

# Looking for the bricks of the life in the interstellar medium: The fascinating world of astrochemistry

VINCENZO BARONE

*Scuola Normale Superiore - Piazza dei Cavalieri 7, I-56126 Pisa, Italy*

CRISTINA PUZZARINI

*Dipartimento di Chimica "Giacomo Ciamician", Università di Bologna  
Via F. Selmi 2, I-40126 Bologna, Italy*

**Summary.** — The discovery in the interstellar medium of molecules showing a certain degree of complexity, and in particular those with a prebiotic character, has attracted great interest. A complex chemistry takes place in space, but the processes that lead to the production of molecular species are a matter of intense discussion, the knowledge still being at a rather primitive stage. Debate on the origins of interstellar molecules has been further stimulated by the identification of biomolecular building blocks, such as nucleobases and amino acids, in meteorites and comets. Since many of the molecules found in space play a role in the chemistry of life, the issue of their molecular genesis and evolution might be related to the profound question of the origin of life itself. Understanding the underlying chemical processes, including the production, reactions and destruction of compounds, requires the concomitant study of spectroscopy, gas-phase reactivity, and heterogeneous processes on dust-grains. The aim of this contribution is to provide a general view of a complex and multifaceted challenge, while focusing on the role played by molecular spectroscopy and quantum-chemical computations. In particular, the derivation of the molecular spectroscopic features and the investigation of gas-phase formation routes of prebiotic species in the interstellar medium are addressed from a computational point of view.

## 1. – Introduction

“Chemistry without numbers is poetry: astrochemistry without numbers is myth” [1]: The exploration of the molecular universe is a long journey through the wonders of astronomy and chemistry (astrochemistry) back to the origin of life on Earth. While

molecular complexity in the universe is undisputed, it is not at all clear how these species could be and still are produced and how chemical evolution proceeds toward biological building-blocks.

Since the discovery of the interstellar gas [2] and dust [3], it has been thought that the extreme conditions of the interstellar medium (ISM, with temperatures ranging between 10 and 100 K, very low density and ionizing radiations) could only allow the presence of atoms. Nevertheless, nearly 80 years ago, the first evidence of molecules in interstellar clouds was obtained: McKellar identified sharp optical absorption lines of CN and CH using the Mount Wilson 100-inch telescope [4]. One year later, Douglas and Herzberg found the  $\text{CH}^+$  cation in interstellar clouds [5]. However, it was only when radioastronomy took hold—nearly fifty years ago—that it became clear that the ISM was home to a diverse array of interesting polyatomic (both organic and inorganic) molecules. Particularly fascinating amongst the molecules discovered in the ISM are the so-called “interstellar complex organic molecules” (iCOMs), such as glycoaldehyde [6], acetamide [7] and methyl acetate [8], that combine multiple functional groups illustrative of the rich chemistry of the ISM. Even more fascinating is the recent discovery of the first chiral molecule in space, methyloxirane [9], here recalling that chirality is the signature of life (at least, as we know it). Currently, the census of interstellar molecular species exceeds two hundreds [10], and the rate of discovery continues at a rapid pace. The presence of a rich chemistry in space has led to the emergence of astrochemistry, an interdisciplinary and multifaceted field at the interface of chemistry, physics, and astronomy, which encompasses astronomical observations and modelling, as well as theoretical and experimental laboratory investigations [1, 11, 12].

Many of the molecules found in the ISM play a role in the chemistry of life; thus, the question of molecular genesis in the ISM might be related to the profound question of the origin of life itself. This aspect of astrochemistry has attracted intense interest in the general scientific community, and spurred the search for biomolecular building blocks in the interstellar space. The progress of observational studies revealed chemical diversity in space, a source-to-source variation of the chemical composition, and led to the concept of chemical evolution (*e.g.*, [11]). Understanding the underlying chemical processes, including the production, reactions and destruction of compounds, is one of the main aims of astrochemistry. As suggested above, understanding these processes can help to address one of the most intriguing questions in all of science, as these surely help to set the stage for the emergence of life to occur, both on Earth and, maybe, elsewhere in the cosmos. Intriguingly, evidence has been presented that nucleobases (the fundamental elements of RNA and DNA) and amino acids exist in other extraterrestrial sources than Earth [13].

Focusing on the mystery of the emergence of life on Earth, two alternative theories have been suggested so far [14]: 1) exogenous delivery and 2) endogenous synthesis. While in the first theory prebiotic molecules are postulated to have come from space on comets, asteroids and meteorites, the latter theory states that the synthesis of simple organic molecules having a potential relation to the origin of life occurred on our planet, starting from simple parent molecules already present, such as liquid water, methane and ammonia. The rationale behind the endogenous theory is the Urey-Miller experiment [15,

16], as it showed that most of the twenty common amino acids, as well as pyrimidines and purines, could be produced from these simple precursor molecules with the help of an electrical discharge, which is an effective laboratory proxy for atmospheric lightning and other energy sources. However, the complexity of a planet cannot be reproduced in a single laboratory experiment, with Titan—the largest moon of Saturn—providing a way-out. Indeed, Titan has been postulated as a model of primitive Earth and it is expected to be a rich source of information on prebiotic organic synthesis in the atmosphere of primitive Earth. Instead, the exogenous delivery theory is based on the fact that prebiotic iCOMs have been observed in the ISM, thus postulating that the prebiotic molecules, formed in the solar nebula and preserved during the early phases of the solar system formation in the body of comets, asteroids, and meteorites, were delivered to Earth by cometary and meteoritic impacts. In any event, it has been pointed out that, whether they were delivered to Earth from space or synthesized from simpler building blocks, prebiotic molecules then evolved to form more complex molecules. While the endogenous theory will not be addressed, the focus will be on the first theory and on the methodology, with particular emphasis on computational chemistry, required for its investigation.

