compounds checking the nature of the obtained structures by diagonalizing their Hessians. In the case of complexes, the basis set superposition error (BSSE) has been corrected using the Counterpoise method.[212,213]

## *Chapter 1. Rotamers of glutamine*

Considering that the glutamine molecule (Figure 1.1) can undergo five different free rotations, up to 70 stable rotamers can be found, exhibiting energies in a wide range.

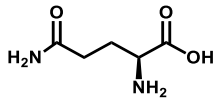



**Figure 1.1.** Glutamine formula.

Measurements in gas phase allowed to detect two rotamers whose rotational constants are shown in Table 1.1.

**Table 1.1.** Measured rotational constants of the detected rotamers of glutamine (MHz).

| | **Rotamer A** | **Rotamer B** |
|---|---|---|
| ***A*** | 2467.70941(14) | 2124.4(28) |
| ***B*** | 727.94433(22) | 847.8152(72) |
| ***C*** | 655.81028(16) | 805.666(21) |

In order to figure out the structures of these two specific rotamers, quantum calculations have been carried out.

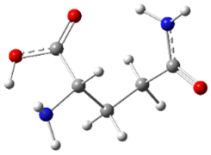
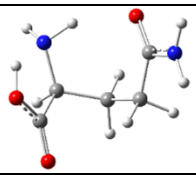
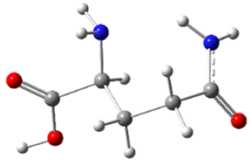
## 1.1. Specific computational details

The calculations were performed with the B3LYP-D3[22,23] and B2PLYP-D3[28] functionals, in conjunction with the m-aug-cc-pVTZ basis set, where *d* functions on hydrogens have been removed.[26,27]

## 1.2. Electronic structures

After the optimization (using the B2PLYP-D3/m-aug-cc-pVTZ level of theory) of a large amount of rotamers, the three most stable ones have been selected (**1**, **2** and **3**) and their relative energies together with their rotational constants have been gathered in Table 1.2. Indeed, one can see in this table that the calculated rotational constants for the third most stable rotamer (**3**) can be matched with the detected rotamer A.

**Table 1.2.** Calculated (B2PLYP-D3/m-aug-cc-pVTZ) rotational constants, relative electronic and zero-point corrected energies (kcal/mol) and structures of the three most stable conformers of glutamine.

| ***Compounds*** | **1** | **2** | **3 = Rotamer A** |
|---|---|---|---|
| ***Relative $E_{el}$*** | 0.00 | 0.52 | 1.81 |
| ***Relative $E_{el}$ + ZPE*** | 0.02 | 0.00 | 1.38 |
| ***A (MHz)*** | 2206.3859 | 1955.4738 | 2498.6829 |
| ***B (MHz)*** | 884.2410 | 899.7865 | 723.9824 |
| ***C (MHz)*** | 733.9179 | 843.3812 | 654.4891 |
| ***Structures*** |  |  |  |

This match led to an issue: why aren't there three detected rotamers? Indeed, both most stable ones should also be detected. To try to answer that question, the hypothesis of an easy interconversion between two of these three rotamers was formed.

## 1.3. Interconversion paths

In order to reduce the computational cost, all the calculations concerning the paths were performed using the B3LYP-D3/m-aug-cc-pVTZ level of theory.

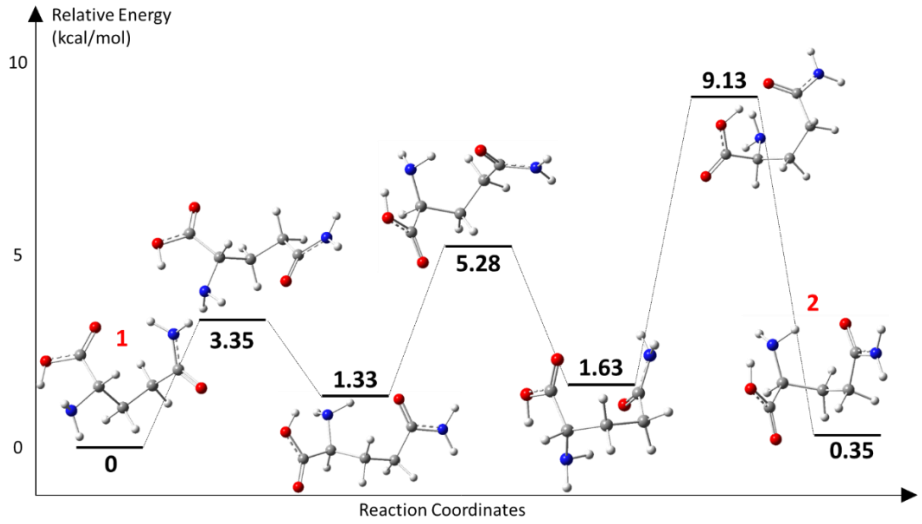



**Figure 1.2.** Interconversion path between rotamers 1 and 2. (B3PLYP-D3/m-aug-cc-pVTZ).

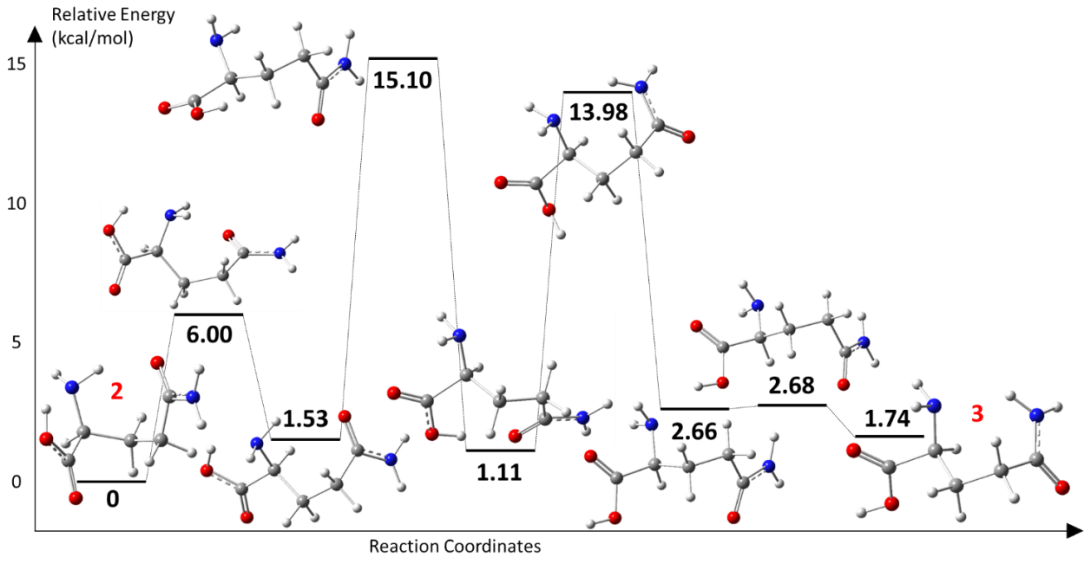



**Figure 1.3.** Interconversion path between rotamers 2 and 3. (B3PLYP-D3/m-aug-cc-pVTZ).

All starting rotamers and intermediates have been checked to be true minima and all transition states exhibited a single imaginary frequency.

As one can see on Figures 1.2, 1.3 and 1.4, the interconversion between rotamers **1** and **2** is much easier than both others. Indeed, it requires only five steps while the interconversion between **2** and **3**, and the one between **1** and **3** require seven or nine steps, respectively. Moreover, this first interconversion does not need to overpass the *ca.* 15 kcal/mol barrier that can be found on both other paths. This can be easily explained by the fact that the geometries of rotamers **1** and **2** are related while the geometry of **3** is different.

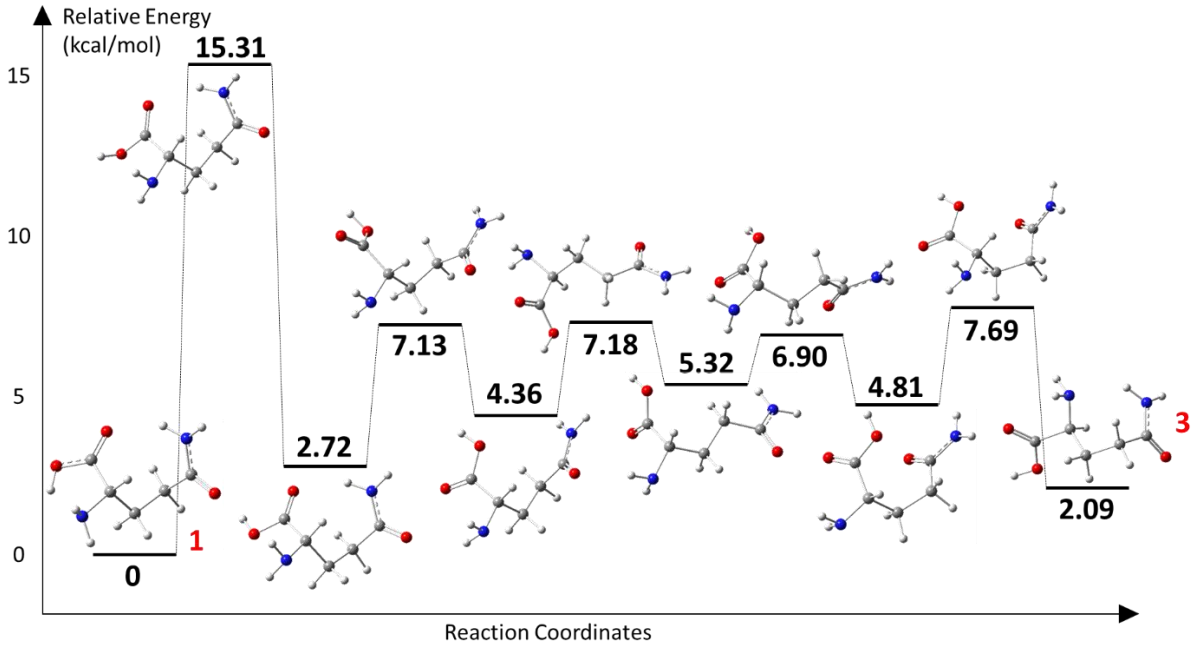



**Figure 1.4.** Interconversion path between rotamers 1 and 3. (B3PLYP-D3/m-aug-cc-pVTZ).

Indeed, the most important feature here is the orientations of the carboxylic acid moiety regarding the $NH_2$ one. In the case of **1** and **2**, it can be seen that the hydrogen atom that belongs to the –COOH moiety forms a hydrogen bond with the nitrogen of $NH_2$. Going from this orientation to the one that can be seen on **3** is actually the step that needs the reach the *ca.* 15 kcal/mol transition state. This "easy" interconversion between **1** and **2** could explain why they are not both detected. Furthermore, the measured values of rotational constants for the rotamer B are close to a mean computed values of **1** and **2**.

## 1.4. Partial conclusion

These interconversion paths have been able to help in giving a possible explanation on the detection on only 2 rotamers of glutamine in gas phase. New computations are ongoing on this topic.

# *Chapter 2. Borderline between reactivity and pre-reactivity of binary mixtures of gaseous carboxylic acids and alcohols*

One of the most important reactions in organic synthesis is esterification:

$$RCOOH + R'OH \rightleftarrows RCOOR' + H_20$$

Generally, a catalyst is needed to attain a satisfactory conversion rate together with the employment of one of the reactants in excess. In the present study, esterification takes place without a catalyst, using a very simple method employing a gas phase 1:1 mixture of carboxylic acids and primary and secondary alcohols.

Several molecular complexes involving carboxylic acids[214–221] have been investigated by rotational spectroscopy, in order to understand the nature of their non-covalent interactions and to have information on their internal dynamics and on their conformational equilibria. Most attention has been paid to the complexes of carboxylic acids, mainly to their dimers[214] and to their adducts with water.[215] Plenty of data have been obtained on the dimers, concerning proton tunneling, Ubbelohde effect and conformational equilibria.[214]

HCOOH (FA) is the prototype of the carboxylic acids family and for this reason it is involved in most of the investigations of carboxylic acids with molecules containing other functional groups, such as its adducts with $H_2O$,[215] $N(CH_3)_3$,[216] anhydrides,[214] formaldehyde ($CH_2O$),[217] formamide ($CH(CO)NH_2$),[218] ethers (dimethylether),[219] ketones (cyclobutanone),[220] esters (isopropyl formate) and azines (pyridine).[221]

However, no MW studies of complexes between carboxylic acids and alcohols are reported. For this reason, the prototype of this kind of complex, that is formic acid-methyl alcohol, has been investigated. Unexpectedly, assigning its rotational spectrum was not possible. Such a failure could be due, among to other reasons, to the complications related to the low $V_3$ barrier underlying the internal rotation of the methyl group, as well as to the inversion of the methyl group from above to below the formic acid plane. In order to understand what was going on, an investigation of the carboxylic acids-alcohols adducts from the formic acid adduct with a series of primary, secondary and tertiary alcohols was started. In detail, supersonic expansions were made, with *ca.* 1% of carboxylic acid and 1% of alcohol in He for the following combinations: HCOOH-$CH_3OH$, HCOOH-$C_2H_5OH$, HCOOH-$(CH_3)_2CHOH$

and HCOOH-$(CH_3)_3COH$, that is formic acid mixed with methyl alcohol and with primary, secondary and tertiary alcohols, respectively. In all cases but the last one, it was not possible to observe the spectra of the adducts, but strong rotational transitions that we discovered to belong to the esters. Then, linear alcohols were replaced with cyclic alcohols, like cyclohexanol (secondary) and 1-methyl-cyclopropanol (tertiary). Again, for the secondary alcohol it was possible to observe only the rotational spectrum of the ester, and for the tertiary alcohol only that of the adduct. Finally, the replacement of HCOOH with carboxylic acids with stronger ($CF_3COOH$) and weaker (pivalic acid) acidity was exploited. In the first case we observed only the ester, while in the second case the experiment did not succeed, because pivalic acid was rapidly obstructing our nozzle and causing serious problems to the operation of the pulsed valve. In summary, all these reactions have been studied:

$HCOOH + CH_3OH \rightarrow ESTER$ (1)

$HCOOH + C_2H_5OH \rightarrow ESTER$ (2)

$HCOOH + (CH_3)_2CHOH \rightarrow ESTER$ (3)

$HCOOH + (CH_3)_3COH \rightarrow ADDUCT$ (4)

$HCOOH + (CH_2)_5CHOH \rightarrow ESTER$ (5)

$HCOOH + (CH_2)_2CCH_3OH \rightarrow ADDUCT$ (6)

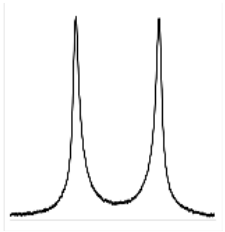
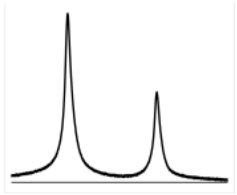
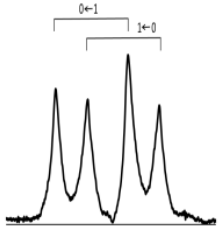
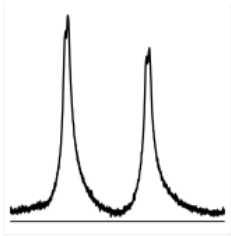
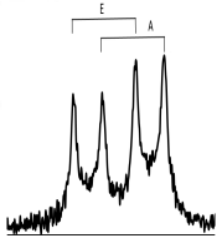
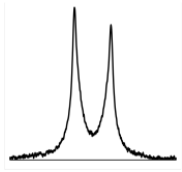
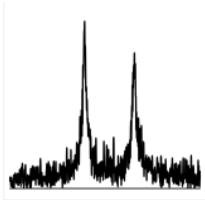
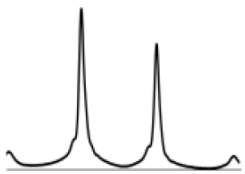
$CF_3COOH + CH_3OH \rightarrow ESTER$ (7)

$(CH_3)_3CCOOH + CH_3OH \rightarrow$ *no results* (8)

## 2.1. Rotational spectra

Figure 2.1 exhibits, as an analysis example, the case of the mixture of formic acid (FA) with isopropanol. In the upper part of the graphic, the expected spectrum of the ester is drawn in blue, while the calculated spectrum of the adduct is indicated in red, negative values. Only strong transitions of the ester have been identified in the rotational spectrum, according to the upper spectrum. In Table 2.1 is summarized the experimental evidence for the observed esters and adducts, respectively. As it usually happens, we observed some unassigned lines (in the broadband spectra there are even more unassigned lines due to spurious signals), but they do not have patterns which can match the expected spectra.

**Table 2.1.** Details (one transition example) of the spectra of the esters (first 5 columns) and of the adducts (last two columns) which have been formed in a supersonic expansion of binary mixture of carboxylic acids with primary and secondary alcohols, and with tertiary alcohols, respectively. All spectra cover a range of 0.4 MHz. The central frequency ($\nu_c$) and the number of accumulation cycles ($N_c$) of each spectrum are given.

| **Esters** | | | | | **Adducts** | |
|---|---|---|---|---|---|---|
| HCOOH + $CH_3OH$ → Methyl formate | HCOOH + $CH_3CH_2OH$ → Ethyl formate | HCOOH + $(CH_3)_2CHOH$ → Isopropyl formate | HCOOH + cy-$C_6H_{11}OH$ → cyclo-hexyl formate | $CF_3COOH$ + $CH_3OH$ → Methyl-trifluoroacetate | HCOOH + $(CH_3)_3COH$ → FA-*tert*Butanol | HCOOH + cy-$C_3H_4CH_3OH$ → FA-1-methyl-cyclopropanol |
|  |  |  |  |  |  |  |
| $4_{13}$←$4_{14}$ transition of the "*E*" species $\nu_c$ = 16037.30 MHz $N_c$ = 200 |  $2_{02}$←$1_{01}$ transitions of *gauche* (upper, $\nu_c$ = 14059.25 MHz $N_c$ = 1000) and *trans* (lower, $\nu_c$ = 10962.45 MHz $N_c$ = 1000) species | $2_{20}$←$2_{11}$ transition. Split into two tunneling components νc = 11836.5 MHz Nc = 50 | $7_{07}$←$6_{06}$ transition $\nu_c$ = 12942.45 MHz $N_c$ = 385 | $5_{05}$←$4_{04}$ transition Split into "*A*" and "*E*" species $\nu_c$ = 14036.80 MHz $N_c$ = 50. | $4_{13}$←$3_{12}$ transition $\nu_c$ = 7645.00 MHz $N_c$ = 1113 | $5_{14}$←$4_{13}$ transition $\nu_c$ = 8961.95 MHz $N_c$ = 504 |

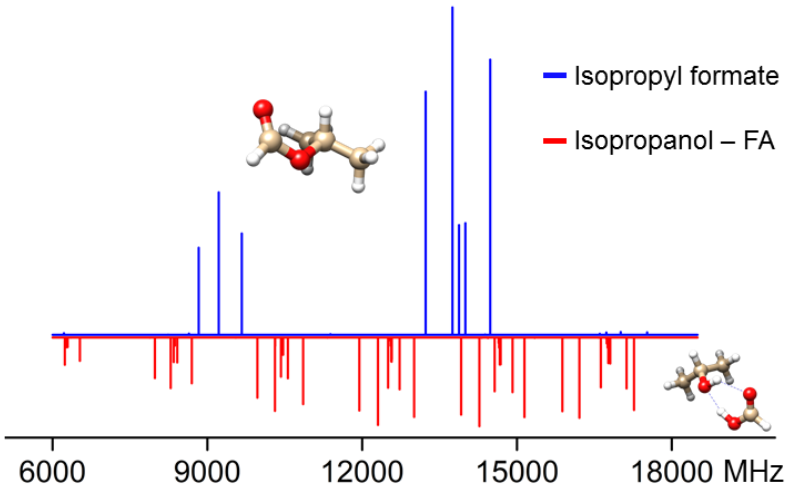



**Figure 2.1.** The rotational spectra expected for the mixture FA-isopropyl alcohol in the case of esterification (blue, positive values) and of formation of the adduct (red, negative values) are extremely different from each other. The experimental observation of one or the other spectrum shows unambiguously if the adduct or the ester is formed.