## 2. – Astronomical spectroscopy

In the field of astrochemistry, molecular spectroscopy plays a central role. Because of the tremendous distances involved, there is no chance to perform direct experiments on astrochemical processes, and detection via interaction of molecules with radiation is the only viable route of investigation [17]. The astronomical observation of the spectroscopic features of a given molecule is the definitive, unequivocal proof of its presence in the astronomical environment under consideration (see, *e.g.*, refs. [9, 18, 19]).

The overwhelming majority of gas-phase chemical species in the cosmos have been discovered via their rotational signatures with the help of radioastronomy, which works in the radio- and micro-wave spectral windows. Indeed, the advent of radioastronomy (in the early 1960s) enabled the detection of new molecules with a trend that has continued at a nearly linear rate. The systematic observation of a given astronomical source leads to line surveys that in principle provide a complete census of the molecular content (provided that they are polar molecules, which is mostly the case). However, the improvement in the observational sensitivity that has characterized the last two decades has posed and still poses challenges in the line assignment. In particular, the increased sensitivity of ALMA (Atacama Large Millimeter/submillimeter Array of radiotelescopes) has been revealing new spectroscopic features in the so-called confusion limit. Therefore, broadband spectral line surveys still show hundreds of features not assignable to transitions of already characterized molecules [20], even if one accounts for transitions of vibrationally excited or rare isotopic species.

Unrecognized features extend across the electromagnetic spectrum, and are not limited to radioastronomy. In the infrared (IR) and ultraviolet (UV) regions of the electromagnetic spectrum, the so-called unidentified infrared emission bands (UIRs) still continue to elude conclusive molecular identification. These emission features are well rec-

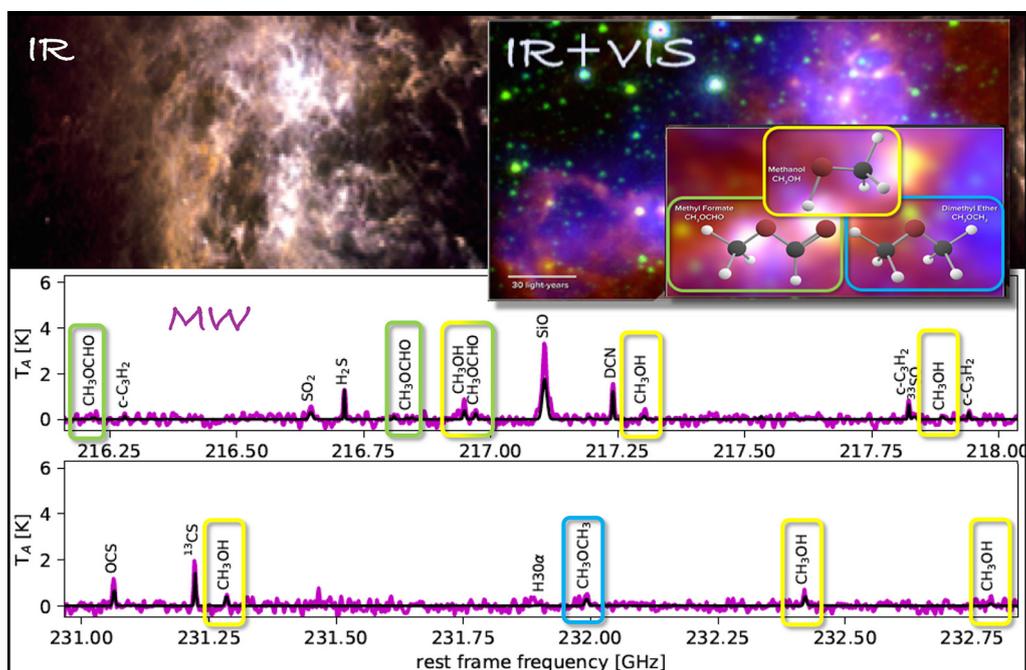


Fig. 1. – The Large Magellanic Cloud observed in different regions of the electromagnetic spectrum. In the top panel, the far-infrared image shows a portion of the galaxy, with the zoom-in images on the right being a combination of mid-infrared data from Spitzer and visible (H-alpha) data from the Blanco 4-meter telescope. In the bottom panel, two Band 6 spectral windows of ALMA are shown: the rotational spectroscopic fingerprints of methanol, dimethyl ether, and methyl formate observed are highlighted [23].

ognized to be due to polycyclic aromatic compounds (PAHs), but no individual molecule has been definitely identified from their analysis. Analogous is the case of the diffuse interstellar bands (DIBs) [21], which are sharp features ranging from the IR to the UV, that are still nearly completely unassigned, with the only undisputed assignment being that of  $C_{60}^+$  [22].

Figure 1 gives a flavor of the interplay of different astronomical techniques: it shows a nice combination of different observations, in different regions of the electromagnetic spectrum, of the Magellanic cloud where the first extragalactic detection of dimethyl ether and methyl formate was reported [23].

### 3. – Computational approaches for astrochemistry: From spectroscopic features to formation pathways

Computational chemistry plays an important role in everything related to laboratory studies. As mentioned above, molecular spectroscopy provides the means for detecting molecules in space: astronomical detections are guided by laboratory experimental mea-

surements, which in turn are supported and complemented by state-of-the-art quantum-chemical (QC) calculations (*e.g.*, refs. [12,24,25]). Because of difficulties in mimicking in the laboratory the extreme conditions that characterize the ISM, accurate computational approaches are of fundamental importance in deriving feasible reaction mechanisms and the associated kinetics (*e.g.*, refs. [26-28]).