The rotational spectra of methylformate and ethyl formate are available in the literature.[222] In all cases the rotational transitions of the monomers (acid and alcohol) were present in the spectra. Their intensities decreased considerably, however, when we observed the ester.

## 2.2. Calculation details and results

In order to interpret the different behaviors of tertiary alcohols, chemical calculations have been performed using the B2PLYP-D3 functional,[24,29] in conjunction with the m-aug-cc-pVTZ basis set[26,27] where *d* functions on hydrogens have been removed. The BSSE has been corrected using the Counterpoise method.[212]

Energies of complexation and esterification have been obtained for the most significant examples listed above, including zero point corrections and, for the complexes, basis set superposition errors. In addition, also the free energy changes at 298 K have been evaluated. The obtained results are shown graphically in Figure 2.2 for FA-methanol, FA-ethanol, FA-isopropanol and FA-*tert*-butanol mixtures, and in Figure 2.3 for FA-cyclohexanol, $CF_3COOH$-ethanol, FA-1-methyl-cyclopropanol and $(CH_3)_3CCOOH$-ethanol

mixture. The starting point, with the acid and alcohol molecules separated from each other, is assumed to be the zero point energy (of free energy) value.

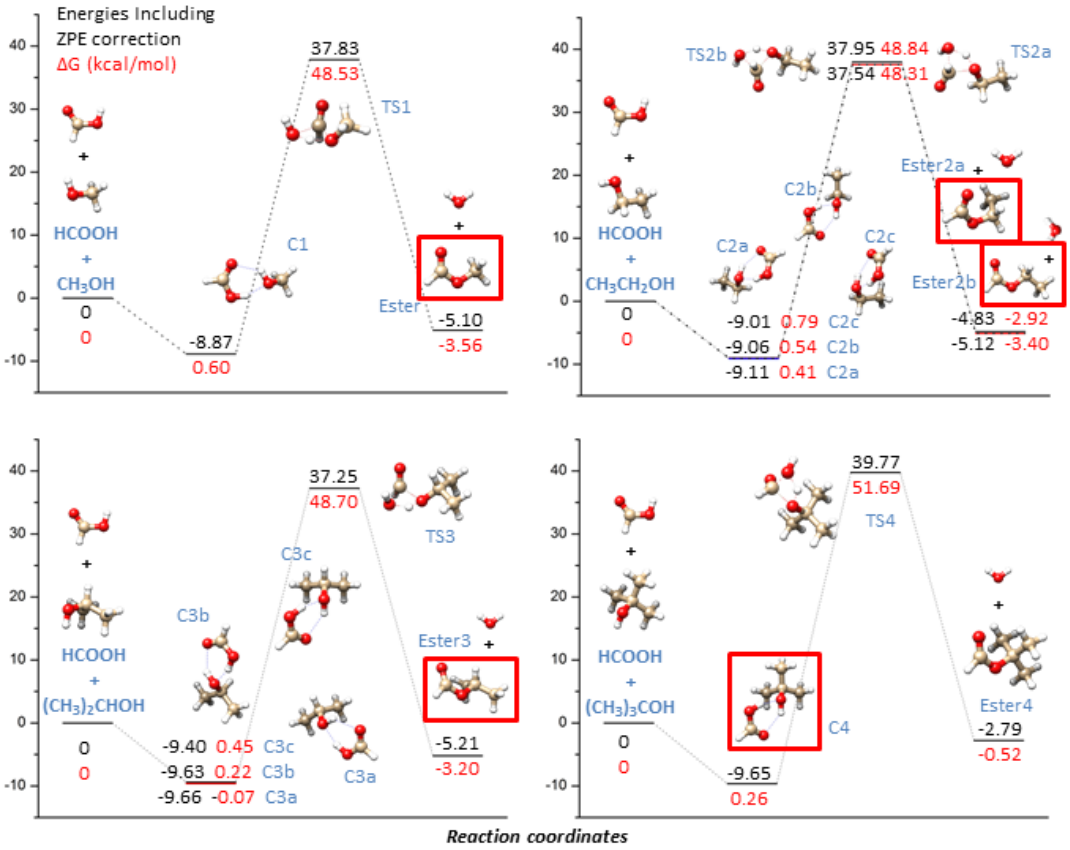



**Figure 2.2.** Energy profiles of the multiple chemical process paths for the FA-methanol, FA-ethanol, FA-isopropanol and FA-*tert*-butanol mixtures. The observed species are indicated with a rectangular frame.

It seems that the $\Delta G$ values are the most significant ones in interpreting the experimental evidence. Their changes ($\Delta G^0$) in going from the adducts to the ester are close to zero for the tertiary alcohols, but *ca.* -3 to -5 kcal/mol for the primary and secondary alcohols. In addition, the transition state values ($\Delta G^{0\ddagger}$) for the esterification process are about 3 kcal/mol higher in the case of tertiary alcohols. We have to outline that the formation of the dimer during the supersonic expansion is not a simple and linear process. For example, when more than one conformation is possible, the constituting molecules can reach the thermal quasi-equilibrium as a result of repeated dissociation and reformation of the dimers in the low

temperature molecular expansion. The same argumentations can be applied to the quasi-equilibrium conditions for the esterification reaction. Therefore, even a small free energy difference between different esterification reactions can turn out in completely different reaction advancement. In addition, it has to be mentioned that organic chemists pay much attention to the fact that in tertiary alcohols steric effects make the OH group less accessible than in the alcohols of the other classes, which results in a reduced propensity to undergo the esterification reaction. However, the first step is the formation of the adduct, and is detected only in the case of tertiary alcohols.

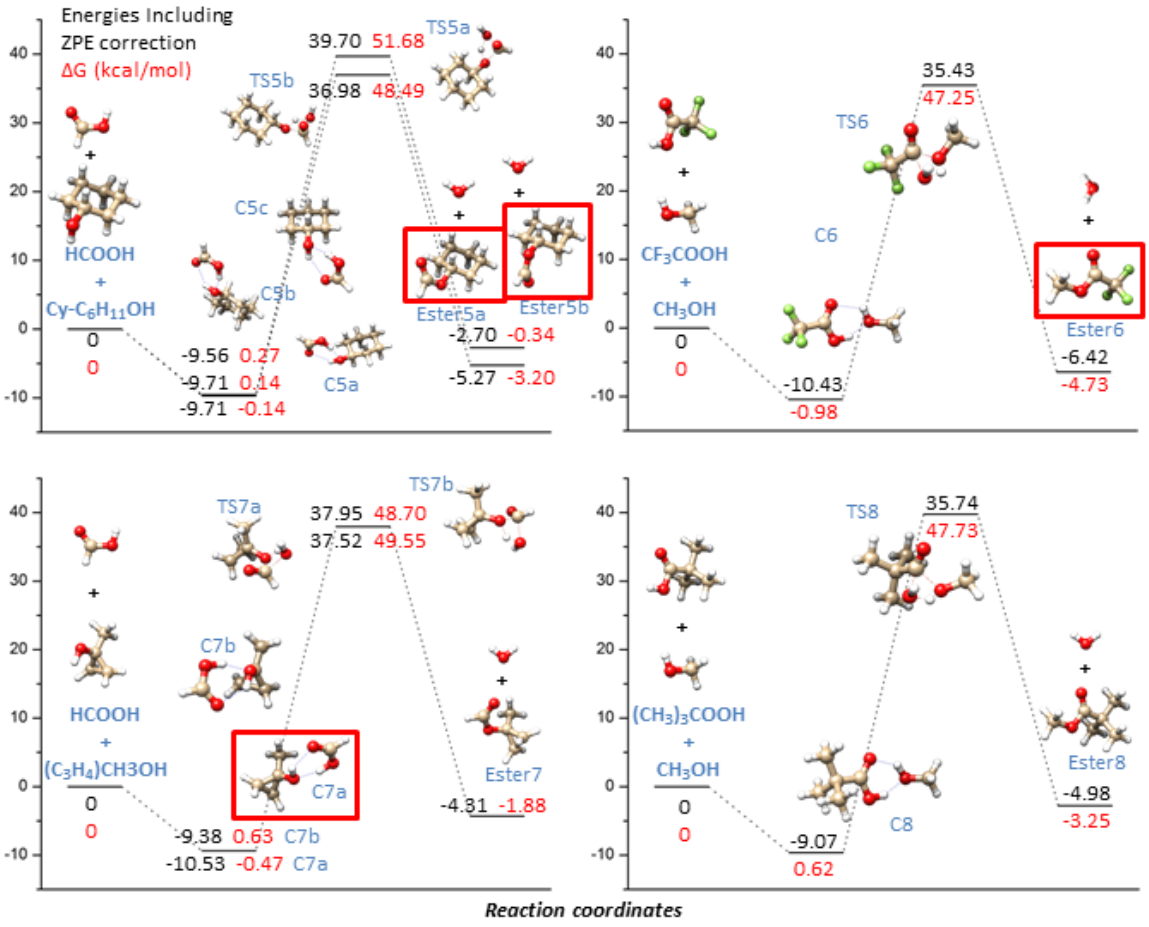



**Figure 2.3.** Energy profiles of the multiple chemical process paths for the FA-cyclohexanol, $CF_3COOH$-ethanol, FA-1-methyl-cyclopropanol and $(CH_3)_3CCOOH$-ethanol mixture. The observed species are indicated with a rectangular frame.

Finally, the fact that a third component, the carrier gas (He in our case), is constituting our mixtures in a concentration of about 99% needs to be taken into account. Its function is to convert the internal energy of the molecules into its own translational kinetic energy, up to reach a supersonic speed. It has often the effect, especially when it has a high atomic mass, to cause a conformational relaxation. When the inter-conversion barrier is smaller than 2kT, the less stable conformer disappears. This factor probably influences the esterification processes but not enough data are available to rationalize this effect.

## 2.3. Partial conclusion

This experimental investigation outlines a sharp cut-off of the esterification reaction in going from primary and secondary to tertiary alcohols. Theoretical molecular dynamics calculations supply an indication in support of the experimental evidence.

Finally, it has to be mentioned that experimentally, the use of fast-mixing nozzleswould probably allow the observation of the rotational spectra of the prereactive adducts also for primary and secondary alcohols.

# *Chapter 3. Interactions in organic complexes*

## 3.1. Non-covalent interactions and internal dynamics in pyridine-ammonia: a combined quantum-chemical and microwave spectroscopy study

Pyridine (Pyr) is a prototype heterocyclic aromatic molecule, of great importance in organic and industrial chemistry. The simultaneous presence of *n* and *π* orbitals makes the molecule particularly suitable for forming non-covalent linkages, and for this reason many of its molecular adducts have been investigated. A detailed knowledge of the typology and strength of these non-bonding interactions has been obtained using high resolution spectroscopy. For example, Pyr is the only molecule whose 1:1 adducts with all stable rare gases (from He to Xe[223]) and with all fluorine containing freons derived from methane ($CH_nF_{4-n}$, n = 0 to 4[224]) have been investigated by means of rotational spectroscopy. The results for these two families are graphically summarized in Figure 3.1.1. One can note that either π- or σ-type complexes are formed, depending on the type of the dominant interaction, dispersive or specific. Figure 3.1.1 shows that complexes with rare gases and methane belong to the first class, whereas those with freons, which can form weak hydrogen or halogen bonds, belong to the second class. Figure 3.1.1 also collects the dissociation energies, which range between 0.5 and 5 kJ/mol for the π-type complexes. It can also be noted that the π-interaction is generally weaker than the hydrogen or halogen bond (Pyr-$CF_4$), which ranges between 10 and 16 kJ/mol, with Pyr acting as donor. Furthermore, σ-type arrangements have been found in molecular adducts of Pyr with other partners, such as CO, $CO_2$, $SO_2$, $SO_3$ and acetylene. All of them are linked to the n non-bonding orbital of Pyr through formal C···N or S···N or C-H···N contacts.[225] In particular, it is possible to mention that Pyr-$SO_3$ is a charge transfer complex, where Pyr donates its pair of electrons to $SO_3$. The resulting S-N bond is thus a covalent one, with a corresponding bonding energy as large as ~120 kJ/mol.[226] Despite the fact that such a large variety of molecular adducts involving Pyr has been studied, the complexes with two very simple and important molecules, namely, water and ammonia, have not been investigated yet. For this reason, it is interesting to investigate the electronic structure and rotational spectrum of Pyr-$NH_3$. This study aims to answer the

following questions: does ammonia interact with the *n* or the *π* electronic system? Is there more than one type of stable conformer? What is the strength of this interaction? How much is the Pyr molecular structure affected by this interaction? Does $NH_3$ freely rotate around its symmetry axis in the complex? The importance of this investigation lies also on the fact that molecular complexes involving hydrogen bonding between nitrogen-containing systems can be considered as prototypes for studying the (N-H···N) interaction in biological systems.

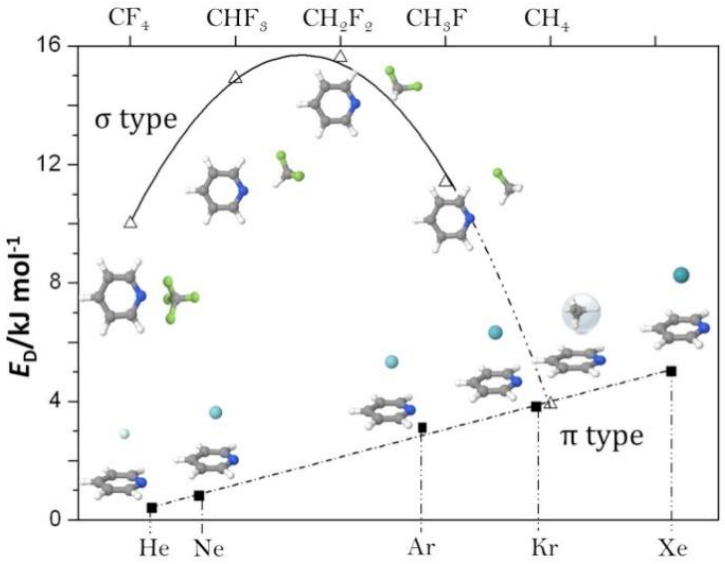



**Figure 3.1.1.** Pyr forms π-type complexes with rare gases and methane, and σ-type complexes with all the remaining molecules so far investigated.

To elucidate the structure of the Pyr-$NH_3$ complex and the type of interaction occurring, a combined theoretical-experimental investigation has been undertaken. Guided by state-of-the-art quantum-chemical calculations, the rotational spectrum of the molecular complex (further complicated by the quadrupolar interactions of the two $^{14}N$ nuclei with the overall rotation) has been recorded and analysed.

### *3.1.1. Computational details and results*

The exploration of the potential energy surface (PES) of the pyridine···ammonia complex has been carried out by employing density functional theory (DFT). Although

standard DFT can handle large molecular systems and its accuracy has been validated for several molecular properties, many density functional approximations fail to correctly describe the long range London dispersion interactions, particularly the correct $-R^{-6}$ dependence of the interaction energy on the intermolecular separation R.[227] As a consequence, dispersion correlation effects have been taken into account by the DFT-D3 correction scheme,[28] which combines reliability and computational effectiveness. Up to now, the DFT-D3 model has been successfully applied to thousands of different systems, including dimers, large supra-molecular complexes and reaction energies/barriers, as well as surface processes.[102,228,229] Within the DFT-D3 approach, Becke-Johnson damping has been used and the exploration of the pyridine···ammonia PES has been carried out using the B3LYP functional[22,23] in conjunction with the SNSD basis set.[68,101] These calculations have led to the identification of three distinct minima (**M1**, **M2** and **M3** in Figure 3.1.2), featuring hydrogen bonds between ammonia and pyridine. All of these minima show a σ-type arrangement, with the **M1** isomer being the most stable because of the double interaction.

A more thorough characterization of the PES, also including the saddle points (**TS1**, **TS2** and **TS3**) connecting the stable structures, has been carried out by using the double-hybrid B2PLYP-D3 functional[24] in conjunction with the m-aug-cc- pVTZ basis set[26,27] in which *d* functions on hydrogen atoms have been removed (referred to in the following as maug-ccpVTZ-*d*H). Figure 3.1.2 provides the graphical representation of the stationary points found on the PES at the B2PLYP-D3/m-aug-ccpVTZ-*d*H level.

An isomer with a π-type arrangement has also been searched on the PES. While an energy minimum has been found at the B3LYP-D3 level, this has not been confirmed by the more reliable B2PLYP-D3/m-aug-cc-pVTZ-*d*H calculations. Therefore, π-type arrangements have not further been considered.

According to the relative energies reported in Figure 3.1.2, the **M1** isomer is the best candidate to be observed in the rotational spectrum and, thus, to be characterized. Experimental observations and spectral analysis can be profitably guided by accurate predictions of the rotational constants, which, in turn require a highly accurate determination of the complex structure. To fulfil the required accuracy, the equilibrium geometry has been determined by means of a composite scheme that accounts

simultaneously for basis-set and electron-correlation effects.[230] In particular, the so-called "cheap geometry" scheme was employed.

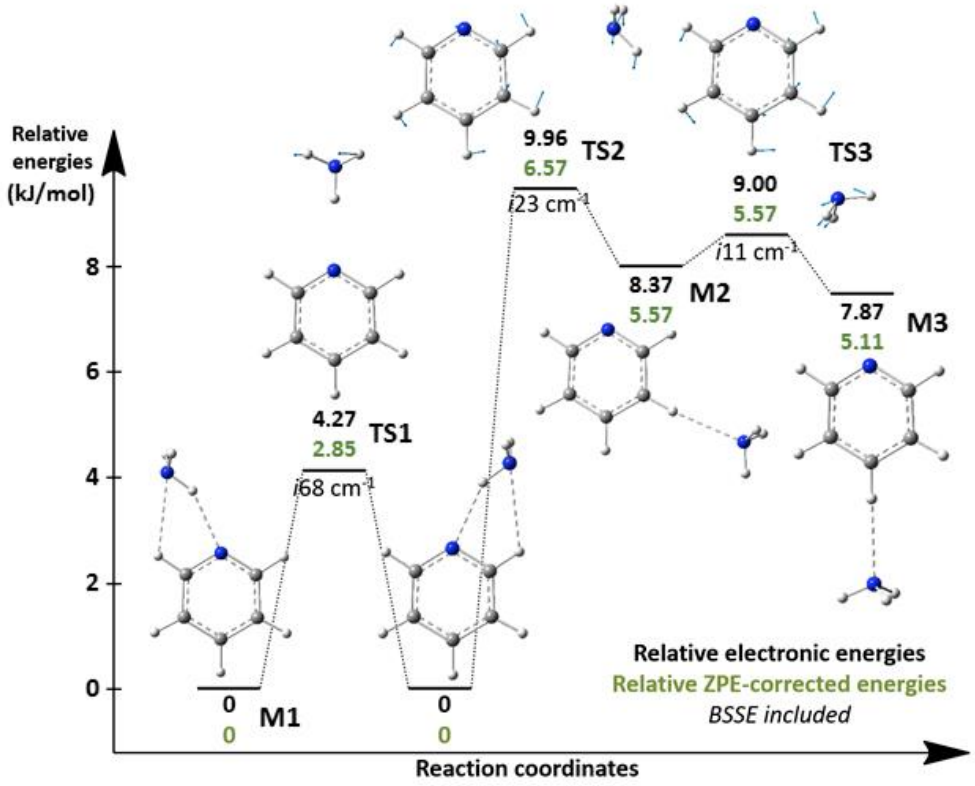



**Figure 3.1.2.** Schematic representation of PES for Pyr-$NH_3$: three different minima (**M1**, **M2**, and **M3**) and the connecting transition state structures (**TS1**, **TS2**, and **TS3**) are shown together with the corresponding relative energies (including the correction for the basis set superposition error, BSSE[212]), as obtained at the B2PLYP-D3/m-aug-cc-pVTZ-*d*H level.