### 3.1. *Quantum-chemical predictions of equilibrium structures and relative energies.* –

The starting point of any characterization is the accurate determination of molecular structures and of their energies. Indeed, when considering molecules with a certain degree of complexity, different conformers or isomers might be present and it is thus important to evaluate accurately their relative stability. In the present context, it is worthwhile mentioning that different isomers of the same molecular system have been detected. Lattelais and coworkers have interpreted the observations of molecular isomers in terms of the so-called “minimum energy principle”, according to which the most thermodynamically stable molecular isomers are those more favored in the chemical processes occurring [29,30]. Even if there are examples that question such a principle (see, *e.g.*, [31,32]), it is fundamental to provide accurate predictions of relative energies for different isomeric forms by means of state-of-the-art QC computations. Accurate energy calculations are also important to support molecular spectroscopy and to investigate reaction kinetics. Concerning the former, rigorous approaches for the resolution of the nuclear problem need to be combined with state-of-the-art electronic structure computations. Moving to formation pathways, at the typical low temperatures of the ISM, rate constants are exquisitely sensitive to energetics and activation barrier heights.

To reach high accuracy in computational approaches, the errors due to the truncation of both basis set and wavefunction, the so-called one- and N-electron errors, respectively, should be reduced as much as possible in QC calculations. If the systems under consideration do not have multi-configurational character, that is, are well described by a single-reference wave function, the coupled-cluster (CC) level of theory employing the CC singles and doubles (CCSD) approximation augmented by a perturbative treatment of triple excitations (CCSD(T)) [33] represents the best compromise between accuracy and computational cost, and this led to define it as the “gold standard” for accurate computations. Even if CCSD(T) provides a good description of the electronic wavefunction, important contributions such as basis-set effects should be taken into account. To this purpose, composite schemes, which rely on the additivity approximation, thus evaluating the various contributions at the best possible level and combining them together, have been developed (see, *e.g.*, refs. [34-41]).

Focusing on energetics, the reference scheme for our strategy is the HEAT approach [35], which allows for achieving high accuracy for thermochemistry, indeed being able to reach the so-called sub-kJ accuracy. Even if different HEAT protocols have been proposed, in all schemes the starting point is the CCSD(T) method within the frozen-core (fc) approximation. The first important corrections are the extrapolation to the complete basis set (CBS) limit, in order to recover the error due to the basis-set truncation, and the contribution of core-valence correlation (CV), evaluated as difference between CCSD(T)

calculations performed correlating all electrons and within the frozen-core approximation using the same basis set. The full CC singles, doubles and triples (CCSDT) [42-44] and CC singles, doubles, triples, quadruples (CCSDTQ) [45] models are employed to further minimize the error associated to the truncation of the wavefunction, with the corrections due to a full treatment of triples (fT) and quadruples (fQ) being included using small basis sets (usually of double-, triple-zeta quality). Overall, the scheme can be shortly denoted as CCSD(T)/CBS+CV+fT+fQ. Other important, even if minor, contributions, are scalar relativistic effects and diagonal Born-Oppenheimer corrections. The computational cost can be largely reduced by retaining only the extrapolation to the CBS limit at the fc-CCSD(T) level and the incorporation of the CV corrections, thus leading to the so-called CCSD(T)/CBS+CV scheme, which is rather well tested in the literature (see, *e.g.*, refs. [46-49]) and is demonstrated to provide results with an accuracy well within 0.5 kcal/mol. To extend the applicability of composite schemes to larger systems, an effective solution is offered by the so-called “cheap” approach [50, 51], which exploits the second-order Møller-Plesset theory (MP2) [52] to perform the extrapolation to the CBS limit and to include CV corrections, while keeping the CCSD(T) method in conjunction with a triple-zeta basis set as starting point. Despite the remarkable reduction of the computational cost (which is at the basis of the denomination of this protocol as “cheap”), it has been demonstrated that this approach provides accurate results (see, *e.g.*, refs. [50, 53]), also for flexible systems [50] and molecular complexes [54]. The situation is more involved for medium-sized open-shell systems, whose effective treatment requires further developments [55-57].

Moving to equilibrium structure determinations, the so-called “gradient” scheme [36, 37, 58] is based on the minimization of the energy gradient set up according to the specific requirements, the target accuracy, and the dimension of the system under consideration. While details on this approach can be found in refs. [24, 36, 37] the CCSD(T)/CBS+CV+fT+fQ scheme is able to provide an accuracy better than 0.001 Å for bond lengths and 0.05 degrees for angles (see, for example, refs. [36, 37, 59, 60] and references therein). However, inclusion of fT and fQ corrections makes this approach computationally very expensive. If these contributions are ignored, thus leading to the CCSD(T)/CBS+CV approach, a composite scheme affordable also for medium-sized molecules is obtained, which is characterized by a limited loss of accuracy. According to the available literature (see, for example, refs. [24, 37, 53, 61]), the corresponding geometrical parameters show an accuracy of 0.001–0.002 Å for distances and 0.05–0.1 degrees for angles. If the additivity approximation is directly applied to structural parameters, thus assuming that they show the same behavior as the energy (see, *e.g.*, refs. [62-65]), the so-called “geometry” composite schemes are obtained. Among them, the approach largely exploited is the scheme denoted as “cheap geometry”, which is entirely analogous to the protocol introduced above for energetics and can be effectively applied for accurate structural determinations of systems that are too large for composite approaches entirely based on CCSD(T) calculations [50, 61, 65, 66]. When the dimension of the system under consideration is such that a composite scheme is computationally too much expensive, the double-hybrid B2PLYP functional [67, 68] in conjunction with a triple-zeta quality

basis set provides a good alternative in predicting equilibrium structures [53, 69-71]. Indeed, the B2PLYP/(aug-)cc-pVTZ level of theory is able to provide better estimates than fc-CCSD(T)/cc-pVTZ, with the accuracy on bond distances ranging from 0.001 Å to 0.003 Å. Furthermore, it is important to note that the choice of the geometry does not affect significantly energetics, thus allowing the application of CCSD(T)-based composite schemes for electronic energy evaluations well beyond the “medium-sized molecule” limit.

**3.2. Quantum-chemical predictions of spectroscopic parameters.** – The starting point for the development of any astrochemical model is the knowledge whether a molecule is present in the astronomical environment considered and, if so, its abundance, with the spectroscopic signatures being the “molecular fingerprints” that allow for the unequivocal proof of the presence of a given chemical species [11, 72]. As briefly mentioned above, the spectroscopic features required to guide astronomical searches are accurately obtained in laboratory spectroscopy studies that are increasingly assisted by QC calculations to support spectral recordings and analyses (see, *e.g.*, refs. [18, 73-82]). Thanks to the developments in theoretical methodologies as well as in hardware facilities, computational spectroscopy is also able to play the role of complementing experiment, thus providing the missing information that cannot be obtained experimentally.

In the following, the methodology for accurate QC calculations of spectroscopic properties is presented and discussed.

**3.2.1. Rotational spectroscopy.** Focusing on pure rotational spectroscopy, the most important terms of the effective rotational Hamiltonian, within the semi-rigid rotor approximation, can be summarized as follows [83]:

$$(1) \quad H_{rot} = H_R + H_{qcd} + H_{scd} + \dots,$$

where

- 1)  $H_R$  is the rigid-rotor Hamiltonian, which contains the equilibrium rotational constants  $B_\tau^{eq}$  ( $\tau$  denoting a generic inertial axis):

$$(2) \quad B_\tau^{eq} = \frac{\hbar^2}{2hcI_{\tau\tau}^{eq}},$$

where  $I_{\tau\tau}^{eq}$  is the  $\tau$ -th diagonal element of the equilibrium inertia tensor,  $\mathbf{I}^{eq}$ ;

- 2)  $H_{qcd}$  and  $H_{scd}$  are the quartic and sextic centrifugal terms, respectively [84];
- 3) the dots refer to the possibility of including higher-order centrifugal terms or different contributions.

The Hamiltonian of eq. (1) does not account for the effect of molecular vibrations, which —however— cannot be neglected. The dependence of the rotational constants on the vibrational motion can be conveniently described by means of vibrational perturbation theory to second order (VPT2) [83, 85]. Focusing on the vibrational ground state,

which is—in the majority of the cases—that of interest, the rotational constant is given by the equilibrium contribution augmented by a corrective term

$$(3) \quad B_{\tau}^0 = B_{\tau}^{eq} + \Delta B_{\tau}^0 = B_{\tau}^{eq} - \frac{1}{2} \sum_{i=1}^N \alpha_{i,\tau},$$

where  $B_{\tau}^0$  denotes a generic rotational constant of the vibrational ground state and  $\alpha_{i,\tau}$ 's are the so-called vibration-rotation interaction constants (the sum running over the vibrational normal modes).

$B_{\tau}^{eq}$  provides the most important contribution to the corresponding rotational constant, the corresponding value indeed accounting for about 90% to 97% of the entire term. From a computational point of view, equilibrium rotational constants are straightforwardly obtained from geometry optimizations,  $B_{\tau}^{eq}$  only depending on the molecular structure and isotopic masses. Anharmonic force field calculations are instead required to obtain the vibrational corrections to equilibrium rotational constants. Moving to the centrifugal-distortion Hamiltonian, the quartic terms only depend on the harmonic part of the potential, while the computation of the sextic centrifugal-distortion constants involves harmonic, anharmonic, and Coriolis perturbation terms, thus requiring anharmonic force field computations. Going in some details, by resorting to the composite schemes previously introduced,  $B_{\tau}^{eq}$ 's can be accurately determined. Based on the results of two recent statistical studies [38, 86], a standard deviation smaller than 0.1% is obtained at the CCSD(T)/CBS+CV+fT+fQ level and also by means of the more affordable CCSD(T)/CBS+CV scheme. Once moving from equilibrium to vibrational ground-state rotational constants, the vibrational corrections need to be evaluated, with methods rooted in the density functional theory (DFT) being effective to this aim. Indeed, even the effective—in terms of computational cost—B3LYP/SNSD level [87-89] has been demonstrated to be suitable to meet the required accuracy [90]. As far as centrifugal distortion constants are concerned, they are obtained as byproducts of the calculations required to evaluate the vibrational corrections to rotational constants, with different levels of theory, ranging from HF-SCF to CCSD(T), providing very similar results [91].

Spectroscopic parameters, obtained following the methodology described above, allow for the prediction of rotational transitions with a relative accuracy of about 0.1% or better in the centimeter/millimeter-wave region, with this upper limit increasing up to 0.15–0.2% in the THz region (see, *e.g.*, refs. [38, 86, 92, 93]). To give an example, the best computed parameters can predict rotational frequencies at 100 GHz and 2 THz with an uncertainty of about 100 MHz and 3–4 GHz, respectively. This accuracy is exemplified by fig. 2 for two extreme cases: a rotational transition of the linear HC<sub>5</sub>N molecule in the centimeter-wave region and a rotational transition in the THz regime for two isotopic species of the HCO<sup>+</sup> cation. For the latter, the comparison between computed (CCSD(T)/CBS+CV+fT+fQ data augmented by vibrational corrections at the CCSD(T)/cc-pCVQZ level) and experimental frequencies is provided.