The starting point is a geometry optimization at the CCSD(T)[231] level (within the frozen-core approximation): as a matter of fact, the coupled-cluster approach including single and double excitations (CCSD) with a perturbative treatment of triple excitations ((T)) is nowadays the method of choice when accurate results are sought. For a large system as the one under consideration, to include the most relevant contributions in the electronic structure treatment without losing computational efficiency, second-order Møller-Plesset perturbation theory (MP2)[232] geometry optimizations are used. In particular, the extrapolation to the complete basis set limit (CBS) has been performed by means of the

consolidated $n^{-3}$ extrapolation equation[147] applied to MP2 optimized geometries obtained with triple- and quadruple-zeta quality basis sets. The contribution of core-valence electron correlation has been included by means of the corresponding correction, $\Delta r$(CV), and the effect of diffuse functions in the basis set ($\Delta r$(aug)) has also been taken into account. On the whole, the best-estimated equilibrium structure has been determined as

$$r(best) = r(CCSD(T)/VTZ) + \Delta r(CBS) + \Delta r(CV) + \Delta r(aug) \quad (1)$$

Starting from the equilibrium rotational constants $B_e^i$, corresponding to the computed equilibrium geometry, best-estimated ground-state rotational constants $B_0^i$ have been obtained by adding the vibrational corrections, $\Delta B_{vib}$, evaluated at the B2PLYP-D3/m-aug-cc-pVTZ-*d*H level

$$B_0^i = B_e^i + \Delta B_{vib} \quad (2)$$

where *i* denotes the principal inertial axis. The vibrational corrections have been computed by means of second-order vibrational perturbation theory (VPT2) applied to the molecular $r_0$-vibrational Hamiltonian expressed in normal coordinates, which requires the calculation of semi-diagonal cubic force constants. Computation of the anharmonic force field (performed at the B2PLYP-D3/m-aug-cc-pVTZ-*d*H level) also provided the quartic centrifugal-distortion constants, for which the symmetric-top reduction (Watson's *S* reduction) has been considered.

A composite scheme analogous to that used for the structural determination has been employed for accurately predicting equilibrium nuclear quadrupole-coupling constants of both nitrogen atoms (one belonging to Pyr and one to NH3).[233,234] Subsequently, the ground-state quadrupole-coupling constants have been obtained by correcting the best-estimated equilibrium values through vibrational contributions evaluated at the B3LYPD3/
SNSD level.[37]

The computed spectroscopic results are summarized in Table 3.1.1, while selected geometrical parameters are given in Table 3.1.2.

### *3.1.2. Rotational spectrum*

Guided by the quantum-chemical calculations described above, the rotational spectrum recording started with the search of the most intense $\mu_a$-type rotational transitions of the most stable isomer (**M1**). Two separated lines, showing a complex hyperfine structure pattern due to the presence of two $^{14}$N nuclei (that possess nuclear spin $I$=1), have been found around 7504.85 and 7505.85 MHz. They have been assigned to the $J_{Ka,Kc} = 3_{0,3} \leftarrow 2_{0,2}$ transition of the $E$ and $A$ states, respectively, underlying the internal rotation of the $NH_3$ group around its $C_3$ axis. Later on, several a-type lines up to $J = 7$ and $K_a = 2$ and some weaker b-type transitions have been assigned and fitted using Pickett's SPFIT program,[235] within the $I^r$-representation of Watson's S reduction. To account for the nuclear spin of the two nitrogen atoms, the coupling schemes $F_1 = I_1 + J$ and $F = I_2 + F_1$ have been employed for both $A$ and $E$ states, where $I_1$ and $I_2$ are the quantum number of the $^{14}$N nuclear spin in pyridine and $NH_3$, respectively. An example of such a doubling is provided by Figure 3.1.3, for the $J_{Ka,Kc} = 6_{0,6} \leftarrow 5_{0,5}$ rotational transition.

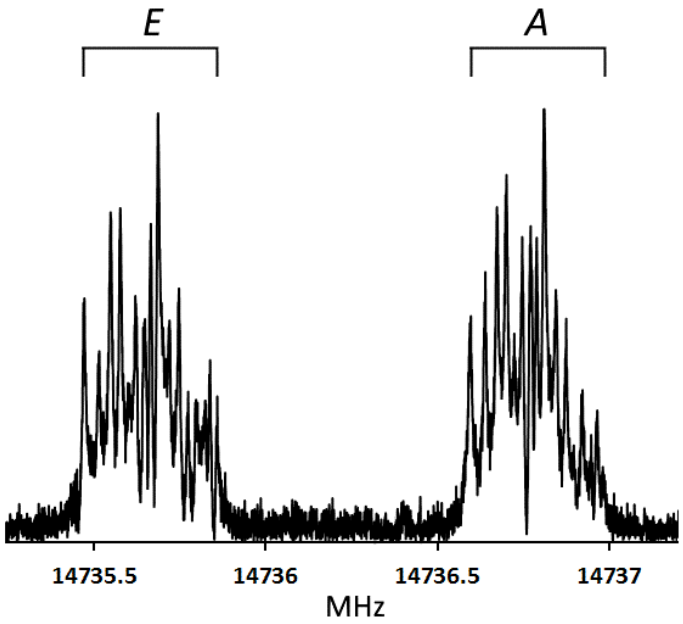



**Figure 3.1.3.** The $J = 6_{0,6} \leftarrow 5_{0,5}$ rotational transition for the observed isomer of pyridine-ammonia showing, for both $E$ and $A$ states, the hyperfine structure due to the presence of two $^{14}$N nuclei and the doubling due to the Doppler effect.

**Table 3.1.1.** Computed and experimental[1] (using SPFIT and XIAM programs) spectroscopic parameters of the M1 PYR-NH3 complex.

| | **Theory** | **SPFIT** | | **XIAM** |
|---|---|---|---|---|
| | | *"A"* | *"E"* | |
| $A$ (MHz) | 5875.20 | 5872.300(1) | 5872.136(1) | 5872.203(2) |
| $B$ (MHz) | 1386.28 | 1387.6497(2) | 1387.3647(2) | 1387.4676(3) |
| $C$ (MHz) | 1127.93 | 1128.7913(1) | 1128.7818(1) | 1128.7904(3) |
| $D_a$ (MHz) | | - | -15.575(2) | |
| $D_b$ (MHz) | | - | -22.23(3) | |
| $D_{pi2J}$ (MHz) | | | | -18(2) |
| $D_J$ (kHz) | 0.459 | 0.5364(6) | | 0.537(1) |
| $D_{JK}$ (kHz) | 7.929 | 7.95(1) | | 7.90(3) |
| $D_K$ (kHz) | -7.147 | | | |
| $d_1$ (kHz) | -0.105 | -0.1192(9) | | -0.121(2) |
| $d_2$ (kHz) | -0.0348 | -0.0357(4) | | -0.0346(9) |
| $X_{aa}$ (MHz) ($^{14}$N Pyr) | -4.00 | -3.746(5) | | |
| $X_{bb} - X_{cc}$ (MHz) ($^{14}$N Pyr) | -2.91 | -2.860(8) | | |
| $X_{ab}$ (MHz) ($^{14}$N Pyr) | 2.24 | 1.6(2) | | |
| $X_{aa}$ (MHz) ($^{14}$N $NH_3$) | 1.43 | 1.437(5) | | |
| $X_{bb} - X_{cc}$ (MHz) ($^{14}$N $NH_3$) | -4.58 | -4.788(8) | | |
| $X_{ab}$ (MHz) ($^{14}$N $NH_3$) | -0.47 | [-0.47][2] | | |
| $\angle$(I,a) (°)[3] | | | | 80.51(2) |
| $I$ (u.Å$^2$)[4] | | | | 2.551(7) |
| $V_3$ (cm$^{-1}$) | | | | 319.8(7) |
| *N5* | | 28 | | 28 |
| $\sigma$[6] (kHz) | | 3.0 | | 2.6 |

[1]Errors in parentheses are expressed in units of the last digit. [2]Fixed at the best-estimated equilibrium value augmented by vibrational corrections. [3]$\angle$(i,b) is the complement to 90° with respect to $\angle$(i,a) ($\angle$(i,c) is 90° because of the $C_S$ symmetry of the **M1** complex), where *i* is the $NH_3$ internal rotation axis. [4]Moment of inertia of NH3 around its C3 axis (α). [5]Number transition frequencies included in the fit. [6]Standard deviation of the fit.

The results of the fit are reported in Table 3.1.1 together with those obtained by fitting the centre frequencies of each transition (corrected for the nuclear hyperfine structure) using the XIAM program,[236] which provides directly the V3 internal rotation barrier. It is worth mentioning that the XIAM approach provides results corresponding to the ideal case of an infinite barrier, thus leading to intermediate values between the *A* and *E* cases, as dealt by the SPFIT code. Note that, for the *E* state, in the SPFIT fit it was also necessary to include terms of the type $D_aP_a$ and $D_bP_b$ in order to take into account the torsion-rotation interaction. Finally, it should be noted that no lines belonging to the other isomers have been observed.

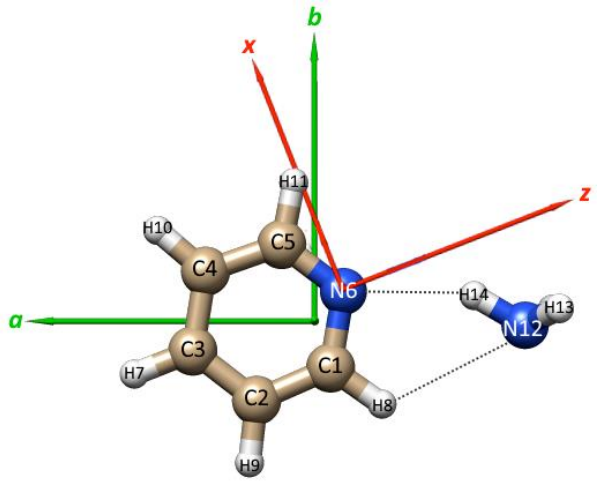



**Figure 3.1.4.** Orientation of the pyridine-ammonia cluster with respect to the principal axes of inertia *a*,*b*,*c* (the *c* axis is perpendicular to the figure plane), also showing the atom numbering and the N-H···N and C-H···N hydrogen bonds. The principal quadrupole axes *x*,*y*,*z* are also depicted, with *y* being parallel to *c*.

In addition to the intrinsic importance of the accurate determination of the spectroscopic parameters of the most stable Pyr-$NH_3$ isomer, Table 3.1.1 shows a very good agreement between experiment and theory. This is a particularly important result because the present study provides the first opportunity to test the reliability and accuracy of the "cheap geometry" scheme also for molecular complexes. While several investigations have demonstrated that this composite approach is able to determine accurate molecular and spectroscopic properties of semi-rigid as well as flexible systems,[233] the present application

to the Pyr-$NH_3$ complex suggests that the same accuracy is also obtainable for molecular adducts. This means that, for instance, the bond distances determined in this study are expected to be accurate within 0.001-0.002 Å. As evident in Table 3.1.1, the best-estimated equilibrium rotational constants augmented by vibrational corrections agree with the experimental ground-state rotational constants well within 0.1%.

In addition to the accurate calculation of the equilibrium structure of the **M1** isomer, from the analysis of the experimental data a partial effective structure has been derived by fitting the N12-N6-C5 angle of Figure 3.1.4 (∠N12N6C5 in Table 3.1.2), while keeping the other parameters fixed at the vibrationally-averaged B2PLYP/m-aug-cc-pVTZ-*d*H values. The results, reproducing the experimental rotational constants with a maximum discrepancy on *A* within 0.1%, together with the hydrogen bond parameters and a selection of the best estimated structural values, are reported in Table 3.1.2.

As already mentioned, the XIAM program directly provides the internal rotation barrier $V_3$. The value obtained, 319.8(7) $cm^{-1}$, is reported in Table 3.1.1. However, a second approach can be used to determine from our data the $V_3$ barrier. In fact, the $D_a$ and $D_b$ terms of Table 3.1.1 are related to the first-order pertubation sum relative to the *E* state, $W_E^{(1)}$, according to the following equations:[9]

$$D_g = F.\rho_g.W_E^{(1)} \qquad g = a,b \qquad (3)$$

where

$$F = \frac{h^2}{I_\alpha(1-\sum_g \frac{\lambda_g^2 I_\alpha}{I_g})} \qquad ; \qquad \rho_g = \frac{\lambda_g I_\alpha}{I_g}$$

with $\lambda_g$ denoting the direction cosines between the internal rotation axis of $NH_3$ and the principal axis $g$ (with $g=a,b$) of the complex, and $I_\alpha$ is the $NH_3$ moment of inertia. The value of $W_E^{(1)}$ (derived from $D_a$ and $D_b$) of the ground torsional level v=0 depends on the reduced barrier value *s*, which is related to both the $V_3$ potential energy barrier and the reduced mass of the motion. In particular, the $V_3$ barrier (in kcal/mol) is calculated via the following expression:

$$V_3 = 0.2146.F.s \qquad (4)$$

Using $D_a$ and $D_b$ of Table 3.1.1, we obtain the values $V_3^{(a)}$ = 321 cm$^{-1}$ and $V_3^{(b)}$= 343 cm$^{-1}$, respectively. These agree with the XIAM fitted $V_3$ value of Table 3.1.1 and are also in reasonably good agreement with the result obtained at the B2PLYP-D3/m-aug-cc-pVTZ-*d*H level including ZPE and BSSE corrections, $V_3$= 238 cm$^{-1}$.

**Table 3.1.2.** Effective $r_0$ and best-estimated equilibrium $r_e$ structural parameters.[1]

| | $r_0$ | $r_e$ |
|---|---|---|
| ∠N12N6C5 | 157.5(1)[2] | 154.6 |
| Derived Hydrogen bond parameters | | |
| $r_{N1H13}$ | 2.326 | 2.208 |
| $r_{N12H11}$ | 2.710 | 2.742 |
| ∠N12H14N6 | 150.5 | 150.8 |
| ∠C1H8N12 | 119.9 | 116.2 |
| $r_{N6C1}$ | 1.344 | 1.337 |
| $r_{N6C5}$ | 1.339 | 1.336 |
| $r_{C1C2}$ | 1.394 | 1.391 |
| $r_{C4C5}$ | 1.393 | 1.390 |
| $r_{C2C3}$ | 1.392 | 1.388 |
| $r_{C3C4}$ | 1.394 | 1.389 |

[1]Errors in parentheses are expressed in units of the last digit. [2]Quadrupole asymmetry parameter: $\eta = (\chi_{xx}-\chi_{yy})/\chi_{zz}$.

Important chemical information can be derived from the quadrupole-coupling analysis thanks to the availability of the full nuclear quadrupole-coupling tensor (in the principal inertia axes) for the $^{14}$N nucleus of pyridine (see Table 3.1.1). From the diagonalization of this tensor (by means of the QDIAG program[237]), the quadrupole coupling constants in the quadrupole principal coordinate system are obtained. The small decrease (up to a maximum of 0.6 MHz) of their values accompanying the formation of the Pyr-$NH_3$ complex (see Table 3.1.3) is a clear diagnostic of a small degree of N-H···N proton transfer. Further structural information can be derived from the rotation angles, obtained from the

matrix diagonalization mentioned above. These are $\theta_{za}$=19(2)° and $\theta_{xa}$=71(2)° ($\theta_{ya}$=90° for symmetry reasons), and provide useful information on the orientation of ammonia in the molecular complex.

**Table 3.1.3.** Principal quadrupole tensor[1] for the $^{14}$N (Pyr) nucleus.

| | **Pyr-NH$_3$** | **Isolated Pyr**[238] |
|---|---|---|
| $X_{zz}$ (MHz) | -4.3(1) | -4.908(3) |
| $X_{xx}$ (MHz) | 1.0(1) | 1.434(3) |
| $X_{yy}$ (MHz) | 3.304(5) | 3.474(3) |
| $\eta^2$ | 0.53(3) | 0.00 |

[1]Errors in parentheses are expressed in units of the last digit. [2]Quadrupole asymmetry parameter: $\eta = (\chi_{xx}-\chi_{yy})/\chi_{zz}$.

### *3.1.3. Dissociation energy*

From the partial effective structure determination it is furthermore possible to evaluate the distance between the effective centres of mass of pyridine and ammonia, $R_{CM}$, which turned out to be 4.48 Å. By assuming a pseudo-diatomic approximation, the force constant ($k_s$) for the intermolecular stretching along the axis connecting the two centres of mass, which is to a rough approximation the a inertial axis, can be estimated as:

$$k_s = \frac{16\pi^4(\mu R_{CM})^2[4B^4+4C^4-(B-C)^2(B+C)^2]}{hD_J} \qquad (5)$$

where $\mu$ is the pseudo-diatomic reduced mass. By using the B, C and $D_J$ values from Table 3.1.1, and $R_{CM}$ derived from the data of Table 3.1.2, a value of 10 N.m$^{-1}$ has been obtained for $k_s$, which corresponds to a harmonic fundamental frequency of 110 cm$^{-1}$. Subsequently, $k_s$ and $R_{CM}$ can be used, within a Lennard-Jones type potential, to determine the dissociation energy ($E_D$):[239]

$$E_D = \frac{1}{72}k_s(R_{CM})^2 \qquad (6)$$

thus leading to a value of 16.7 kJ/mol.

A more rigorous determination of the dissociation energy can be obtained via quantum chemistry. To accurately compute $E_D$, single-point energy calculations at the best-estimated equilibrium geometry of the complex and of the isolated non-interacting fragments, have been carried out by means of a composite scheme similar to that employed for geometry. Then, the best-estimated dissociation energy has been evaluated as the difference between the electronic energy of the Pyr-$NH_3$ complex and those of the isolated monomers, with the BSSE taken into account via the counterpoise correction (CP) procedure, proposed by Boys and Bernardi.[212] The composite approach have led to a value of 16.7 kJ/mol, with an estimated uncertainty around 1-2 kJ/mol. It is noteworthy that, despite the limitations of the pseudo-diatomic approximation, the cancellation of errors can help to get a dissociation energy obtained by this approximation in remarkable agreement with the quantum-chemical result. Subsequently, electronic energies have been corrected for the anharmonic ZPE contribution calculated at the B3LYP- D3/SNSD level by adopting the resonance-free perturbative approach within the hybrid degeneracy corrected second-order perturbation theory (HDCPT2). ZPE energy turns out to reduce the dissociation energy by 5.7 kJ/mol, thus leading to a ZPE-corrected value for the dissociation energy of 10.9 kJ/mol.

### *3.1.4. Partial conclusion*

The joint experimental-computational investigation reported here provides incontrovertible evidence that the Pyr-$NH_3$ complex shows a σ-type bonding structure in which ammonia interacts via two hydrogen bonds within the plane of the pyridine ring, with a small degree of proton transfer from ammonia to pyridine. No evidence was obtained that supports the formation of a stable π-bound adduct. For the first time, the so-called “cheap” composite scheme, purposely set up for accurately describing the electronic structure and spectroscopic properties of small biomolecules, has been applied to a weakly bonded molecular complex and its accuracy and reliability tested by means of a high-resolution MW spectroscopic study. The results reported demonstrated the suitability of this approach also for accurately studying molecular complexes. In detail, the molecular structure has been determined by means of a computational approach able to provide structural parameters with an estimated accuracy of about 0.001-0.002 Å for bond distances and of about 0.1° for

angles. The binding energy has also been determined by means of a composite scheme analogous to that employed for the evaluation of equilibrium geometry. The best-estimated, ZPE and BSSE corrected, dissociation energy is 10.9 kJ/mol and can be considered a reliable reference value, undoubtedly fulfilling chemical accuracy requirements. This value is perfectly in line with those previously derived for σ-type pyridine -$CH_nF_{4-n}$ freons complexes.

## 3.2. On the competition between weak O-H···F and C-H···F hydrogen bonds, in cooperation with C-H···O contacts, in the difluoromethane – tert-butyl alcohol cluster

Weak hydrogen bond (WHB) is, together with hydrogen (HB) and halogen bond (HaB), the most widespread non-covalent interaction encountered in nature, governing processes such as molecular association,[240,241] microsolvation of organic molecules,[221,242] molecular recognition[243] and many others. Their X-H···Y layered contacts (where X and Y are the hydrogen bond donor and acceptor, respectively) are characterized by similar directionality of the HB, showing, however, blue shifts of the stretching frequencies of the X-H groups and dissociation energies of only a few $kJ.mol^{-1}$. Among WHBs, the interactions involving organic fluorine as hydrogen bond acceptor are of great interest because of the role they play, for example, in pharmaceutical, in functional materials[244] and in polymerization[245] fields.