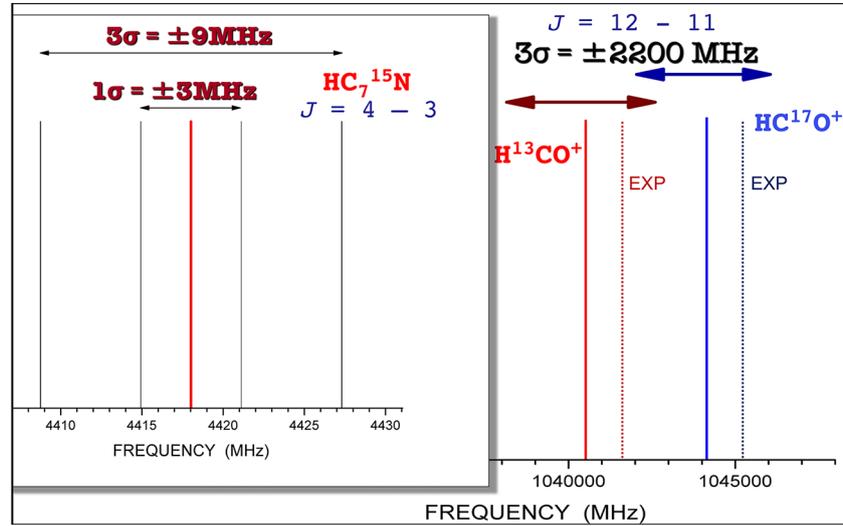


Fig. 2. – On the left: the  $J = 4 \leftarrow 3$  rotational transition of the main isotopic species of  $\text{HC}_5\text{N}$ . On the right: the  $J = 12 \leftarrow 11$  rotational transition of  $\text{H}^{13}\text{CO}^+$  and  $\text{HC}^{17}\text{O}^+$ .

**3.2.2. Vibrational spectroscopy.** Vibrational spectra are generally interpreted in terms of the harmonic oscillator model possibly scaling the computed frequencies by empirical factors in order to account for anharmonicity and other limitations due to the computational approach. However, in recent years, thanks to hardware and software advances, it has become possible to employ more sophisticated models that take explicit account of anharmonicity and ro-vibrational couplings. While small-sized molecules can be treated using very refined Hamiltonians and numerical approaches, VPT2 based on cartesian normal modes still represents the method of choice for medium-sized systems. VPT2 approaches are rooted in the Watson Hamiltonian [94] and fourth-order Taylor expansions of the potential energy [95], leading to the following vibrational Hamiltonian ( $H_{vib}$ ) expressed in terms of normal modes ( $\mathbf{q}$ ), their conjugate momenta ( $\mathbf{p}$ ), harmonic wavenumbers ( $\omega$ ), cubic and quartic reduced force constants ( $k$ ), equilibrium rotational constants ( $\mathbf{B}^{eq}$ ) and Coriolis couplings ( $\zeta$ ):

$$(4) \quad H_{vib} = \frac{1}{2} \sum_{i=1}^N \omega_i (\mathbf{p}_i^2 + \mathbf{q}_i^2) + \frac{1}{6} \sum_{i,j,k=1}^N k_{ijk} \mathbf{q}_i \mathbf{q}_j \mathbf{q}_k + \frac{1}{24} \sum_{i,j,k,l=1}^N k_{ijkl} \mathbf{q}_i \mathbf{q}_j \mathbf{q}_k \mathbf{q}_l + \sum_{\tau} \mathbf{B}_{\tau}^{eq} \sum_{i,j,k,l=1}^N \zeta_{ij,\tau} \zeta_{kl,\tau} \sqrt{\frac{\omega_i \omega_k}{\omega_j \omega_l}} \mathbf{q}_i \mathbf{p}_j \mathbf{q}_k \mathbf{p}_l + U,$$

where  $i, j, k, l$  run on the vibrational modes, while  $\tau$  runs on the inertial axes;  $U$  is a mass-dependent contribution, which leads to a nearly constant shift in the spectrum that has negligible spectroscopic importance.

The second-order perturbative solution of the Schrödinger equation associated to the Hamiltonian of eq. (4) leads to an analytical expression for the energy levels  $E_m$  (in  $\text{cm}^{-1}$ ), with  $m$  being a generic vibrational state:

$$(5) \quad E_m = E_0 + \sum_{i=1}^N v_i^m \omega_i + \sum_{i,j=1}^N \chi_{ij} \left[ v_i^m v_j^m + \frac{1}{2} (v_i^m + v_j^m) \right],$$

where  $E_0$  is the zero-point vibrational energy (ZPE),  $v_i^m$  is the number of quanta associated to state  $m$  of mode  $i$ , and the  $\chi$  matrix elements are explicit functions of cubic and quartic force constants together with Coriolis couplings [96,97]. Transition energies from the ground state,  $\nu_m$ , are then straightforwardly obtained as  $E_m - E_0$  difference.