Even if the weakness of organic fluorine as hydrogen bond acceptor (HBa) is well established,[246–248] the formation of C-H···F-C or O-H···F-C WHBs is the mechanism at the basis of the crystal packing of fluoroaromatic compounds[248] or the stabilization in solid, alkane solution,[249] and in the gas phase of some fluoroalcohol compounds,[250,251] respectively, for the latter in association with O-H···O HBs. In this respect, when a competition is possible between these two WHBs involving organic fluorine as HBa, which of the aforementioned non-covalent interactions is favorite? Aiming at answering to this question, a combined quantum-chemical and rotational study of the difluoromethane (DFM) - *tert*-butyl alcohol (TBA) cluster is reported, where the competition between C-H···F-C and O-H···F-C WHBs, in combination with C-H···O contacts, is possible. The capability of rotational spectroscopy in a supersonic expansion with high-level quantum-chemical calculations were combined to explore the potential energy surface of the dimer and to unveil the genuine nature of these WHBs in an environment free from solvent and matrix effects. Furthermore, thanks to a previous microwave characterization of the O-H···F-C WHB, together with two C-H···O interactions, in the DFM-water complex,[240,252,253] the influence of substituent (tert-butyl of TBA vs hydrogen of water) on the cluster bonding energy can be also directly evaluated.

### *3.2.1. Computational details*

This study began with exploratory quantum-chemical calculations of the relevant stationary points on the potential energy surface (PES) of the DFM-TBA complex employing density functional theory (DFT). To this purpose, the B3LYP functional[22,23] augmented by dispersion corrections according to Grimme's DFT-D3 scheme[28] in conjunction with the polarized double-$\xi$ SNSD basis set[68] was used. It needs to be noted that the D3 correction, which has been tested over a wide range of molecular systems,[102,228,229,254–256] is fundamental for describing the thermochemistry of van der Waals adducts with DFT, because most of the density functionals fail to correctly describe long-range dispersion correlation effects.

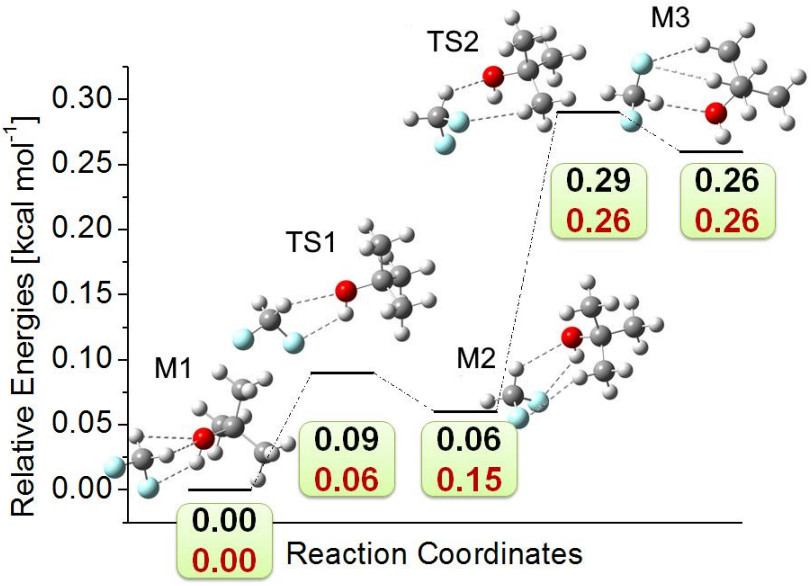



**Figure 3.2.1.** Potential energy surface of the DFM – TBA cluster showing the structures of the three minima (**M1**, **M2**, **M3**) and the two transition states (**TS1**, **TS2**). The best-estimated relative electronic energies (black values) and the best-estimated relative electronic energies corrected by anharmonic ZPE computed from hybrid B2PLYP-D3/m-aug-cc-pVTZ-dH/B3LYP-D3/SNSD force field (red values) are also reported.

Geometry optimizations were carried out first, and then at the optimized structure harmonic frequencies were computed through analytical hessians.[257] In addition, for each

identified stationary point, in order to compute anharmonic contributions to vibrational frequencies, but also vibrational corrections to rotational constants (of interest for the minima structures), cubic and quartic semi-diagonal force constants were calculated in a normal mode representation by numerical differentiation of analytic quadratic force constants (a step size of 0.01 Å was used for displacements).

In a subsequent step, for all stationary points, optimized structures and harmonic force fields were further refined using the double-hybrid B2PLYP-D3 functional[24] coupled to the m-aug-cc-pVTZ basis set,[26,27] in which *d*-functions on H atoms were removed. In the following, this basis set is referred to as m-aug-cc-pVTZ-dH. Figure 3.2.1 provides a graphical overview of the stationary points located on the PES and of their interconnection. To evaluate improved zero-point vibrational energies (ZPE) beyond the harmonic approximation, B2PLYP-D3/m-aug-cc-pVTZ-dH harmonic frequencies were then coupled with B3LYP-D3/SNSD cubic and semi-diagonal quartic force constants to formulate a hybrid B2PLYP/B3LYP anharmonic force field. To determine ZPE, the hybrid degeneracy corrected perturbation theory to second order (HDCPT2) was employed.[37]

The equilibrium structures at the B2PLYP-D3/m-aug-cc-pVTZ-dH level were used as starting point for computing vibrational ground-state rotational constants and accurate electronic energies. In fact, equilibrium structures straightforwardly provide equilibrium rotational constants $B_e^\alpha$, which were corrected by adding the vibrational contributions $\Delta B_0^\alpha$, thus obtaining the ground-state rotational constants $B_0^\alpha$:

$$B_0^\alpha = B_e^\alpha + B_0^\alpha \qquad (1)$$

where α denotes the principal inertia axis ($\alpha$ = $a$, $b$, $c$; thus $B_0^\alpha$ = $A_0$, $B_0$, $C_0$). Vibrational corrections were obtained by means of second-order vibrational perturbation theory (VPT2)[258] applied to the molecular roto-vibrational Watson Hamiltonian expressed in normal coordinates.[259]

Accurate electronic energies were obtained as single-point computations at the B2PLYP-D3/m-aug-cc-pVTZ-dH optimized structures by means of the composite scheme described here below. In this scheme, the different contributions are evaluated at the highest level possible, yet maintaining the overall computational cost affordable in terms of both hardware requirements and time. This computational protocol takes the coupled-cluster

theory employing the singles and doubles approximation augmented by a perturbative treatment of triple excitations CCSD(T),[141,231] as a starting point. The energy at the CCSD(T)/cc-pVTZ[260] level is then corrected for basis-set truncation error and core-correlation effects. The latter two contributions are evaluated at second-order Møller-Plesset perturbation theory (MP2).[232] The best estimate of the electronic energy is then given by the equation:[163,234]

$$E_{best} = E(CCSD(T)/VTZ) + \Delta E(MP2/CBS) + \Delta E(MP2/CV) \quad (2)$$

The $\Delta E$(MP2/CBS) is obtained by extrapolating the Hartree-Fock (HF-SCF) electronic energy $E_{\infty}^{HF-SCF}$, to the complete basis set (CBS) limit according to the $e^{-Cn}$ functional form[146] using the cc-pVnZ with n = T, Q, 5 basis sets,[145,260,261] and the MP2 correlation contribution by means of the $n^{-3}$ formula[147] employing the cc-pVTZ and cc-pVQZ basis sets. $\Delta E$(MP2/CV) represents the core-valence correlation correction, and is obtained as difference of MP2/cc-pCVTZ[261] energies evaluated correlating all electrons (a.e.) and within the frozen-core (f.c.) approximation:

$$\Delta E(CV) = E[(MP2/pCVnZ(a.e.)] - E[MP2/pCVnZ(f.c.)] \quad (3)$$

The same protocol was also applied to obtain best-estimate electronic energies of the isolated fragments, *i.e.*, DFM and TBA, in order to evaluate best-estimate dissociation energies as the difference between the electronic energy of the DFM-TBA complex and those of the isolated monomers. The basis set superposition error (BSSE) was taken into account by means of the counterpoise correction (CP) procedure, introduced by Boys and Bernardi.[212]

Finally, for the most stable isomer, the cubic and semi-diagonal quartic force field was also evaluated at the B2PLYP-D3/m-aug-cc-pVTZ-dH level in order to derive the partial effective structure $r_0$.

All DFT, HF-SCF and MP2 calculations were carried out by using a development version of the Gaussian software,[21] while CCSD(T) computations were performed with the CFOUR program package.[143]

### *3.2.2. Experimental methods and Rotational spectrum*

**Table 2.2.1.** Rotational parameters and dipole moment components for the **M1**, **M2** and **M3** isomers of DFM-TBA.

| | **M1** | | **M2** | **M3** |
|---|---|---|---|---|
| | *Experiment*[a] | *Theory*[b] | *Theory*[b] | *Theory*[b] |
| $A$ (MHz) | 3602.299(3) | 3570 | 3137 | 3064 |
| $B$ (MHz) | 615.0150(1) | 601 | 780 | 772 |
| $C$ (MHz) | 597.1948(1) | 584 | 742 | 732 |
| $D_J$ (kHz) | 0.1630(1) | 0.15 | 1.33 | 0.98 |
| $D_{JK}$ (kHz) | 7.672(2) | 7.63 | 7.45 | 6.29 |
| $D_K$ (kHz) | -2.8(7) | -3.86 | -6.90 | -4.28 |
| $d_1$ (Hz) | -5.2(2) | -5.78 | -104.37 | -76.24 |
| $d_2$ (Hz) | 4.3(1) | 4.47 | 3.73 | 3.46 |
| $\sigma$ [c] (kHz) | 2.8 | | | |
| $N$ [d] | 105 | | | |
| $\lvert\mu_a\rvert$[e] (D) | | 2.55 | 1.49 | 0.34 |
| $\lvert\mu_b\rvert$[e] (D) | | 0.63 | 0.49 | 0.04 |
| $\lvert\mu_c\rvert$[e] (D) | | 0.00 | 0.78 | 0.42 |

[a]Standard error in parentheses in units of the last digit.
[b]B2PLYP-D3/aug-cc-pVTZ-dH equilibrium parameters. Equilibrium rotational constants have been augmented by vibrational corrections at the B3LYP-D3/SNSD level.
[c]Standard deviation of the fit.
[d]Number of fitted transitions.
[e]Absolute values of the dipole moment components: B2PLYP-D3/aug-cc-pVTZ-dH equilibrium values augmented by vibrational corrections at the B3LYP-D3/SNSD level.

The rotational spectrum of the 1:1 DFM:TBA cluster has been recorded using a COBRA-type[262] pulsed-jet Fourier-transform microwave (FTMW) spectrometer,[263,264] operating in the 6.5 - 18.5 GHz range. A gas mixture at 1% of $CH_2F_2$ (DFM) in helium at a stagnation pressure of 0.2 MPa has been passed over a sample of *tert*-butyl alcohol (TBA), used without further purification, and expanded into the Fabry-Perot cavity, through the pulsed valve (General valve, series 9, nozzle diameter 0.5 mm). In accordance with the quantum-chemical calculations, the study was started with the search of the rotational transitions belonging to the isomer **M1**, which is predicted to be the most stable one (see

Figure 3.2.1), having a large value of the dipole moment component $\mu_a$ (2.6 D). Three rotational transitions around 7218.9, 7270.3 and 7325.8 MHz have been first observed and assigned to the J = 6 ← 5 band, with $K_a$ = 0, 1, of the **M1** isomer. Later on, 93 more $\mu_a$-R-type lines up to $J_{upper}$ = 15 and $K_a$ = 6 and other 9 much weaker $\mu_b$-R-type transitions have been identified and fitted using Pickett's SPFIT program[235] within Watson's semirigid Hamiltonian (*S*-reduction; $I^r$-representation). The results of the fit are reported in Table 2.2.1. No $\mu_c$-type lines have been observed, in agreement with the $C_s$ symmetry of the cluster. The search for rotational transitions belonging to the **M2** and **M3** isomers failed due to plausible relaxation processes toward **M1**, taking place when the interconversion barrier values are lower than 2*k*T at room temperature,[265] as suggested by the data reported on Figure 3.2.1. The rest frequencies have been obtained as the arithmetic mean of the two component frequencies. The estimated accuracy of frequency measurements is better than 3 kHz, and the resolution is better than 7 kHz.

### *3.2.3. Results*

Table 3.2.1 collects the experimental rotational parameters of the **M1** isomer as obtained in the fitting procedure described above. This table also reports the corresponding computed values together with those of the **M2** and **M3** isomers. It needs to be noted that the experimental rotational constants as well as the $D_J$ and $D_{JK}$ quartic centrifugal-distortion terms have been determined to a very good precision, while the other centrifugal-distortion constants are affected by larger uncertainties, ranging from 3% to 25%. Furthermore, a good agreement can also be noted between the computed values and the experimental counterparts: *A* is predicted well within 1%, while deviations of about 2% are observed for *B* and *C*. According to the results of Table 3.2.1 as well as based on the relative energies collected in Figure 3.2.1, there is no doubt concerning which isomer has been experimentally observed. Indeed, the isomerization path graphically summarized in Figure 3.2.1 explains the experimental outcome: at the equilibrium, both the **M3**-**M2** and **M2**-**M1** interconversion barriers are predicted to be only 0.03 kcal/mol at the best-estimated (composite scheme, see Section 3.2.1) level. Even if the **M2** isomer is predicted to be less stable than **M1** by only 0.06 kcal/mol, the small interconversion barrier suggests that it easily relaxes to **M1**, thus

preventing its observation. Even more significant are the conclusions that can be drawn once the equilibrium relative energies are corrected for ZPE (hybrid force field, see Section 3.2.1). Indeed, ZPE-corrected energies (see Figure 3.2.1) suggest that the vibrational ground state of the **M2** isomer is above the **TS1** barrier, while that of **M3** is as high in energy as **TS2**. The overall picture thus suggests that both **M2** and **M3** relax to **M1**, as experiment tends to support.

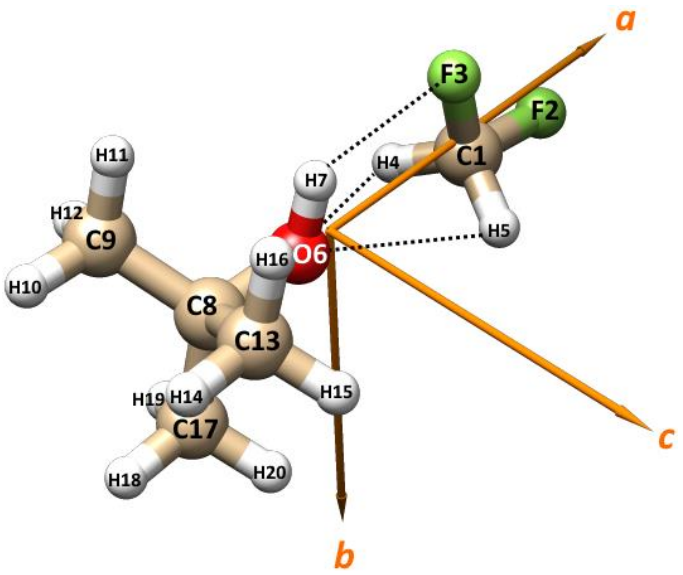



**Figure 3.2.2.** Structure (with atom numbering) of the **M1** isomer of DFM-TBA orientated with respect to the principal axes of inertia (*a*,*b*,*c*)

According to the molecular structures depicted in Figure 3.2.1, the energies of the **M1**, **M2** and **M3** isomers are related to the number and type of WHBs involved in the cluster. In the case of **M3**, two C-H···F and one C-H···O interactions are established. The stabilization in energy moving to the **M2** isomer can be explained in terms of the replacement of one C-H···F WHB with one O-H···F interaction. A further stabilization is then obtained once also the remaining C-H···F WHB is replaced by an additional O-H···F bond. We can thus conclude that the most stable structure (**M1**) is obtained when the largest number of C-H···O contacts are established together with the O-H···F WHB.

In the context of this study, the more interesting parameter is the partial effective structure ($r_0$) of **M1** that can be derived from our data. This has been obtained by fitting two

geometrical parameters (the distance $r_{O6\cdots C1}$ and angle O6···C1-F3 of Figure 3.2.2) while keeping the other ones fixed at their vibrationally averaged B2PLYP-D3/m-aug-cc-pVTZ-dH values. In this way, the experimental rotational constants are perfectly reproduced, with the largest discrepancy, that on the *A* rotational constant, being well within 0.1%. The resulting values are reported in Table 3.2.2, where they are compared with the ones relative to the equilibrium structure ($r_e$).

**Table 3.2.2.** Comparison between the fitted effective structure and the B2PLYP-D3/m-aug-cc-pVTZ-dH calculated values of the observed (**M1**) DFM-TBA cluster.

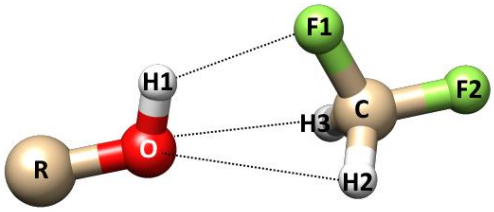
| | $\mathbf{r_0}$ | | $\mathbf{r_e}$ |
|---|---|---|---|
| | *Exp* | *Theory* | |
| $r_{O6\cdots C1}$ (Å) | 2.960(2) | 2.993 | 2.966 |
| ∠O6···C1-F3 (°) | 73.7(1) | 74.6 | 74.2 |

The conclusion that can be drawn is that the data of Table 3.2.2 further confirms that the level of theory chosen is able to well describe the structure of the molecular complex.

**Table 3.2.3.** Comparison of the parameters involved in the WHBs of the DFM-TBA and DFM-water complexes (R = H- or *tert*-butyl-).

| | DFM-TBA | | DFM-water | |
|---|---|---|---|---|
| | $r_0$ | $r_e$ | $r_0$[a] | $r_e$ |
| rH1···F1 (Å) | 2.384 | 2.360 | 2.20(1) | 2.306 |
| $r_{H2\cdots O} = r_{H3\cdots O}$ (Å) | 2.781 | 2.784 | - | 2.834 |
| ∠O-H1···F1 (°) | 113.1 | 115.6 | 135(3) | 117.8 |
| ∠C-F1···H1 | 97.0 | 96.2 | 93(1) | 98.4 |

[a]From Ref. [253]




The results obtained here allow to discuss the hydrogen bond parameters in comparison with those of the DFM-water complex. The relevant parameters are collected in Table 3.2.3. According to the partial $r_0$ structure of Table 3.2.3, it is noted that when going from water to TBA as counterpart molecule in the cluster with DFM, the O-H···F distance increases by about 0.2 Å (2.20 Å *vs.* 2.384 Å). This is only in apparent contradiction with the stronger binding energy of the DFM-TBA complex. In fact, it is seen that at the B2PLYP/m-

aug-cc-pVTZ-dH level, the O-H···F distance increases by a smaller extent (2.306 Å *vs.* 2.360 Å). Furthermore, for the C-H···O interaction, a shrinkage of the corresponding length is predicted when going from DFM-water to DFM-TBA (2.834 Å *vs.* 2.784 Å). Furthermore, in the case of the water complexation of DFM, the internal rotation of water around its $C_2$ axis generates a tilted angle of 23° between the water plane and the ∠FCF angle of DFM, instead $C_s$ symmetry is observed for the DFM-TBA adduct. The effect of the aforementioned motion affects, in particular, the O-H1···F1 angle value, while only a slight difference is noted for the C1-F1···H1 angle (see Table 2.3.3) of both the DFM-water and DFM-TBA complexes.