Analytical expressions for the intensities of the different vibrational spectroscopies including mechanical, electric and, possibly, magnetic anharmonicities can be obtained by a double second-order perturbative approach [98]. Together with increased accuracy, these expressions give access to the intensities of overtones and combination bands, which vanish at the harmonic level. However, two critical issues of VPT2 are the presence of large amplitude motions (LAMs) and of nearly degenerate vibrational states. Concerning the first problem, loosely coupled LAMs can be treated individually by one-dimensional variational approaches and integrated with reduced-dimensionality VPT2 approaches [99]. More strongly coupled situations require, except for very small systems, further developments including the employment of generalized internal coordinates [100]. Moving to resonances, their impact on the vibrational energies can be effectively treated in the framework of the so-called generalized (G) VPT2 model in which, their contribution is neglected in the first perturbative step (deperturbed (D) VPT2) and then reintroduced in a successive reduced-dimensionality variational evaluation. An alternative approach, known as degeneracy-corrected perturbation theory (DCPT2) [101], replaces all potentially divergent terms with non-resonant forms and has been recently improved to lead to the hybrid HDCPT2 scheme [102], which combines DCPT2 and VPT2 results to overcome some shortcomings of the former approach far from resonances. This is the method of choice for thermodynamic and/or kinetic studies, which require comparable treatments for different stationary points (see next section). None of these procedures is necessary for the evaluation of  $E_0$  since the corresponding expression can be rearranged in a resonance-free formulation [103]. On the other hand, the situation is more involved for intensities, since additional types of resonances come into play involving also electric and magnetic contributions. A general automatic procedure for their treatment is still lacking although significant progresses have been made also in this connection [96,98].

As far as computational requirements are concerned, the composite schemes discussed in the preceding sections can also be applied to the accurate computation of harmonic force fields. According to the literature on this topic (see, for example, refs. [56,61,104]), approaches based on CCSD(T) are able to provide harmonic frequencies with an accuracy of  $5\text{--}10 \text{ cm}^{-1}$ . However, their applicability is limited to small- to medium-sized molecules. As in the case of equilibrium structures, double-hybrid functionals (*e.g.*, B2PLYP) in

conjunction with triple-zeta quality basis sets are the methods of choice for larger systems, showing an accuracy of 8–15 cm<sup>-1</sup>. Due to the computational requirements, cubic and semi-diagonal quartic force constants are usually evaluated at a lower level of theory. In this connection, the B3LYP global hybrid functional in conjunction with a double-zeta basis set is usually adequate, with B2PLYP clearly performing better. Accuracy can be improved by means of hybrid approaches, where the harmonic part of the potential is determined at the CCSD(T) level or even employing a composite scheme, whereas B3LYP or B2PLYP are used for the anharmonic part: the mean absolute errors (MAEs) for a large panel of fundamentals as well as overtone and combination bands are 6–8 (B3LYP) and 5–7 cm<sup>-1</sup> (B2PLYP) (see, for example, ref. [76]), respectively. Once again, the situation is more involved for intensities due to the lack of systematic studies devoted to the validation of different approaches applicable to medium- to large-sized molecules. However, the available literature suggests that the same hybrid protocols can be applied also in this connection, provided that basis sets are carefully chosen (*e.g.*, some electric properties require diffuse polarization functions).

**3.3. Thermochemistry and kinetics.** – Within the time-independent approach followed in this contribution, the link between microscopic (electronic energies, rotational and vibrational levels, etc.) and macroscopic (thermodynamic functions, reaction rates) quantities is provided by the partition function  $Z(T)$  of energy minima and the corresponding quantity,  $Z(T)^\ddagger$ , defined for transition states by removing the motion along the reaction coordinate (RC) and the coupling of RC with perpendicular vibrations. At not too high temperatures, vibrations can be considered separable from translations and rotations, the latter two motions being treated within a classical approximation. In this framework, the overall partition function is the product of vibrational ( $Z_V$ ), rotational ( $Z_r$ ) and translational ( $Z_t$ ) contributions, which —within the rigid-rotor harmonic-oscillator approximation— have well-known analytical expressions. In particular, the vibrational term is given by

$$(6) \quad Z_V(T) = \frac{e^{-\beta E_0}}{\prod_i (1 - e^{-\beta \Delta_i})},$$

where  $\beta = 1/kT$ ,  $E_0$  is the ZPE contribution, and  $\Delta_i$  is the lowest excitation energy of mode  $i$ . In the framework of VPT2, both  $E_0$  and  $\Delta_i$  have analytical expressions, so that the leading anharmonic contributions can be included by simply replacing the harmonic values of these two quantities with their VPT2 counterparts. This model, known as simple perturbation theory (SPT) [105], provides remarkably accurate results at a reasonable cost. The same model can be employed for the bound vibrations of TSs leading to the same expression for the canonical reaction rate obtained at the harmonic level in the framework of the transition state theory (TST) [106], provided that tunneling and non-classical reflection effects are neglected:

$$(7) \quad k(T) = \frac{1}{T} \frac{Z(T)^\ddagger}{Z(T)} e^{-\beta V^\ddagger},$$

where  $V^\ddagger$  is the bare potential barrier. If tunneling and non-classical reflection effects are computed by one-dimensional models (*e.g.*, parabolic or Eckart barriers), eq. (7) remains unchanged except for a so-called temperature-dependent tunneling coefficient, which multiplies its right-hand side. More refined models, including mode-mode and mode-RC couplings can be effectively introduced by employing the VPT2 expressions in the framework of the semi-classical (SC) approximation (in which quantum numbers correspond to actions and motion along the RC is governed by an imaginary action) [106, 107]. In this case, it is nearly equivalent to start from microcanonical rate constants (the TS theory in this ensemble corresponds to the Rice-Ramsperger-Kassel-Marcus (RRKM) theory [108]) and then obtain, if needed, the canonical results by means of a numerical Laplace transform. Several studies are now available reporting successful applications of the SCTST-VPT2 model to the evaluation of reaction rates for quite large systems, also employing reduced-dimensionality models in which anharmonic contributions are evaluated only for the RC and for a few number of modes strongly coupled with it [106]. Clearly, all the limitations discussed in the preceding section for VPT2 also apply to these computations, especially if one considers that RC is intrinsically a large amplitude motion. For instance, deep tunneling effects require specific extensions [106].

Reactions with vanishing (or very small) barriers (*e.g.*, formation of van der Waals complexes in the first stage of several reactions of astrochemical interest) cannot be treated in the framework of the conventional transition state theory, but require ad-hoc models ranging from the quite demanding variational extension of TST [109] to the much simpler capture theory [110]. In both cases, calculations are performed at various long-range distances of the reactants, but VTST requires evaluation of energy derivatives (in principle up to the fourth order) at each point, whereas capture theory only requires energies, which are then fitted to a simple  $1/R^6$  expression. The resulting effective potential is then used to perform capture calculations, under the assumption that each successful capture leads to the intermediate.

After the reaction rates for all the elementary steps of the studied reaction have been computed, the time evolution of the concentrations for the different species can be obtained by means of the master equation approach [111, 112]. The  $N$  species present in the overall mechanism give rise to a  $N \times N$  rate coefficient matrix  $\mathbf{K}$  representing  $N$  coupled first-order or pseudo-first-order differential equations. The first eigenvalue issuing from diagonalization of this rate matrix is equal to zero, and the corresponding eigenvector gives the equilibrium Boltzmann distribution of all the species entering the general reaction network [112].

#### 4. – Exogenous delivery: prebiotic molecules in the interstellar medium

In the early seventies, molecular synthesis through gas-phase ion-neutral reactions was proposed and successfully rationalized molecular abundances observed in interstellar clouds. Later, the importance of neutral-neutral reactions was recognized, even in low-temperature conditions. However, as observational capabilities advanced over the past decade, iCOMs began to be detected in regions where gas-phase reactions did not

contribute significantly to chemical processes. After the failure of gas-phase chemistry in explaining the detected abundances of iCOMS, the gas-grain chemistry was introduced and took over. This assumes that iCOMs are mostly synthesized on grain surfaces during the so-called warm-up phase, that is, when various radicals trapped in the grain mantles acquire mobility and can recombine into larger molecules (*e.g.*, refs [113,114]). However, recent detections of iCOMs in cold environments (see, *e.g.*, refs. [115-119]) have casted doubt on the exclusive role of grain-surface chemistry. Indeed, it is clear that gas-phase reactions play a role, in particular in cold environments (see, *e.g.*, refs. [120-122]). It is therefore well accepted that gas-phase reactions have been overlooked. In this context, formamide provides an important and compelling example. Despite its wide distribution in the universe, the mechanism by which formamide is formed is controversial and a subject of intense debate: both gas-phase (*e.g.*, [27,123-125]) and grain surface mechanisms (*e.g.*, [114]) have been proposed. However, limited are the astronomical observations able to resolve the issue [126]. Indeed, understanding the underlying chemical processes, including the production, reactions and destruction of compounds, requires the concomitant study of gas-phase reactivity and heterogeneous processes on dust-grains. It is interesting to note that most of the last dozens of species detected in space were not predicted by any chemical model, thus indicating the deficiency of current knowledge and the need of extensive work.

**4.1. Spectroscopic characterization.** – Focusing on the identification of biomolecular building blocks in the ISM, the “holy grail” is glycine, the simplest amino acid. Glycine (found in meteorites and comets) is strongly believed to be present in the ISM but has not yet been detected. However, a deep investigation of its formation routes might suggest one or more stable, simpler byproducts to be spectroscopically characterized and to be used as reliable “proxy” of glycine. To better understand this concept, let us consider the gas-phase formation route of formamide. According to [27], in addition to the latter, stable products of the suggested pathway are the Z- and E-methanimidic acid isomers. Their spectroscopic characterization (still missing) would allow for their astronomical searches in environments where formamide is thought to be present. Then, a positive detection would confirm the suggested formation pathway of formamide. The required spectroscopic characterization can be obtained by means of an integrated experiment-theory approach. High-level QC computations are performed in order to provide accurate predictions for the spectroscopic parameters, which are then used to simulate rotational spectra. These simulations are then employed to guide rotational spectroscopy measurements, whose assignment —guided by the QC simulation— finally leads to the spectroscopic data needed for supporting astronomical spectra. To exploit this interplay of experiment and theory, we rely on the VMS-ROT software [127], developed in the framework of the VMS project [76]. Figure 3 provides a schematic representation of the integrated experiment-theory approach briefly explained above.

**4.2. Gas-phase chemistry.** – The derivation of a feasible and accessible reactive potential energy surface (PES) leading to the iCOM of interest is the starting point for the