The last piece of information that can be obtained from this study is the dissociation energy of the complex. Based on the isomer structure depicted in Figure 3.2.3, the so-called "pseudo-diatomic approximation" can be assumed to estimate the force constant ($k_s$) of the intermolecular stretching between the centers of mass of DFM and TBA, because its axis is almost parallel to the *a*-axis of the cluster. According to this approximation, $k_s$ can be written as:

$$k_s = \frac{16\pi^4(\mu R_{CM})^2[4B^4+4C^4-(B-C)^2(B+C)^2]}{hD_J} \qquad (4)$$

where $\mu$ is the pseudo-diatomic reduced mass, $R_{CM}$ the distance between the centers of mass of the two monomers (4.73 Å), and $B$, $C$ and $D_J$ the experimental values reported in Table 3.2.1. The $k_s$ value obtained by applying Eq. 4 is 9.0 N.m$^{-1}$. Assuming a Lennard-Jones type potential, the dissociation energy is calculated according to the equation:[266]

$$E_D = \frac{1}{72} k_S (R_{CM})^2 \qquad (5)$$

thus obtaining a value of 4.02 kcal.mol$^{-1}$. This value is in good agreement with the best-estimated ZPE-corrected value, 3.56 kcal.mol$^{-1}$, which is expected to have an accuracy of ~0.5 kcal.mol$^{-1}$.[234] Such an agreement suggests that, despite its limitations, the "pseudo-diatomic approximation" is able to provide a reliable determination of the dissociation energy. It is noted that this value is much larger than that estimated for DFM-water (1.79 kcal.mol$^{-1}$).[253] This suggests that the size of the residual groups of the R-O-H hydrogen bond donor sensibly affects the stability of the clusters linked by one O-H···F and two C-H···O WHBs. This suggestion was further supported by computations at the B2PLYP-D3/m-aug-cc-pVTZ-dH

level. At this level of theory, the dissociation energy corrected by the ZPE contribution (evaluated using the B2PLYP/B3LYP hybrid force field) is 2.17 kcal.mol$^{-1}$, to be compared with the value of 3.19 kcal.mol$^{-1}$ obtained for DFM-TBA at the same level of theory.

### *3.2.4. Partial conclusion*

The 1:1 complex of *tert*-butyl alcohol with difluoromethane has been accurately characterized by means of experimental and computational techniques. Starting from an accurate sampling of the potential energy surface, the structure and energetics of the complex have been obtained from the investigation of its rotational spectrum, which has been guided and complemented by accurate quantum-chemical calculations. The $C_s$ symmetric **M1** isomer, whose stability is due to the formation of one O-H···F and two C-H···O WHBs, has been experimentally confirmed to be the global minimum. The failure in the observation of the rotational spectra of the **M2** and **M3** isomers, whose monomers are linked by one O-H···F, one C-H···F, one C-H···O and two C-H···F, one C-H···O WHBs, respectively, has been plausibly ascribed to the relaxation processes toward the **M1** isomer taking place in the supersonic expansion.

The H → tert-butyl substitution effect, observed when going from DFM-water to DFM-TBA 1:1 cluster, has been evaluated, resulting in an increase of the bonding energy of the adduct. The weakness of the hydrogen bonds established in the complex is further confirmed by the energy dissociation, which has been accurately computed to be 3.6 kcal.mol-1, value also confirmed by an analysis based on the so-called “pseudo-diatomic approximation”.

# *General Conclusions.*

In this thesis, after simulating vibrational and optical properties of several organic compounds including one (or more) C≡N moiety(ies) and transition metal complexes in order to evaluate and confirm the adequacy of the utilized level of theory, the issue of the formation of COMs in ISM was faced. Indeed, as they are detected in several spots and can be considered as precursors for the building blocks of life, they are of major interest in prebiotic chemistry. The main questions this thesis therefore focused on are: How are the COMs formed in the hard conditions of ISM? Even though their formation was mostly believed to occur on dust grain, does surface chemistry really need to be always invoked? Different molecules of high prebiotic interest were considered as test cases: cyanomethanimine isomers (*E*-C-, *Z*-C- and N), formamide, glycolaldehyde and acetic acid.

In the case of cyanomethanimine, whose *E*-C- isomer was the only one detected in ISM, a gas-phase formation path starting from CN and methanimine as precursor was confirmed to be viable at the very low temperature and pressure conditions of ISM. However, this reaction should lead to all three isomers, *Z*-C- in a slightly larger amount that *E*-C- and N only as traces. Considering that the tool used to detect this compound was microwave spectroscopy, and that this technique is highly dependent on the dipole moment of the studied compound (the line strength is proportional to the square of the dipole moment value), this parameter was verified for cyanomethanimine isomers. It turned out that even if the *Z*-C- isomer should have been formed in a larger amount than its *E*-C- isomer, the low value of its dipole moment could have kept it from being detected by microwave spectroscopy. As far as formamide is concerned, two different gas-phase formation paths were proposed. The first one, assuming OH and methanimine as precursors, was found feasible in ISM but leads preferentially to other compounds of prebiotic interest, such as $H_2CN$ and methanimidic acid isomers. On the other hand, when using $NH_2$ and formaldehyde as precursors, the obtained path is both possible in ISM and efficient in forming formamide. In order to further validate this path, its mono-deuteration was studied and gave results in accordance with observation. Moreover, new models and observations of abundances and locations tend to confirm it. Finally, glycolaldehyde and acetic acid were considered. $O(^3P)$ and two types of hydroxyethyl radicals ($^{\bullet}CH_2CH_2OH$ for glycolaldehyde and $CH_3{}^{\bullet}CHOH$ for

acetic acid) were used and the resulting paths were found to be feasible in ISM. The rate constants obtained from these paths showed that even though glycolaldehyde and acetic acid are not the primary products, their formations are not negligible. Moreover, when the rate constants concerning glycolaldehyde are included in an astrochemical model, the resulting abundances match the observed ones.

Thanks to current space missions, such as New Horizons or OSIRIS-REx for example, together with different ongoing projects, like SOLIS or PRIMOS survey, new COMs are expected to be detected in the upcoming years. ALMA and NOEMA telescopes are also important allies to reach this goal. Moreover, it can be assumed that some new detection tools will be developed, which could lead to a more accurate knowledge of the compositions of some spots of space. In addition to the already detected COMs whose formation in ISM is not yet known, the ones which will undoubtedly be detected in the next years will provide an extensive field of research. Accordingly, collaboration between astronomers, experimental and theoretical chemists, as well as surface and gas-phase chemists, will be needed to try to understand where these rather complex molecules come from and, who knows, the origin of life on Earth.

# *References*

(1) Sarafian, A. R.; Nielsen, S. G.; Marschall, H. R.; McCubbin, F. M.; Monteleone, B. D. Early Accretion of Water in the Inner Solar System from a Carbonaceous Chondrite-like Source. *Science,* **2014**, *346* (6209), 623–626.

(2) Altwegg, K.; Balsiger, H.; Bar-Nun, A.; Berthelier, J. J.; Bieler, A.; Bochsler, P.; Briois, C.; Calmonte, U.; Combi, M.; De Keyser, J.; Eberhardt, P.; Fiethe, B.; Fuselier, S.; Gasc, S.; Gombosi, T. I.; Hansen, K. C.; Hässig, M.; Jäckel, A.; Kopp, E.; Korth, A.; LeRoy, L.; Mall, U.; Marty, B.; Mousis, O.; Neefs, E.; Owen, T.; Rème, H.; Rubin, M.; Sémon, T.; Tzou, C.-Y.; Waite, H.; Wurz, P. 67P/Churyumov-Gerasimenko, a Jupiter Family Comet with a High D/H Ratio. *Science,* **2015**, *347* (6220).

(3) Schmandt, B.; Jacobsen, S. D.; Becker, T. W.; Liu, Z.; Dueker, K. G. Dehydration Melting at the Top of the Lower Mantle. *Science,* **2014**, *344* (6189), 1265–1268.

(4) Genda, H.; Ikoma, M. Origin of the Ocean on the Earth: Early Evolution of Water D/H in a Hydrogen-Rich Atmosphere. *Icarus* **2008**, *194*, 42–52.

(5) Kissel, J.; Brownlee, D. E.; Buchler, K.; Clark, B. C.; Fechtig, H.; Grun, E.; Hornung, K.; Igenbergs, E. B.; Jessberger, E. K.; Krueger, F. R.; Kuczera, H.; McDonnell, J. A. M.; Morfill, G. M.; Rahe, J.; Schwehm, G. H.; Sekanina, Z.; Utterback, N. G.; volk, H. J.; Zook, H. A. Composition of Comet Halley Dust Particles from Giotto Observations. *Nature* **1986**, *321* (6067s), 336–337.

(6) Kissel, J.; Sagdeev, R. Z.; Bertaux, J. L.; Angarov, V. N.; Audouze, J.; Blamont, J. E.; Buchler, K.; Evlanov, E. N.; Fechtig, H.; Fomenkova, M. N.; von Hoerner, H.; Inogamov, N. A.; Khromov, V. N.; Knabe, W.; Krueger, F. R.; Langevin, Y.; Leonas, V. B.; Levasseur-Regourd, A. C.; Managadze, G. G.; Podkolzin, S. N.; Shapiro, V. D.; Tabaldyev, S. R.; Zubkov, B. V. Composition of Comet Halley Dust Particles from Vega Observations. *Nature* **1986**, *321* (6067s), 280–282.

(7) Fomenkova, M. N.; Chang, S.; Mukhin, L. M. Carbonaceous Components in the Comet Halley Dust. *Geochim. Cosmochim. Acta* **1994**, *58* (20), 4503–4512.

(8) Mumma, M. J.; Charnley, S. B. The Chemical Composition of Comets—Emerging Taxonomies and Natal Heritage. *Annu. Rev. Astron. Astrophys.* **2011**, *49* (1), 471–524.

(9) Fray, N.; Bardyn, A.; Cottin, H.; Altwegg, K.; Baklouti, D.; Briois, C.; Colangeli, L.;

Engrand, C.; Fischer, H.; Glasmachers, A.; Grün, E.; Haerendel, G.; Henkel, H.; Höfner, H.; Hornung, K.; Jessberger, E. K.; Koch, A.; Krüger, H.; Langevin, Y.; Lehto, H.; Lehto, K.; Le Roy, L.; Merouane, S.; Modica, P.; Orthous-Daunay, F.-R.; Paquette, J.; Raulin, F.; Rynö, J.; Schulz, R.; Silén, J.; Siljeström, S.; Steiger, W.; Stenzel, O.; Stephan, T.; Thirkell, L.; Thomas, R.; Torkar, K.; Varmuza, K.; Wanczek, K.-P.; Zaprudin, B.; Kissel, J.; Hilchenbach, M. High-Molecular-Weight Organic Matter in the Particles of Comet 67P/Churyumov–Gerasimenko. *Nature* **2016**, *advance online publication*.

(10) Halfen, D. T.; Ilyushin, V.; Ziurys, L. M. Formation of Peptide Bonds in Space: A Comprehensive Study of Formamide and Acetamide in Sgr B2(N). *Astrophys. J.* **2011**, *743* (2001), 60.

(11) Pino, S.; Sponer, J.; Costanzo, G.; Saladino, R.; Mauro, E. From Formamide to RNA, the Path Is Tenuous but Continuous. *Life* **2015**, *5*, 372–384.

(12) Saladino, R.; Crestini, C.; Pino, S.; Costanzo, G.; Di Mauro, E. Formamide and the Origin of Life. *Phys. Life Rev.* **2012**, *9* (1), 84–104.

(13) Jørgensen, J. K.; Favre, C.; Bisschop, S. E.; Bourke, T. L.; van Dishoeck, E. F.; Schmalzl, M. Detection of the Simplest Sugar, Glycolaldehyde, in a Solar-Type Protostar with ALMA. *Astrophys. J. Lett.* **2012**, *757* (1), L4.

(14) Altwegg, K.; Balsiger, H.; Bar-Nun, A.; Berthelier, J.-J.; Bieler, A.; Bochsler, P.; Briois, C.; Calmonte, U.; Combi, M. R.; Cottin, H.; De Keyser, J.; Dhooghe, F.; Fiethe, B.; Fuselier, S. A.; Gasc, S.; Gombosi, T. I.; Hansen, K. C.; Haessig, M.; Jäckel, A.; Kopp, E.; Korth, A.; Le Roy, L.; Mall, U.; Marty, B.; Mousis, O.; Owen, T.; Rème, H.; Rubin, M.; Sémon, T.; Tzou, C.-Y.; Hunter Waite, J.; Wurz, P. Prebiotic Chemicals -Amino Acid and Phosphorus- in the Coma of Comet 67P/Churyumov-Gerasimenko. *Sci. Adv.* **2016**, *2* (5).

(15) Hudson, R. L.; Moore, M. H. Laboratory Studies of the Formation of Methanol and Other Organic Molecules by Water+Carbon Monoxide Radiolysis: Relevance to Comets, Icy Satellites, and Interstellar Ices. *Icarus* **1999**, *140* (2), 451–461.

(16) Wada, A.; Mochizuki, N.; Hiraoka, K. Methanol Formation from Electron-Irradiated Mixed $H_2O/CH_4$ Ice at 10 K. *Astrophys. J.* **2006**, *644* (1), 300.

(17) Whittet, D. C. B.; Cook, A. M.; Herbst, E.; Chiar, J. E.; Shenoy, S. S. Observational Constraints on Methanol Production in Interstellar and Preplanetary Ices. *Astrophys. J.*

**2011**, *742* (1), 28.

(18) Vasyunin, A. I.; Herbst, E. Reactive Desorption and Radiative Association as Possible Drivers of Complex Molecule Formation in the Cold Interstellar Medium. *Astrophys. J.* **2013**, *769* (1), 34.

(19) Ceperley, D. M.; Alder, B. J. Ground State of the Electron Gas by a Stochastic Method. *Phys. Rev. Lett.* **1980**, *45* (7), 566–569.

(20) Becke, A. D. A New Mixing of Hartree–Fock and Local Density-functional Theories. *J. Chem. Phys.* **1993**, *98* (2), 1372–1377.

(21) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. R.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr. J. A. Peralta, J. R.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V; Cioslowski, J.; Fox, D. J. GDVH37p, **2014**.

(22) 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* **1988**, *37* (2), 785–789.

(23) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652.

(24) Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation. *J. Chem. Phys.* **2006**, *124* (3), 34108.

(25) Biczysko, M.; Panek, P.; Scalmani, G.; Bloino, J.; Barone, V. Harmonic and Anharmonic Vibrational Frequency Calculations with the Double-Hybrid B2PLYP Method: Analytic Second Derivatives and Benchmark Studies. *J. Chem. Theory Comput.* **2010**, *6*, 2115–2125.

(26) Papajak, E.; Leverentz, H. R.; Zheng, J.; Truhlar, D. G. Efficient Diffuse Basis Sets: Cc-pVxZ+ and Maug-Cc-pVxZ. *J. Chem. Theory Comput.* **2009**, *5* (5), 1197–1202.

(27) 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* (2), 1007.

(28) 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* (15), 154104.

(29) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32* (7), 1456–1465.

(30) Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *33* (12), 8822–8824.

(31) Perdew, J. P.; Burke, K.; Wang, Y. Generalized Gradient Approximation for the Exchange-Correlation Hole of a Many-Electron System. *Phys. Rev. B* **1996**, *54* (23), 16533–16539.

(32) Dunning Jr., T. H.; Hay, P. J. Gaussian Basis Sets for Molecular Calculations. In *Methods of Electronic Structure Theory SE - 1*; Schaefer III, H., Ed.; Modern Theoretical Chemistry; Springer US, **1977**; Vol. 3, pp 1–27.

(33) Wadt, W. R.; Hay, P. J. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for Main Group Elements Na to Bi. *J. Chem. Phys.* **1985**, *82* (1), 284–298.

(34) Latouche, C.; Baiardi, A.; Barone, V. Virtual Eyes Designed for Quantitative Spectroscopy of Inorganic Complexes: Vibronic Signatures in the Phosphorescence Spectra of Terpyridine Derivatives. *J. Phys. Chem. B* **2015**, *119* (24), 7253–7257.

(35) Baiardi, A.; Latouche, C.; Bloino, J.; Barone, V. Accurate yet Feasible Computations of Resonance Raman Spectra for Metal Complexes in Solution: $[Ru(bpy)_3]^{2+}$ as a Case Study. *Dalt. Trans.* **2014**, *43* (47), 17610–17614.

(36) Barone, V. Anharmonic Vibrational Properties by a Fully Automated Second-Order Perturbative Approach. *J. Chem. Phys.* **2005**, *122*, 14108.

(37) Bloino, J.; Biczysko, M.; Barone, V. General Perturbative Approach for Spectroscopy, Thermodynamics, and Kinetics: Methodological Background and Benchmark Studies. *J. Chem. Theory Comput.* **2012**, *8* (3), 1015–1036.

(38) Barone, V.; Biczysko, M.; Bloino, J. Fully Anharmonic IR and Raman Spectra of Medium-Size Molecular Systems: Accuracy and Interpretation. *Phys. Chem. Chem. Phys.* **2014**, *16* (5), 1759–1787.

(39) Bloino, J.; Biczysko, M.; Santoro, F.; Barone, V. General Approach to Compute Vibrationally Resolved One-Photon Electronic Spectra. *J. Chem. Theory Comput.* **2010**, *6* (4), 1256–1274.

(40) Dennington, R.; Keith, T.; Millam, J. GaussView, Version 5, **2009**.

(41) Skripnikov, L. A Computer Program to Analyse and Visualise Quantum-Chemical Calculations, **2012**.

(42) Barone, V.; Cossi, M.; Tomasi, J. A New Definition of Cavities for the Computation of Solvation Free Energies by the Polarizable Continuum Model. *J. Chem. Phys.* **1997**, *107* (8), 3210.

(43) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. New Developments in the Polarizable Continuum Model for Quantum Mechanical and Classical Calculations on Molecules in Solution. *J. Chem. Phys.* **2002**, *117* (1), 43.

(44) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105* (8), 2999–3094.

(45) Licari, D.; Baiardi, A.; Biczysko, M.; Egidi, F.; Latouche, C.; Barone, V. Implementation of a Graphical User Interface for the Virtual Multifrequency Spectrometer: The VMS-Draw Tool. *J. Comput. Chem.* **2014**, *36* (5), 321-334.

(46) Osman, O. I. Experimental and Theoretical Investigation of the Pyrolysis Products of Iminodiacetonitrile, $(N{\equiv}CCH_2)_2NH$. *J. Phys. Chem. A* **2014**, *118* (46), 10934–10943.

(47) Morin, M. S. T.; Lu, Y.; Black, D. A.; Arndtsen, B. A. Copper-Catalyzed Petasis-Type Reaction: A General Route to α-Substituted Amides From Imines, Acid Chlorides, and Organoboron Reagents. *J. Org. Chem.* **2012**, *77* (4), 2013–2017.

(48) Zhang, E.; Tian, H.; Xu, S.; Yu, X.; Xu, Q. Iron-Catalyzed Direct Synthesis of Imines from Amines or Alcohols and Amines via Aerobic Oxidative Reactions under Air. *Org. Lett.* **2013**, *15* (11), 2704–2707.

(49) Morin, M. S. T.; St-Cyr, D. J.; Arndtsen, B. A.; Krenske, E. H.; Houk, K. N. Modular Mesoionics: Understanding and Controlling Regioselectivity in 1,3-Dipolar

Cycloadditions of Münchnone Derivatives. *J. Am. Chem. Soc.* **2013**, *135* (46), 17349–17358.

(50) Ferris, J. P.; Hagan Jr, W. J. HCN and Chemical Evolution: The Possible Role of Cyano Compounds in Prebiotic Synthesis. *Tetrahedron* **1984**, *40* (7), 1093–1120.

(51) Snyder, L. E.; Buhl, D. Observations of Radio Emission from Interstellar Hydrogen Cyanide. *Astrophys. J.* **1971**, *163*, L47.

(52) Balucani, N. Elementary Reactions of N Atoms with Hydrocarbons: First Steps towards the Formation of Prebiotic N-Containing Molecules in Planetary Atmospheres. *Chem. Soc. Rev.* **2012**, *41* (16), 5473–5483.

(53) Balucani, N. Elementary Reactions and Their Role in Gas-Phase Prebiotic Chemistry. *Int. J. Mol. Sci.* **2009**, *10* (5), 2304–2335.

(54) Boulanger, E.; Anoop, A.; Nachtigallova, D.; Thiel, W.; Barbatti, M. Photochemical Steps in the Prebiotic Synthesis of Purine Precursors from HCN. *Angew. Chemie Int. Ed.* **2013**, *52* (31), 8000–8003.

(55) Webster, O. W. Cyanocarbons: A Classic Example of Discovery-Driven Research. *J. Polym. Sci. Part A Polym. Chem.* **2002**, *40* (2), 210–221.

(56) Miller, J. S.; Epstein, A. J. Organic and Organometallic Molecular Magnetic Materials-Designer Magnets. *Angew. Chemie Int. Ed. English* **1994**, *33* (4), 385–415.

(57) Her, J.-H.; Stephens, P. W.; Davidson, R. A.; Min, K. S.; Bagnato, J. D.; van Schooten, K.; Boehme, C.; Miller, J. S. Weak Ferromagnetic Ordering of the $Li^{+}[TCNE]^{\bullet-}$ (TCNE = Tetracyanoethylene) Organic Magnet with an Interpenetrating Diamondoid Structure. *J. Am. Chem. Soc.* **2013**, *135* (48), 18060–18063.

(58) Tucker, M. J.; Getahun, Z.; Nanda, V.; DeGrado, W. F.; Gai, F. A New Method for Determining the Local Environment and Orientation of Individual Side Chains of Membrane-Binding Peptides. *J. Am. Chem. Soc.* **2004**, *126* (16), 5078–5079.

(59) Lindquist, B. A.; Furse, K. E.; Corcelli, S. A. Nitrile Groups as Vibrational Probes of Biomolecular Structure and Dynamics: An Overview. *Phys. Chem. Chem. Phys.* **2009**, *11* (37), 8119–8132.

(60) Maienschein-Cline, M. G.; Londergan, C. H. The CN Stretching Band of Aliphatic Thiocyanate Is Sensitive to Solvent Dynamics and Specific Solvation. *J. Phys. Chem. A*

**2007**, *111* (40), 10020–10025.

(61) Barone, V.; Baiardi, A.; Biczysko, M.; Bloino, J.; Cappelli, C.; Lipparini, F. Implementation and Validation of a Multi-Purpose Virtual Spectrometer for Large Systems in Complex Environments. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12404–12422.

(62) Latouche, C.; Barone, V. Computational Chemistry Meets Experiments for Explaining the Behavior of Bibenzyl: A Thermochemical and Spectroscopic (IR, Raman and NMR) Investigation. *J. Chem. Theory Comput.* **2014**, *10* (12), 5586–5592.

(63) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98* (45), 11623–11627.

(64) Andersson, M. P.; Uvdal, P. New Scale Factors for Harmonic Vibrational Frequencies Using the B3LYP Density Functional Method with the Triple-ζ Basis Set 6-311+G(d,p). *J. Phys. Chem. A* **2005**, *109* (12), 2937–2941.

(65) Merrick, J. P.; Moran, D.; Radom, L. An Evaluation of Harmonic Vibrational Frequency Scale Factors. *J. Phys. Chem. A* **2007**, *111* (45), 11683–11700.

(66) Bloino, J.; Barone, V. A Second-Order Perturbation Theory Route to Vibrational Averages and Transition Properties of Molecules: General Formulation and Application to Infrared and Vibrational Circular Dichroism Spectroscopies. *J. Chem. Phys.* **2012**, *136*, 124108.

(67) Piccardo, M.; Bloino, J.; Barone, V. Generalized Vibrational Perturbation Theory for Rotovibrational Energies of Linear, Symmetric and Asymmetric Tops: Theory, Approximations, and Automated Approaches to Deal with Medium-to-Large Molecular Systems. *Int. J. Quantum Chem.* **2015**, *115* (15), 948–982.

(68) Carnimeo, I.; Puzzarini, C.; Tasinato, N.; Stoppa, P.; Charmet, A. P.; Biczysko, M.; Cappelli, C.; Barone, V. Anharmonic Theoretical Simulations of Infrared Spectra of Halogenated Organic Compounds. *J. Chem. Phys.* **2013**, *139* (7).

(69) Latouche, C.; Palazzetti, F.; Skouteris, D.; Barone, V. High Accuracy Vibrational Computations for Transition Metal Complexes Including Anharmonic Corrections: Ferrocene, Ruthenocene and Osmocene as Test Cases. *J. Chem. Theory Comput.* **2014**, *10* (10), 4565–4573.

(70) Vazart, F.; Calderini, D.; Skouteris, D.; Latouche, C.; Barone, V. Re-Assessment of the Thermodynamic, Kinetic, and Spectroscopic Features of Cyanomethanimine Derivatives: A Full Anharmonic Perturbative Treatment. *J. Chem. Theory Comput.* **2015**.

(71) *IR-Cyclobutanecarbonitrile, visited on 22th of June 2015*, **2015**, http://webbook.nist.gov/cgi/cbook.cgi?ID=C75127{&}Ty.

(72) *IR-2-propenenitrile, visited on 22th of June 2015*, **2015**, http://webbook.nist.gov/cgi/cbook.cgi?ID=C75127{&}Ty.

(73) *IR-Benzonitrile, visited on 22th of June 2015*, **2015**, http://webbook.nist.gov/cgi/cbook.cgi?ID=C75127{&}Ty.

(74) Binev, I. Y.; Binev, I. G. Possibilities of Detection and Determination of Propanedinitrile (Malononitrile) as a Carbanion by Means of IR Spectroscopy. *Comptes Rendus l'Academie Bulg. des Sci.* **2000**, *53* (7), 7:67.

(75) *IR-Buntanedinitrile, visited on 22th of June 2015*, **2015**, http://webbook.nist.gov/cgi/cbook.cgi?ID=C75127{&}.

(76) Long, D. A.; George, W. O. Spectroscopic Studies Ot Compounds Containing the -CN group. *Spectrochim. Acta* **1963**, *19* (11), 1717–1729.

(77) D'Andrade, B. W.; Forrest, S. R. White Organic Light-Emitting Devices for Solid-State Lighting. *Adv. Mater.* **2004**, *16* (18), 1585–1595.

(78) Sun, Y.; Forrest, S. R. High-Efficiency White Organic Light Emitting Devices with Three Separate Phosphorescent Emission Layers. *Appl. Phys. Lett.* **2007**, *91* (26), 263503.

(79) Chen, C.-Y.; Pootrakulchote, N.; Chen, M.-Y.; Moehl, T.; Tsai, H.-H.; Zakeeruddin, S. M.; Wu, C.-G.; Grätzel, M. A New Heteroleptic Ruthenium Sensitizer for Transparent Dye-Sensitized Solar Cells. *Adv. Energy Mater.* **2012**, *2* (12), 1503–1509.

(80) Latouche, C.; Lanoë, P.-H.; Williams, J. a. G.; Guerchais, V.; Boucekkine, A.; Fillaut, J.-L. Switching of Excited States in Cyclometalated Platinum Complexes Incorporating Pyridyl-Acetylide Ligands (Pt–CC–py): A Combined Experimental and Theoretical Study. *New. J. Chem.* **2011**, *35* (10), 2196–2202.

(81) Fillaut, J.; Akdas-kilig, H.; Dean, E.; Latouche, C.; Boucekkine, A. Switching of Reverse Charge Transfers for a Rational Design of an OFF–ON Phosphorescent

Chemodosimeter of Cyanide Anions. *Inorg. Chem.* **2013**, *52*, 4890–4897.

(82) Savel, P.; Latouche, C.; Roisnel, T.; Akdas-Kilig, H.; Boucekkine, A.; Fillaut, J.-L. Cyclometalated platinum(II) with Ethynyl-Linked Azobenzene Ligands: An Original Switching Mode. *Dalton Trans.* **2013**, *42* (48), 16773–16783.

(83) Okada, S.; Okinaka, K.; Iwawaki, H.; Furugori, M.; Hashimoto, M.; Mukaide, T.; Kamatani, J.; Igawa, S.; Tsuboyama, A.; Takiguchi, T.; Ueno, K. Substituent Effects of Iridium Complexes for Highly Efficient Red OLEDs. *Dalt. Trans.* **2005**, No. 9, 1583–1590.

(84) Lamansky, S.; Djurovich, P.; Murphy, D.; Abdel-Razzaq, F.; Kwong, R.; Tsyba, I.; Bortz, M.; Mui, B.; Bau, R.; Thompson, M. E. Synthesis and Characterization of Phosphorescent Cyclometalated Iridium Complexes. *Inorg. Chem.* **2001**, *40* (7), 1704–1711.

(85) Lepeltier, M.; Le Bozec, H.; Guerchais, V.; Lee, T. K.-M.; Lo, K. K.-W. Tris-Cyclometalated Iridium(III) Styryl Complexes and Their Saturated Analogues: Direct Functionalization of Ir(4-Me-ppy)$_3$ and Hydrogen Transfer Process. *Organometallics* **2005**, *24* (24), 6069–6072.

(86) Wu, S.-H.; Burkhardt, S. E.; Yao, J.; Zhong, Y.-W.; Abruña, H. D. Near-Infrared Absorbing and Emitting RuII−Pt(II) Heterodimetallic Complexes of Dpdpz (Dpdpz = 2,3-Di(2-Pyridyl)-5,6-Diphenylpyrazine). *Inorg. Chem.* **2011**, *50* (9), 3959–3969.

(87) Culham, S.; Lanoë, P.-H.; Whittle, V. L.; Durrant, M. C.; Williams, J. A. G.; Kozhevnikov, V. N. Highly Luminescent Dinuclear Platinum(II) Complexes Incorporating Bis-Cyclometallating Pyrazine-Based Ligands: A Versatile Approach to Efficient Red Phosphors. *Inorg. Chem.* **2013**, *52* (19), 10992–11003.

(88) Green, K.; Gauthier, N.; Sahnoune, H.; Halet, J.-F.; Paul, F.; Fabre, B. Covalent Immobilization of Redox-Active Fe($\kappa^2$-dppe)($\eta^5$-$C_5Me_5$)-Based π-Conjugated Wires on Oxide-Free Hydrogen-Terminated Silicon Surfaces. *Organometallics* **2013**, *32* (19), 5333–5342.

(89) Sahnoune, H.; Baranová, Z.; Bhuvanesh, N.; Gladysz, J. A.; Halet, J.-F. A Metal-Capped Conjugated Polyyne Threaded through a Phenanthroline-Based Macrocycle. Probing beyond the Mechanical Bond to Interactions in Interlocked Molecular Architectures. *Organometallics* **2013**, *32* (21), 6360–6367.

(90) Makhoul, R.; Sahnoune, H.; Davin, T.; Kahlal, S.; Dorcet, V.; Roisnel, T.; Halet, J.-F.; Hamon, J.-R.; Lapinte, C. Proton-Controlled Regioselective Synthesis of [Cp*(dppe)Fe–C≡C-1-($\eta^6$-$C_{10}H_7$)Ru($\eta^5$-Cp](PF$_6$) and Electron-Driven Haptotropic Rearrangement of the ($\eta^5$-Cp)Ru$^+$ Arenophile. *Organometallics* **2014**, *33* (18), 4792–4802.

(91) Green, K.; Gauthier, N.; Sahnoune, H.; Argouarch, G.; Toupet, L.; Costuas, K.; Bondon, A.; Fabre, B.; Halet, J.-F.; Paul, F. Synthesis and Characterization of Redox-Active Mononuclear Fe($\kappa^2$-dppe)($\eta^5$-$C_5Me_5$)-Terminated π-Conjugated Wires. *Organometallics* **2013**, *32* (15), 4366–4381.

(92) Latouche, C.; Liu, C. W.; Saillard, J.-Y. Encapsulating Hydrides and Main-Group Anions in d$^{10}$-Metal Clusters Stabilized by 1,1-Dichalcogeno Ligands. *J. Clust. Sci.* **2014**, *25* (1), 147–171.

(93) Liao, J.-H.; Latouche, C.; Li, B.; Kahlal, S.; Saillard, J.-Y.; Liu, C. W. A Twelve-Coordinated Iodide in a Cuboctahedral silver(I) Skeleton. *Inorg. Chem.* **2014**, *53* (4), 2260–2267.

(94) Latouche, C.; Lin, Y.-R.; Tobon, Y.; Furet, E.; Saillard, J.-Y.; Liu, C.-W.; Boucekkine, A. Au-Au Chemical Bonding Induced by UV Irradiation of Dinuclear Gold(I) Complexes: A Computational Study with Experimental Evidence. *Phys. Chem. Chem. Phys.* **2014**, *16* (47), 25840–25845.

(95) Schulze, M.; Steffen, A.; Würthner, F. Near-IR Phosphorescent Ruthenium(II) and Iridium(III) Perylene Bisimide Metal Complexes. *Angew. Chemie Int. Ed.* **2015**, *54* (5), 1570–1573.

(96) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. A Direct MP2 Gradient Method. *Chem. Phys. Lett.* **1990**, *166* (3), 275–280.

(97) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. Semi-Direct Algorithms for the MP2 Energy and Gradient. *Chem. Phys. Lett.* **1990**, *166* (3), 281–289.

(98) Head-Gordon, M.; Head-Gordon, T. Analytic MP2 Frequencies without Fifth-Order Storage. Theory and Application to Bifurcated Hydrogen Bonds in the Water Hexamer. *Chem. Phys. Lett.* **1994**, *220* (1–2), 122–128.

(99) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 Energy Evaluation by Direct Methods. *Chem. Phys. Lett.* **1988**, *153* (6), 503–506.

(100) Sæbø, S.; Almlöf, J. Avoiding the Integral Storage Bottleneck in LCAO Calculations of

Electron Correlation. *Chem. Phys. Lett.* **1989**, *154* (1), 83–89.

(101)http://compchem.sns.it/sites/default/files/download/gaussian/basis_sets/SNSD.gbs **2013**.

(102) Goerigk, L.; Grimme, S. Efficient and Accurate Double-Hybrid-Meta-GGA Density Functionals-Evaluation with the Extended GMTKN30 Database for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2011**, *7* (2), 291–309.

(103) Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A. A Complete Basis Set Model Chemistry. V. Extensions to Six or More Heavy Atoms. *J. Chem. Phys.* **1996**, *104* (7).

(104) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. A Complete Basis Set Model Chemistry. VII. Use of the Minimum Population Localization Method. *J. Chem. Phys.* **2000**, *112* (15).

(105) Miller, W. H.; Hernandez, R.; Handy, N. C.; Jayatilaka, D.; Willetts, A. Ab Initio Calculation of Anharmonic Constants for a Transition State, with Application to Semiclassical Transition State Tunneling Probabilities. *Chem. Phys. Lett.* **1990**, *172* (1), 62–68.

(106) Daniel P. Zaleski and Nathan A. Seifert and Amanda L. Steber and Matt T. Muckle and Ryan A. Loomis and Joanna F. Corby and Oscar Martinez, J.; Crabtree, K. N.; Jewell, P. R.; Hollis, J. M.; Lovas, F. J.; Vasquez, D.; Nyiramahirwe, J.; Sciortino, N.; Johnson, K.; McCarthy, M. C.; Remijan, A. J.; Pate, B. H. Detection of E-Cyanomethanimine toward Sagittarius B2(N) in the Green Bank Telescope PRIMOS Survey. *Astrophys. J. Lett.* **2013**, *765* (1), L10.

(107) Adrio, J.; Carretero, J. C. Novel Dipolarophiles and Dipoles in the Metal-Catalyzed Enantioselective 1,3-Dipolar Cycloaddition of Azomethine Ylides. *Chem. Commun.* **2011**, *47* (24), 6784–6794.

(108) Erkkilä, A.; Majander, I.; Pihko, P. M. Iminium Catalysis. *Chem. Rev.* **2007**, *107* (12), 5416–5470.

(109) Ma, J.-A. Catalytic Asymmetric Synthesis of α- and β-Amino Phosphonic Acid Derivatives. *Chem. Soc. Rev.* **2006**, *35* (7), 630–636.

(110) Nielsen, M.; Worgull, D.; Zweifel, T.; Gschwend, B.; Bertelsen, S.; Jorgensen, K. A.

Mechanisms in Aminocatalysis. *Chem. Commun.* **2011**, *47* (2), 632–649.

(111) Balucani, N.; Bergeat, A.; Cartechini, L.; Volpi, G. G.; Casavecchia, P.; Skouteris, D.; Rosi, M. Combined Crossed Molecular Beam and Theoretical Studies of the N($^2$D) + $CH_4$ Reaction and Implications for Atmospheric Models of Titan. *J. Phys. Chem. A* **2009**, *113* (42), 11138–11152.

(112) Balucani, N.; Leonori, F.; Petrucci, R.; Stazi, M.; Skouteris, D.; Rosi, M.; Casavecchia, P. Formation of Nitriles and Imines in the Atmosphere of Titan: Combined Crossed-Beam and Theoretical Studies on the Reaction Dynamics of Excited Nitrogen Atoms N($^2$D) with Ethane. *Faraday Discuss.* **2010**, *147* (0), 189–216.

(113) Puzzarini, C.; Biczysko, M.; Barone, V. Accurate Harmonic/Anharmonic Vibrational Frequencies for Open-Shell Systems: Performances of the B3LYP/N07D Model for Semirigid Free Radicals Benchmarked by CCSD(T) Computations. *J. Chem. Theory Comput.* **2010**, *6* (3), 828–838.

(114) CN Stretching from HCN, Last Time Viewed: December 12th 2014.

(115) Wang, F.; Landau, D. Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States. *Phys. Rev. Lett.* **2001**, *86* (10), 2050–2053.

(116) Basire, M.; Parneix, P.; Calvo, F. Quantum Anharmonic Densities of States Using the Wang-Landau Method. *J. Chem. Phys.* **2008**, *129* (8), 81101.

(117) Nguyen, T. L.; Barker, J. R. Sums and Densities of Fully Coupled Anharmonic Vibrational States: A Comparison of Three Practical Methods. *J. Phys. Chem. A* **2010**, *114* (10), 3718–3730.

(118) Evans, R. A.; Lorencak, P.; Ha, T. K.; Wentrup, C. HCN Dimers: Iminoacetonitrile and N-Cyanomethanimine. *J. Am. Chem. Soc.* **1991**, *113* (19), 7261–7276.

(119) Stein, S. E.; Rabinovitch, B. S. Accurate Evaluation of Internal Energy Level Sums and Densities Including Anharmonic Oscillators and Hindered Rotors. *J. Chem. Phys.* **1973**, *58* (6).

(120) Troe, J.; Ushakov, V. G. Anharmonic Rovibrational Numbers and Densities of States for $H_2O$, $H_2CO$, and $H_2O_2$. *J. Phys. Chem. A* **2009**, *113* (16), 3940–3945.

(121) Puzzarini, C.; Ali, A.; Biczysko, M.; Barone, V. Accurate Spectroscopic Characterization of Protonated Oxirane: A Potential Prebiotic Species in Titan's Atmosphere. *Astrophys.*

*J.* **2014**, *792* (2), 118.

(122) Puzzarini, C.; Biczysko, M.; Bloino, J.; Barone, V. Accurate Spectroscopic Characterization of Oxirane: A Valuable Route To Its Identification In Titan's Atmosphere And The Assignment Of Unidentified Infrared Bands. *Astrophys. J.* **2014**, *785* (2), 107.

(123) Chakrabarti, S.; Chakrabarti, S. K. Can DNA Bases Be Produced during Molecular Cloud Collapse? *A&A* **2000**, *354*, L6–L8.

(124) Smith, I. W. M.; Talbi, D.; Herbst, E. The Production of HCN Dimer and More Complex Oligomers in Dense Interstellar Clouds. *Astron. Astrophys.* **2001**, *369* (2), 611–615.

(125) Yim, M. K.; Choe, J. C. Dimerization of HCN in the Gas Phase: A Theoretical Mechanistic Study. *Chem. Phys. Lett.* **2012**, *538*, 24–28.

(126) Jung, S. H.; Choe, J. C. Mechanisms of Prebiotic Adenine Synthesis from HCN by Oligomerization in the Gas Phase. *Astrobiology* **2013**, *13* (5), 465–475.

(127) Basiuk, V. A.; Bogillo, V. I. Theoretical Study of Amino Acid Precursor Formation in the Interstellar Medium. 2. Reaction of Methylenimine with CN Radical. *Adv. Sp. Res.* **2002**, *30* (6), 1445–1450.

(128) Bada, J. L.; Cleaves, H. J. Ab Initio Simulations and the Miller Prebiotic Synthesis Experiment. *Proc. Natl. Acad. Sci.* **2015**, *112* (4), E342.

(129) Jeilani, Y. A.; Nguyen, H. T.; Cardelino, B. H.; Nguyen, M. T. Free Radical Pathways for the Prebiotic Formation of Xanthine and Isoguanine from Formamide. *Chem. Phys. Lett.* **2014**, *598* (0), 58–64.

(130) Saladino, R.; Botta, G.; Pino, S.; Costanzo, G.; Di Mauro, E. Genetics First or Metabolism First? The Formamide Clue. *Chem. Soc. Rev.* **2012**, *41* (16), 5526.

(131) Barone, V.; Latouche, C.; Skouteris, D.; Vazart, F.; Balucani, N.; Ceccarelli, C.; Lefloch, B. Gas Phase Formation of the Prebiotic Molecule Formamide: Insights from New Quantum Computations. *Mon. Not. R. Astron. Soc.* **2015**, *453*, L31–L35.

(132) Gottlieb, C. A.; Palmer, P.; Rickard, L. J.; Zuckerman, B. Studies of Interstellar Formamide. *Astrophys. J.* **1973**, *182*, 699.

(133) Motiyenko, R. A.; Tercero, B.; Cernicharo, J.; Margulès, L. Rotational Spectrum of Formamide up to 1 THz and First ISM Detection of Its $\nu_{12}$ Vibrational State. *A&A* **2012**,

*548*, A71.

(134) Kahane, C.; Ceccarelli, C.; Faure, a.; Caux, E. Detection of Formamide, the Simplest But Crucial Amide, in a Solar-Type Protostar. *Astrophys. J.* **2013**, *763*, L38.

(135) Redondo, P.; Barrientos, C.; Largo, A. Some Insights Into Formamide Formation Through Gas-Phase Reactions in the Interstellar Medium. *Astrophys. J.* **2014**, *780* (2011), 181.

(136) Dawley, M. M.; Pirim, C.; Orlando, T. M. Thermal Processing of Formamide Ices on Silicate Grain Analogue. *J. Phys. Chem. A* **2014**, *118* (7), 1220–1227.

(137) Dawley, M. M.; Pirim, C.; Orlando, T. M. Radiation Processing of Formamide and Formamide: Water Ices on Silicate Grain Analogue. *J. Phys. Chem. A* **2014**, *118* (7), 1228–1236.

(138) Jones, B. M.; Bennett, C. J.; Kaiser, R. I. Mechanistical Studies on the Production of Formamide ($H_2NCHO$) within Interstellar Ice Analogs. *Astrophys. J.* **2011**, *734* (2), 78.

(139) Kaňuchová, Z.; Urso, R. G.; Baratta, G. A.; Brucato, J. R.; Palumbo, M. E.; Strazzulla, G. Synthesis of Formamide and Isocyanic Acid after Ion Irradiation of Frozen Gas Mixtures. *A&A* **2016**, *585*, A155.

(140) Ali, M. A.; Barker, J. R. Comparison of Three Isoelectronic Multiple-Well Reaction Systems: OH + $CH_2O$, OH + $CH_2CH_2$, and OH + $CH_2NH$. *J. Phys. Chem. A* **2015**, *119* (28), 7578–7592.

(141) 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* (6), 479–483.

(142) Watts, J. D.; Gauss, J.; Bartlett, R. J. Coupled-cluster Methods with Noniterative Triple Excitations for Restricted Open-shell Hartree–Fock and Other General Single Determinant Reference Functions. Energies and Analytical Gradients. *J. Chem. Phys.* **1993**, *98* (11).

(143) Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G. CFOUR A Quantum Chemical Program Package. **2011**; with contributions from A. A. Auer, R. J. Bartlett, U. Benedikt, C. Berger, D. E. Bernholdt, Y. J. Bomble, O. Christiansen, M. Heckert, O. Heun, C. Huber, T.-C. Jagau, D. Jonsson, J. Juselius, K. Klein, W. J. Lauderdale, F. Lipparini, D. Matthews,

T. Metzroth, L. A. Mueck, D. P. O'Neill, D. R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, S. Stopkowicz, A. Tajti, J. Vazquez, F. Wang, J. D. Watts and the integral packages MOLECULE (J. Almloef and P. R. Taylor), PROPS (P. R. Taylor), ABACUS (T. Helgaker, H. J. Aa. Jensen, P. Jorgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wuellen. For the current version, see http://www.cfour.de (accessed Oct 4, 2016)..

(144) Heckert, M.; Kállay, M.; Tew, D. P.; Klopper, W.; Gauss, J. Basis-Set Extrapolation Techniques for the Accurate Calculation of Molecular Equilibrium Geometries Using Coupled-Cluster Theory. *J. Chem. Phys.* **2006**, *125* (4).

(145) Woon, D. E.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. V. Core-valence Basis Sets for Boron through Neon. *J. Chem. Phys.* **1995**, *103* (11).

(146) Feller, D. The Use of Systematic Sequences of Wave Functions for Estimating the Complete Basis Set, Full Configuration Interaction Limit in Water. *J. Chem. Phys.* **1993**, *98* (9).

(147) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-Set Convergence of Correlated Calculations on Water. *J. Chem. Phys.* **1997**, *106* (23).

(148) Noga, J.; Bartlett, R. J. The Full CCSDT Model for Molecular Electronic Structure. *J. Chem. Phys.* **1987**, *86* (12).

(149) Scuseria, G. E.; Schaefer, H. F. A New Implementation of the Full CCSDT Model for Molecular Electronic Structure. *Chem. Phys. Lett.* **1988**, *152* (4), 382–386.

(150) Kállay, M.; Surján, P. R. Higher Excitations in Coupled-Cluster Theory. *J. Chem. Phys.* **2001**, *115* (7).

(151) M. Kàllay Z. Rolik, I. L..; Kornis, B. MRCC, a Quantum Chemical Program Suite. See Also Z. Rolik and M. Kàllay, J. Chem. Phys. 135, 104111 (**2011**) as well as www.mrcc.hu.

(152) Aquino, N.; Campoy, G.; Yee-Madeira, H. The Inversion Potential for $NH_3$ Using a DFT Approach. *Chem. Phys. Lett.* **1998**, *296* (1–2), 111–116.

(153) von Roos, O. Position-Dependent Effective Masses in Semiconductor Theory. *Phys. Rev. B* **1983**, *27* (12), 7547–7552.

(154) Bonatsos, D.; Georgoudis, P. E.; Minkov, N.; Petrellis, D.; Quesne, C. Bohr Hamiltonian

with a Deformation-Dependent Mass Term for the Kratzer Potential. *Phys. Rev. C* **2013**, *88* (3), 34316.

(155) Miller, W. H.; Handy, N. C.; Adams, J. E. Reaction Path Hamiltonian for Polyatomic Molecules. *J. Chem. Phys.* **1980**, *72* (1).

(156) Barone, V.; Grand, A.; Minichino, C.; Subra, R. Vibrational Modulation Effects on the Hyperfine Coupling Constants of Fluoromethyl Radicals. *J. Chem. Phys.* **1993**, *99* (9).

(157) Minichino, C.; Barone, V. From Concepts to Algorithms for the Characterization of Reaction Mechanisms. $H_2CS$ as a Case Study. *J. Chem. Phys.* **1994**, *100* (5).

(158) Colbert, D. T.; Miller, W. H. A Novel Discrete Variable Representation for Quantum Mechanical Reactive Scattering via the S-matrix Kohn Method. *J. Chem. Phys.* **1992**, *96* (3).

(159) Leonori, F.; Skouteris, D.; Petrucci, R.; Casavecchia, P.; Rosi, M.; Balucani, N. Combined Crossed Beam and Theoretical Studies of the $C(^1D) + CH_4$ Reaction. *J. Chem. Phys.* **2013**, *138*, 24311.

(160) Leonori, F.; Petrucci, R.; Balucani, N.; Casavecchia, P.; Rosi, M.; Skouteris, D.; Berteloite, C.; Le Picard, S. D.; Canosa, A.; Sims, I. R. Crossed-Beam Dynamics, Low-Temperature Kinetics, and Theoretical Studies of the Reaction $S(^1D)+C_2H_4$. *J. Phys. Chem. A* **2009**, *113*, 15328–15345.

(161) Vazart, F.; Latouche, C.; Skouteris, D.; Balucani, N.; Barone, V. Cyanomethanimine Isomers in Cold Interstellar Clouds: Insights from Electronic Structure and Kinetic Calculations. *Astrophys. J.* **2015**, *810* (2), 111.

(162) Weston, R. E.; Schwartz, H. E. *Chemical Kinetics*; Prentice-Hall Inc.: Englewood Cliffs,NJ.

(163) 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* (10), 2058–2082.

(164) Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. The Performance of Composite Schemes and Hybrid CC/DFT Model in Predicting Structure, Thermodynamic and Spectroscopic Parameters: The Challenge of the Conformational Equilibrium in Glycine. *Phys. Chem. Chem. Phys.* **2013**, *15* (25), 10094–10111.

(165) Puzzarini, C.; Penocchio, E.; Biczysko, M.; Barone, V. Molecular Structure and Spectroscopic Signatures of Acrolein: Theory Meets Experiment. *J. Phys. Chem. A* **2014**, *118* (33), 6648–6656.

(166) Barone, V.; Biczysko, M.; Bloino, J.; Cimino, P.; Penocchio, E.; Puzzarini, C. CC/DFT Route toward Accurate Structures and Spectroscopic Features for Observed and Elusive Conformers of Flexible Molecules: Pyruvic Acid as a Case Study. *J. Chem. Theory Comput.* **2015**, *11* (9), 4342–4363.

(167) Webook.nist.gov. IR-Formamide, visited on 20th of May 2015.

(168) Puzzarini, C. Isomerism of Cyanomethanimine: Accurate Structural, Energetic, and Spectroscopic Characterization. *J. Phys. Chem. A* **2015**, *119* (47), 11614–11622.

(169) 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* (10), 4689–4707.

(170) Vazart, F.; Latouche, C.; Cimino, P.; Barone, V. Accurate Infrared (IR) Spectra for Molecules Containing the C≡N Moiety by Anharmonic Computations with the Double Hybrid B2PLYP Density Functional. *J. Chem. Theory Comput.* **2015**, *11* (9), 4364–4369.

(171) Bunkan, A. J. C.; Tang, Y.; Selleva, S. R.; Nielsen, C. J. Atmospheric Gas Phase Chemistry of $CH_2$=NH and HNC. A First- Principles Approach. *J. Phys. Chem. A*, **2014**, *118* (28), 5279–5288.

(172) Ceccarelli, C.; Caselli, P.; Bockelée-Morvan, D.; Mousis, O.; Pizzarello, S.; Robert, F.; Semenov, D. *Protostars and Planets VI*, University.; Beuther, H., Klessen, R. S., Dullemond, C. P., Henning, T., Eds.; University of Arizona Press, **2014**.

(173) Linsky, J. L. D/H and Nearby Interstellar Cloud Structures. *Space Sci. Rev.* **2007**, *130* (1), 367–375.

(174) Ceccarelli, C.; Castets, A.; Loinard, L.; Caux, E.; Tielens, A. G. G. M. Detection of Doubly Deuterated Formaldehyde towards the Low Luminosity Protostar IRAS 16293-2422. *A&A.* **1998**, *338*, L43–L46.

(175) Parise, B.; Ceccarelli, C.; Tielens, A. G. G. M.; Castets, A.; Caux, E.; Lefloch, B.; Maret, S. Testing Grain Surface Chemistry: A Survey of Deuterated Formaldehyde and Methanol in Low-Mass Class 0 Protostars. *A&A* **2006**, *453* (3), 949–958.

(176) Caselli, P.; Ceccarelli, C. Our Astrochemical Heritage. *Astron. Astrophys. Rev.* **2012**, *20* (1), 56.

(177) Coutens, A.; Jørgensen, J. K.; van der Wiel, M. H. D.; Müller, H. S. P.; Lykke, J. M.; Bjerkeli, P.; Bourke, T. L.; Calcutt, H.; Drozdovskaya, M. N.; Favre, C.; Fayolle, E. C.; Garrod, R. T.; Jacobsen, S. K.; Ligterink, N. F. W.; Öberg, K. I.; Persson, M. V.; van Dishoeck, E. F.; Wampfler, S. F. The ALMA-PILS Survey: First Detections of Deuterated Formamide and Deuterated Isocyanic Acid in the Interstellar Medium. *A&A* **2016**, *590*, L6.

(178) Jørgensen, J. K.; van der Wiel, M. H. D.; Coutens, A.; Lykke, J. M.; Müller, H. S. P.; van Dishoeck, E. F.; Calcutt, H.; Bjerkeli, P.; Bourke, T. L.; Drozdovskaya, M. N.; Favre, C.; Fayolle, E. C.; Garrod, R. T.; Jacobsen, S. K.; Öberg, K. I.; Persson, M. V.; Wampfler, S. F. The ALMA Protostellar Interferometric Line Survey (PILS) - First Results from an Unbiased Submillimeter Wavelength Line Survey of the Class 0 Protostellar Binary IRAS 16293-2422 with ALMA. *A&A* **2016**, *595*, A117.

(179) Jaber Al-Edhari, A.; Ceccarelli, C.; Kahane, C.; Viti, S.; Balucani, N.; Caux, E.; Faure, A.; Lefloch, B.; Lique, F.; Mendoza, E.; Quenard, D.; Wiesenfeld, L. History of the Solar-Type Protostar IRAS 16293–2422 as Told by the Cyanopolyynes. *A&A* **2017**, *597*, A40.

(180) Balucani, N.; Skouteris, D.; Leonori, F.; Petrucci, R.; Hamberg, M.; Geppert, W. D.; Casavecchia, P.; Rosi, M. Combined Crossed Beam and Theoretical Studies of the N($^2$D) + $C_2H_4$ Reaction and Implications for Atmospheric Models of Titan. *J. Phys. Chem. A* **2012**, *116* (43), 10467–10479.

(181) Song, L.; Kastner, J. Formation of the Prebiotic Molecule $NH_2CHO$ on Astronomical Amorphous Solid Water Surfaces: Accurate Tunneling Rate Calculations. *Phys. Chem. Chem. Phys.* **2016**, *18* (42), 29278–29285.

(182) Vazart, F.; Calderini, D.; Puzzarini, C.; Skouteris, D.; Barone, V. State-of-the-Art Thermochemical and Kinetic Computations for Astrochemical Complex Organic Molecules: Formamide Formation in Cold Interstellar Clouds as a Case Study. *J. Chem. Theory Comput.* **2016**, *12* (11), 5385–5397.

(183) Mendoza, E.; Lefloch, B.; López-Sepulcre, A.; Ceccarelli, C.; Codella, C.; Boechat-Roberty, H. M.; Bachiller, R. Molecules with a Peptide Link in Protostellar Shocks: A Comprehensive Study of L1157. *Mon. Not. R. Astron. Soc.* **2014**, *445* (1), 151–161.

(184) Gueth, F.; Guilloteau, S.; Bachiller, R. SiO Shocks in the L1157 Molecular Outflow. *A&A* **1998**, *333*, 287–297.

(185) Podio, L.; Codella, C.; Gueth, F.; Cabrit, S.; Maury, A.; Tabone, B.; Lefèvre, C.; Anderl, S.; André, P.; Belloche, A.; Bontemps, S.; Hennebelle, P.; Lefloch, B.; Maret, S.; Testi, L. First Image of the L1157 Molecular Jet by the CALYPSO IRAM-PdBI Survey. *A&A* **2016**, *593*, L4.

(186) Busquet, G.; Lefloch, B.; Benedettini, M.; Ceccarelli, C.; Codella, C.; Cabrit, S.; Nisini, B.; Viti, S.; Gómez-Ruiz, A. I.; Gusdorf, A.; di Giorgio, A. M.; Wiesenfeld, L. The CHESS Survey of the L1157-B1 Bow-Shock: High and Low Excitation Water Vapor. *A&A* **2014**, *561*, A120.

(187) Lefloch, B.; Cabrit, S.; Busquet, G.; Codella, C.; Ceccarelli, C.; Cernicharo, J.; Pardo, J. R.; Benedettini, M.; Lis, D. C.; Nisini, B. The CHESS Survey of the L1157-B1 Shock Region: CO Spectral Signatures of Jet-Driven Bow Shocks. *Astrophys. J. Lett.* **2012**, *757* (2), L25.

(188) Codella, C.; Benedettini, M.; Beltrán, M. T.; Gueth, F.; Viti, S.; Bachiller, R.; Tafalla, M.; Cabrit, S.; Fuente, A.; Lefloch, B. Methyl Cyanide as Tracer of Bow Shocks in L1157-B1. *A&A* **2009**, *507* (2), L25–L28.

(189) Gueth, F.; Guilloteau, S.; Bachiller, R. A Precessing Jet in the L1157 Molecular Outflow. *A&A* **1996**, *307*, 891.

(190) Bachiller, R.; Pérez Gutiérrez, M.; Kumar, M. S. N.; Tafalla, M. Chemically Active Outflow L 1157. *A&A* **2001**, *372* (3), 899–912.

(191) Benedettini, M.; Busquet, G.; Lefloch, B.; Codella, C.; Cabrit, S.; Ceccarelli, C.; Giannini, T.; Nisini, B.; Vasta, M.; Cernicharo, J.; Lorenzani, A.; di Giorgio, A. M. The CHESS Survey of the L1157-B1 Shock: The Dissociative Jet Shock as Revealed by Herschel–PACS. *A&A* **2012**, *539*, L3.

(192) Codella, C.; Fontani, F.; Ceccarelli, C.; Podio, L.; Viti, S.; Bachiller, R.; Benedettini, M.; Lefloch, B. Astrochemistry at Work in the L1157-B1 Shock: Acetaldehyde Formation. *Mon. Not. R. Astron. Soc. Lett.* **2015**, *449* (1), L11–L15.

(193) Boogert, A. C. A.; Gerakines, P. A.; Whittet, D. C. B. Observations of the Icy Universe. *Annu. Rev. Astron. Astrophys.* **2015**, *53* (1), 541–581.

(194) Podio, L.; Lefloch, B.; Ceccarelli, C.; Codella, C.; Bachiller, R. Molecular Ions in the

Protostellar Shock L1157-B1. *A&A* **2014**, *565*, A64.

(195) Enrique-Romero, J.; Rimola, A.; Ceccarelli, C.; Balucani, N. The (Impossible?) Formation of Acetaldehyde on the Grain Surfaces: Insights from Quantum Chemical Calculations. *Mon. Not. R. Astron. Soc. Lett.* **2016**, *459* (1), L6–L10.

(196) Vastel, C.; Ceccarelli, C.; Lefloch, B.; Bachiller, R. The Origin of Complex Organic Molecules in Prestellar Cores. *Astrophys. J. Lett.* **2014**, *795* (1), L2.

(197) Balucani, N.; Ceccarelli, C.; Taquet, V. Formation of Complex Organic Molecules in Cold Objects: The Role of Gas-Phase Reactions. *Mon. Not. R. Astron. Soc. Lett.* **2015**, *449* (1), L16–L20.

(198) Sorrell, W. H. Origin of Amino Acids and Organic Sugars in Interstellar Clouds. *Astrophys. J. Lett.* **2001**, *555* (2), L129.

(199) Mehringer, D. M.; Snyder, L. E.; Miao, Y.; Lovas, F. J. Detection and Confirmation of Interstellar Acetic Acid. *Astrophys. J. Lett.* **1997**, *480* (1), L71.

(200) Remijan, A.; Snyder, L. E.; Liu, S.-Y.; Mehringer, D.; Kuan, Y.-J. Acetic Acid in the Hot Cores of Sagitarrius B2(N) and W51. *Astrophys. J.* **2002**, *576* (1), 264.

(201) Shiao, Y.-S. J.; Looney, L. W.; Remijan, A. J.; Snyder, L. E.; Friedel, D. N. First Acetic Acid Survey with CARMA in Hot Molecular Cores. *Astrophys. J.* **2010**, *716* (1), 286.

(202) Hollis, J. M.; Lovas, F. J.; Jewell, P. R. Interstellar Glycolaldehyde: The First Sugar. *Astrophys. J. Lett.* **2000**, *540* (2), L107.

(203) Beltrán, M. T.; Codella, C.; Viti, S.; Neri, R.; Cesaroni, R. First Detection of Glycolaldehyde Outside the Galactic Center. *Astrophys. J. Lett.* **2009**, *690* (2), L93.

(204) Calcutt, H.; Viti, S.; Codella, C.; Beltrán, M. T.; Fontani, F.; Woods, P. M. A High-Resolution Study of Complex Organic Molecules in Hot Cores. *Mon. Not. R. Astron. Soc.* **2014**, *443* (4), 3157–3173.

(205) Garrod, R. T.; Herbst, E. Formation of Methyl Formate and Other Organic Species in the Warm-up Phase of Hot Molecular Cores. *A&A* **2006**, *457* (3), 927–936.

(206) Occhiogrosso, A.; Viti, S.; Modica, P.; Palumbo, M. E. A Study of Methyl Formate in Astrochemical Environments. *Mon. Not. R. Astron. Soc.* **2011**, *418* (3), 1923–1927.

(207) Woods, P. M.; Slater, B.; Raza, Z.; Viti, S.; Brown, W. A.; Burke, D. J. Glycolaldehyde Formation via the Dimerisation of the Formyl Radical. *ArXiv e-prints* **2013**.

(208) Burke, D. J.; Puletti, F.; Brown, W. A.; Woods, P. M.; Viti, S.; Slater, B. Glycolaldehyde, Methyl Formate and Acetic Acid Adsorption and Thermal Desorption from Interstellar Ices. *Mon. Not. R. Astron. Soc.* **2015**, *447* (2), 1444–1451.

(209) Johnson, T. J.; Sams, R. L.; Profeta, L. T. M.; Akagi, S. K.; Burling, I. R.; Yokelson, R. J.; Williams, S. D. Quantitative IR Spectrum and Vibrational Assignments for Glycolaldehyde Vapor: Glycolaldehyde Measurements in Biomass Burning Plumes. *J. Phys. Chem. A*, **2013**, *117* (20), 4096–4107

(210) *IR-Acetic Acid, visited on 25th of January 2016*, 2016.

(211) Lefloch, B.; Ceccarelli, C.; Codella, C.; Favre, C.; Podio, L.; Vastel, C.; Viti, S.; Bachiller, R. L1157-B1, a Factory of Complex Organic Molecules in a Solar-Type Star-Forming Region. *Mon. Not. R. Astron. Soc. Lett.* **2017**, *469* (1), L73–L77.

(212) 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* (4), 553–566.

(213) Simon, S.; Duran, M.; Dannenberg, J. J. How Does Basis Set Superposition Error Change the Potential Surfaces for Hydrogen-bonded Dimers? *J. Chem. Phys.* **1996**, *105* (24).

(214) Feng, G.; Favero, L. B.; Maris, A.; Vigorito, A.; Caminati, W.; Meyer, R. Proton Transfer in Homodimers of Carboxylic Acids: The Rotational Spectrum of the Dimer of Acrylic Acid. *J. Am. Chem. Soc.* **2012**, *134* (46), 19281–19286.

(215) Priem, D.; Ha, T.-K.; Bauder, A. Rotational Spectra and Structures of Three Hydrogen-Bonded Complexes between Formic Acid and Water. *J. Chem. Phys.* **2000**, *113* (1), 169–175.

(216) Mackenzie, R. B.; Dewberry, C. T.; Leopold, K. R. The Trimethylamine–Formic Acid Complex: Microwave Characterization of a Prototype for Potential Precursors to Atmospheric Aerosol. *J. Phys. Chem. A* **2016**, *120* (14), 2268–2273.

(217) Gou, Q.; Favero, L. B.; Bahamyirou, S. S.; Xia, Z.; Caminati, W. Interactions between Carboxylic Acids and Aldehydes: A Rotational Study of HCOOH–$CH_2O$. *J. Phys. Chem. A* **2014**, *118* (45), 10738–10741.

(218) Daly, A. M.; Sargus, B. A.; Kukolich, S. G. Microwave Spectrum and Structural Parameters for the Formamide-Formic Acid Dimer. *J. Chem. Phys.* **2010**, *133* (17),

174304.

(219) Evangelisti, L.; Spada, L.; Li, W.; Ciurlini, A.; Grabow, J.-U.; Caminati, W. Shape of the Adduct Formic Acid–Dimethyl Ether: A Rotational Study. *J. Phys. Chem. A* **2016**, *120* (18), 2863–2867.

(220) Evangelisti, L.; Spada, L.; Li, W.; Blanco, S.; Lopez, J. C.; Lesarri, A.; Grabow, J.-U.; Caminati, W. A Butterfly Motion of Formic Acid and Cyclobutanone in the 1 : 1 Hydrogen Bonded Molecular Cluster. *Phys. Chem. Chem. Phys.* **2017**, *19* (1), 204–209.

(221) Spada, L.; Gou, Q.; Giuliano, B. M.; Caminati, W. Interactions between Carboxylic Acids and Heteroaromatics: A Rotational Study of Formic Acid–Pyridine. *J. Phys. Chem. A* **2016**, *120* (27), 5094–5098.

(222) Duan, C.; Carvajal, M.; Yu, S.; Pearson, J. C.; Drouin, B. J.; Kleiner, I. THz Extended Spectrum of the Monodeuterated Methyl Formate ($DCOOCH_3$). *A&A* **2015**, *576*, A39.

(223) Tanjaroon, C.; Jäger, W. High-Resolution Microwave Spectrum of the Weakly Bound Helium-Pyridine Complex. *J. Chem. Phys.* **2007**, *127* (3), 34302.

(224) Maris, A.; Favero, L. B.; Velino, B.; Caminati, W. Pyridine-$CF_4$: A Molecule with a Rotating Cap. *J. Phys. Chem. A* **2013**, *117* (44), 11289–11292.

(225) Bettens, R. P. A.; Bauder, A. The Microwave Spectrum and Structure of the pyridine–CO Complex. *J. Chem. Phys.* **1995**, *102* (4), 1501–1509.

(226) Hunt, S. W.; Leopold, K. R. Molecular and Electronic Structure of $C_5H_5N$–$SO_3$: Correlation of Ground State Physical Properties with Orbital Energy Gaps in Partially Bound Lewis Acid–Base Complexes. *J. Phys. Chem. A* **2001**, *105* (22), 5498–5506.

(227) Grimme, S. Density Functional Theory with London Dispersion Corrections. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1* (2), 211–228.

(228) Delle Piane, M.; Corno, M.; Ugliengo, P. Does Dispersion Dominate over H-Bonds in Drug–Surface Interactions? The Case of Silica-Based Materials As Excipients and Drug-Delivery Agents. *J. Chem. Theory Comput.* **2013**, *9* (5), 2404–2415.

(229) Delle Piane, M.; Vaccari, S.; Corno, M.; Ugliengo, P. Silica-Based Materials as Drug Adsorbents: First Principle Investigation on the Role of Water Microsolvation on Ibuprofen Adsorption. *J. Phys. Chem. A* **2014**, *118* (31), 5801–5807.

(230) Puzzarini, C. Accurate Molecular Structures of Small- and Medium-Sized Molecules. *Int.*

*J. Quantum Chem.* **2016**, *116* (21), 1513–1519.

(231) Purvis III, G. D.; Bartlett, R. J. A Full Coupled-Cluster Singles and Doubles Model: The Inclusion of Disconnected Triples. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

(232) Moller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46* (7), 618–622.

(233) 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* (39), 16965–16975.

(234) Puzzarini, C.; Biczysko, M.; Barone, V.; Largo, L.; Peña, I.; Cabezas, C.; Alonso, J. L. Accurate Characterization of the Peptide Linkage in the Gas Phase: A Joint Quantum-Chemical and Rotational Spectroscopy Study of the Glycine Dipeptide Analogue. *J. Phys. Chem. Lett.* **2014**, *5* (3), 534–540.

(235) Pickett, H. M. The Fitting and Prediction of Vibration-Rotation Spectra with Spin Interactions. *J. Mol. Spectrosc.* **1991**, *148* (2), 371–377.

(236) H, H.; H, D. The Microwave Spectrum of Trans-2,3-Dimethyloxirane in Torsional Excited States. *Zeitschrift f{ü}r Naturforschung A*, 1996, *51*, 923.

(237) Giuliano, B. M.; Evangelisti, L.; Maris, A.; Caminati, W. Weak Hydrogen Bonds in σ-1,4-Difluorobenzene-Ammonia: A Rotational Study. *Chem. Phys. Lett.* **2010**, *485* (1), 36–39.

(238) Kisiel, Z.; Białkowska-Jaworska, E.; Pszczółkowski, L.; Milet, A.; Struniewicz, C.; Moszynski, R.; Sadlej, J. Structure and Properties of the Weakly Bound Trimer $(H_2O)_2HCl$ Observed by Rotational Spectroscopy. *J. Chem. Phys.* **2000**, *112* (13), 5767–5776.

(239) Cazzoli, G.; Dore, L.; Puzzarini, C. The Hyperfine Structure of the Inversion-Rotation Transition JK = $1_0 \leftarrow 0_0$ of $NH_3$ Investigated by Lamb-Dip Spectroscopy. *A{&}A* **2009**, *507* (3), 1707–1710.

(240) Feng, G.; Evangelisti, L.; Cacelli, I.; Carbonaro, L.; Prampolini, G.; Caminati, W. Oligomers Based on Weak Hydrogen Bond Networks: A Rotational Study of the Tetramer of Difluoromethane. *Chem. Commun.* **2014**, *50* (2), 171–173.

(241) Tatamitani, Y.; Liu, B.; Shimada, J.; Ogata, T.; Ottaviani, P.; Maris, A.; Caminati, W.; Alonso, J. L. Weak, Improper, C–O···H–C Hydrogen Bonds in the Dimethyl Ether Dimer. *J. Am. Chem. Soc.* **2002**, *124* (11), 2739–2743.

(242) Pérez, C.; Neill, J. L.; Muckle, M. T.; Zaleski, D. P.; Peña, I.; Lopez, J. C.; Alonso, J. L.; Pate, B. H. Water–Water and Water–Solute Interactions in Microsolvated Organic Complexes. *Angew. Chemie Int. Ed.* **2015**, *54* (3), 979–982.

(243) Bissantz, C.; Kuhn, B.; Stahl, M. A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem.* **2010**, *53* (14), 5061–5084.

(244) Berger, R.; Resnati, G.; Metrangolo, P.; Weber, E.; Hulliger, J. Organic Fluorine Compounds: A Great Opportunity for Enhanced Materials Properties. *Chem. Soc. Rev.* **2011**, *40* (7), 3496–3508.

(245) Kui, S. C. F.; Zhu, N.; Chan, M. C. W. Observation of Intramolecular C-H···F-C Contacts in Non-Metallocene Polyolefin Catalysts: Model for Weak Attractive Interactions between Polymer Chain and Noninnocent Ligand. *Angew. Chemie* **2003**, *115* (14), 1666–1670.

(246) Howard, J. A. K.; Hoy, V. J.; O'Hagan, D.; Smith, G. T. How Good Is Fluorine as a Hydrogen Bond Acceptor? *Tetrahedron* **1996**, *52* (38), 12613–12622.

(247) Dunitz, J. D.; Taylor, R. Organic Fluorine Hardly Ever Accepts Hydrogen Bonds. *Chem. Eur. J.* **1997**, *3* (1), 89–98.

(248) Thalladi, V. R.; Weiss, H.-C.; Bläser, D.; Boese, R.; Nangia, A.; Desiraju, G. R. C–H···F Interactions in the Crystal Structures of Some Fluorobenzenes. *J. Am. Chem. Soc.* **1998**, *120* (34), 8702–8710.

(249) Barbarich, T. J.; Rithner, C. D.; Miller, S. M.; Anderson, O. P.; Strauss, S. H. Significant Inter- and Intramolecular O-H···F-C Hydrogen Bonding. *J. Am. Chem. Soc.* **1999**, *121* (17), 4280–4281.

(250) Liu, X.; Borho, N.; Xu, Y. Molecular Self-Recognition: Rotational Spectra of the Dimeric 2-Fluoroethanol Conformers. *Chem. Eur. J.* **2009**, *15* (1), 270–277.

(251) Thomas, J.; Liu, X.; Jäger, W.; Xu, Y. Unusual H-Bond Topology and Bifurcated H-Bonds in the 2-Fluoroethanol Trimer. *Angew. Chemie Int. Ed.* **2015**, *54* (40), 11711–11715.

(252) Caminati, W.; Melandri, S.; Schnell, M.; Banser, D.; Grabow, J.-U.; Alonso, J. L. The

Fourier Transform Rotational Spectrum of Difluoromethane–water: Internal Motion of Water. *J. Mol. Struct.* **2005**, *742* (1), 87–90.

(253) Caminati, W.; Melandri, S.; Rossi, I.; Favero, P. G. The C–F···H–O Hydrogen Bond in the Gas Phase. Rotational Spectrum and Ab Initio Calculations of Difluoromethane-Water. *J. Am. Chem. Soc.* **1999**, *121* (43), 10098–10101.

(254) Tasinato, N.; Grimme, S. Unveiling the Non-Covalent Interactions of Molecular Homodimers by Dispersion-Corrected DFT Calculations and Collision-Induced Broadening of Ro-Vibrational Transitions: Application to $(CH_2F_2)_2$ and $(SO_2)_2$. *Phys. Chem. Chem. Phys.* **2015**, *17* (8), 5659–5669.

(255) Sure, R.; Grimme, S. Comprehensive Benchmark of Association (Free) Energies of Realistic Host–Guest Complexes. *J. Chem. Theory Comput.* **2015**, *11* (8), 3785–3801.

(256) Tasinato, N.; Moro, D.; Stoppa, P.; Charmet, A. P.; Toninello, P.; Giorgianni, S. Adsorption of $F_2CCFCl$ on $TiO_2$ Nano-Powder: Structures, Energetics and Vibrational Properties from DRIFT Spectroscopy and Periodic Quantum Chemical Calculations. *Appl. Surf. Sci.* **2015**, *353*, 986–994.

(257) Biczysko, M.; Panek, P.; Scalmani, G.; Bloino, J.; Barone, V. Harmonic and Anharmonic Vibrational Frequency Calculations with the Double-Hybrid B2PLYP Method : *J. Chem. Theory Comput.* **2010**, *6* (7), 2115–2125.

(258) Amos, R. D.; Handy, N. C.; Green, W. H.; Jayatilaka, D.; Willetts, A.; Palmieri, P. Anharmonic Vibrational Properties of $CH_2F_2$: A Comparison of Theory and Experiment. *J. Chem. Phys.* **1991**, *95* (11), 8323–8336.

(259) 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* (12), 1065–1076.

(260) Kendall, R. A.; Dunning Jr., T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

(261) Peterson, K. A.; Jr., T. H. D. Accurate Correlation Consistent Basis Sets for Molecular Core–valence Correlation Effects: The Second Row Atoms Al–Ar, and the First Row Atoms B–Ne Revisited. *J. Chem. Phys.* **2002**, *117* (23), 10548–10560.

(262) Grabow, J.; Stahl, W.; Dreizler, H. A Multioctave Coaxially Oriented Beam-resonator Arrangement Fourier-transform Microwave Spectrometer. *Rev. Sci. Instrum.* **1996**, *67* (12), 4072–4084.

(263) Balle, T. J.; Flygare, W. H. Fabry–Perot Cavity Pulsed Fourier Transform Microwave Spectrometer with a Pulsed Nozzle Particle Source. *Rev. Sci. Instrum.* **1981**, *52* (1), 33–45.

(264) Caminati, W.; Millemaggi, A.; Alonso, J. L.; Lesarri, A.; López, J. C.; Mata, S. Molecular Beam Fourier Transform Microwave Spectrum of the Dimethylether–xenon Complex: Tunnelling Splitting and $^{131}$Xe Quadrupole Coupling Constants. *Chem. Phys. Lett.* **2004**, *392* (1), 1–6.

(265) Ruoff, R. S.; Klots, T. D.; Emilsson, T.; Gutowsky, H. S. Relaxation of Conformers and Isomers in Seeded Supersonic Jets of Inert Gases. *J. Chem. Phys.* **1990**, *93* (5), 3142–3150.

(266) Novick, S. E.; Harris, S. J.; Janda, K. C.; Klemperer, W. Structure and Bonding of KrClF: Intermolecular Force Fields in Van Der Waals Molecules. *Can. J. Phys.* **1975**, *53* (19), 2007–2015.

# *Publications list*

1-Skouteris, D; Balucani, N.; Ceccarelli, C.; Vazart, F.; Puzzarini, C.; Barone, V.; Codella, C.; Lefloch, B. *Astrophys. J.*, **2017**, *submitted;*

2- Codella, C.; Ceccarelli, C.; Caselli, P.; Balucani, N.; Barone, V.; Fontani, F.; Lefloch, B.; Podio, L.; Viti, S.; Feng, S.; Bachiller, R.; Bianchi, E.; Dulieu, F.; Jiménez-Serra, I.; Holdship, J.; Neri, R.; Pineda, J.-E.; Pon, A.; Sims, I.; Spezzano, S.; Vasyunin, A.; Alves, F.; Bizzocchi, L.; Bottinelli, S.; Caux, E.; Chacón-Tanarro, A.; Choudhury, R.; Coutens, A.; Favre, C.; Hily-Blant, P.; Kahane, C.; Jaber Al-Edhari, A.; Laas, J.; López-Sepulcre, A.; Ospina, J.; Oya, Y.; Punanova, A.; Puzzarini, C.; Rimola, C.; Sakai, N.; Skouteris, D.; Taquet, V.; Testi, L.; Theulé, P.; Ugliengo, P.; Vastel, C.; Vazart, F.; Wiesenfeld, L.; Yamamoto, *S. A&A*, **2017**, *submitted*;

3- Spada, L.; Tasinato, N.; Bosi, G.; Vazart, F.; Barone, V.; Puzzarini, C. *J. Mol. Spectrosc.*, **2017**, *DOI: http://dx.doi.org/10.1016/j.jms.2017.04.001*;

4- Evangelisti, L.; Spada, L.; Li, W.; Vazart, F.; Barone, V.; Caminati, W. *Angew. Chem. Int. Ed.*, **2017**, *56*, 1-5;

5- Spada, L.; Tasinato, N.; Vazart, F.; Barone, V.; Caminati, W.; Puzzarini, C. *Chem. Eur. J.*, **2017**, *DOI: 10.1002/chem.201606014*;

6- Skouteris, D.; Vazart, F.; Ceccarelli, C.; Balucani, N.; Puzzarini, C; Barone, V. *Mon. Not. R. Astron. Soc.*, **2017**, *arXiv:1701.06138*;

7- Vazart, F.; Calderini, D.; Puzzarini, C.; Skouteris, D.; Barone, V. *J. Chem. Theory Comput.*, **2016**, *12 (11)*, 5385-5397;

8- Vazart, F.; Savel, P.; Latouche, C.; Barone, V.; Camerel, F.; Roisnel, T.; Fillaut, J.-L., Akdas-Kilig, H.; Achard, M. *Dalton Trans.*, **2016**, *45*, 6566-6573;

9- Vazart, F.; Latouche, C.; Cimino, P., Barone, V. *J. Chem. Theory Comput*, **2015**, *11 (9)*, 4364-4369;

10- Barone, V.; Latouche, C.; Skouteris, D.; Vazart, F.; Balucani, N.; Ceccarelli, C.; Lefloch, B. *Mon. Not. R. Astron. Soc.*, **2015**, *453*, L31-L35;

11- Vazart, F.; Latouche, C.; Skouteris, D.; Balucani, N.; Barone, V. *Astrophys. J.*, **2015**, *810 (2)*, 111;

12- Vazart, F.; Latouche, C. *Theor. Chem. Acc.*, **2015**, *134 (12)*, 144;

13- Vazart, F.; Latouche, C.; Bloino, J.; Barone, V. *Inorg. Chem.*, **2015**, *54 (11)*, 5588-5595;

14- Vazart, F.; Calderini, D.; Skouteris, D.; Latouche, C.; Barone, V. *J. Chem. Theory Comput.*, **2015**, *11 (3)*, 1165-1171.