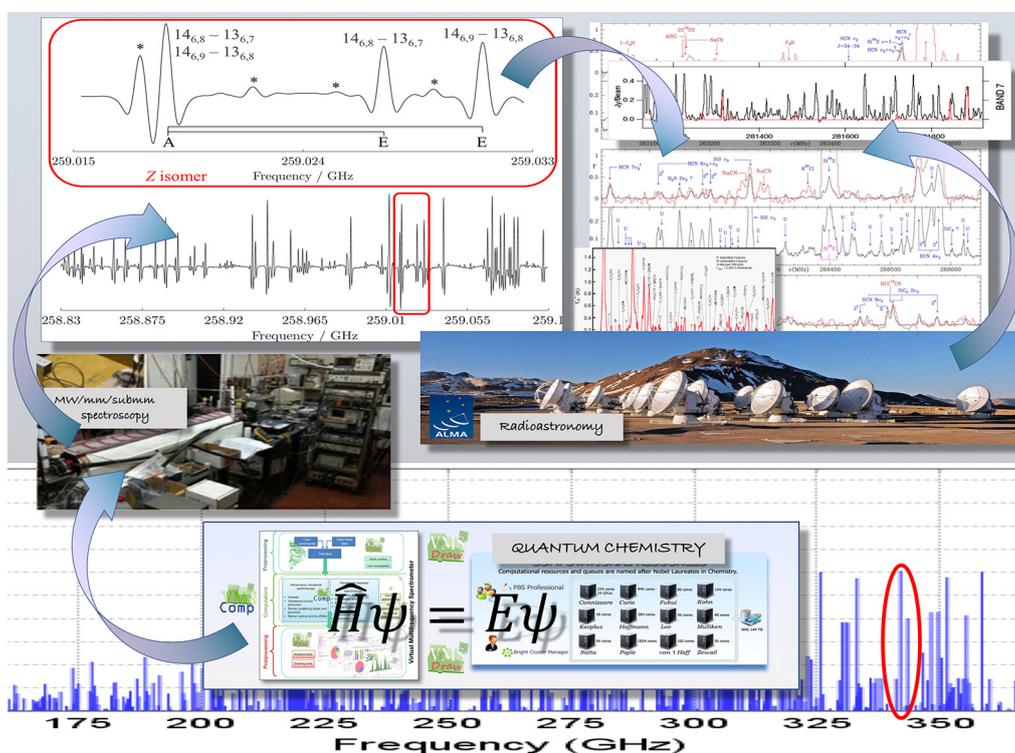


Fig. 3. – Synergism of laboratory, experimental and computational, spectroscopy applied to astronomical observations.

complete characterization of gas-phase formation pathways, with the potential precursors being selected among the molecular species already detected in space. The first step is the identification of all stationary points (minima and transition states) along the reactive PES with a cost-effective computational model. Among the different routes identified, only those that are feasible in the typical conditions of the astronomical environment under consideration will be then further investigated. Since the ISM is characterized by harsh conditions with extremely cold (down to 10 K) regions where the density is extremely low (of the order of  $10^4$  particles/cm<sup>3</sup>), accessible chemical routes are those for which all transition states lie below the energy of the reactants, that is only transition states should be present. In the following step, the selected reaction schemes are better characterized by means of an effective computational strategy, thus requiring the accurate computation of structural, energetic, and vibrational features of all the intermediates and transition states involved. In this step, we usually rely on the B2PLYP functional (also including the D3BJ empirical dispersion correction [128, 129]) in conjunction with a triple-zeta quality basis set. Indeed, this level of theory is able to provide reliable and accurate molecular structures, on top of which accurate energies are then evaluated by means of the composite schemes previously introduced. In parallel, DFT can be used

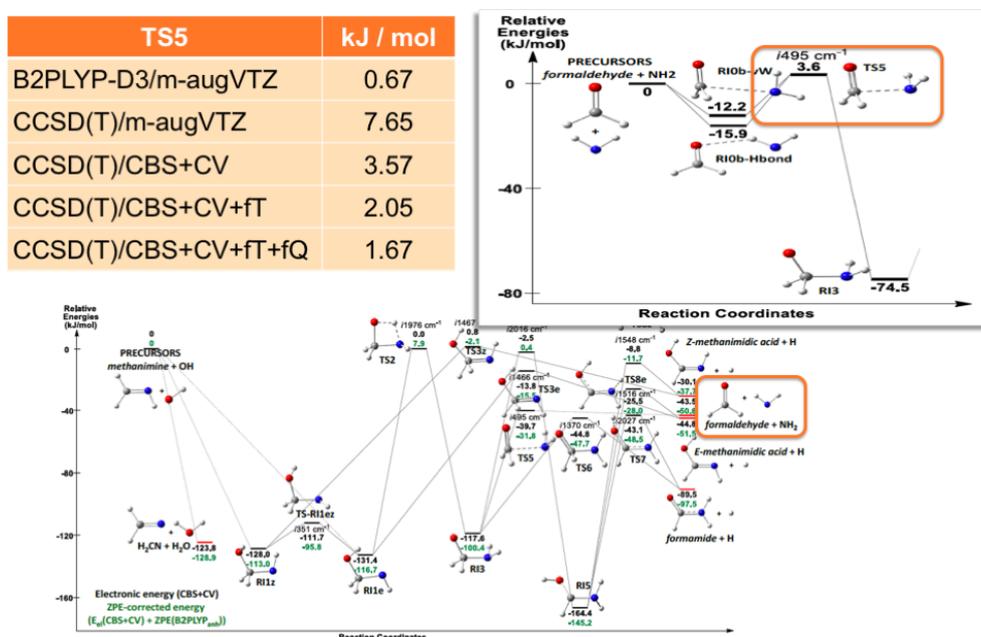


Fig. 4. – Gas-phase formation route of formamide: the impact of the level of theory on reaction barriers.

to accurately compute ZPE and anharmonic vibrational frequencies within the VPT2 approach. This strategy leads to the accurate thermochemical characterization of the formation pathway, which is then followed by accurate kinetic calculations carried out as detailed above.

If required, to further inspect the magnitude of a barrier height with respect to the energy of reactants, the effects due to a full treatment of triple and quadruple excitations can be accounted for. Figure 4, based on the results of ref. [27], shows the importance of such corrections in determining whether a barrier height leads to discard a path or not.

**4.3. Grain and gas-grain chemistry.** – In addition to gas-phase chemistry, as already mentioned, grain-surface chemistry —occurring on bare dust grains or in ice mantles— is of fundamental importance for the formation of iCOMs [130]. Dust grains —that are part of the interstellar matter— play a double role: by scattering and adsorbing interstellar radiations, they protect molecules from photodissociation and photoionization, but they are also actively involved in the production of further molecules. In fact, dust grains absorbing molecules and atoms on their surface provide the opportunity for chemical reactions to take place. In molecular clouds, the grain cores, mainly consisting of non-volatile silicates and carbonaceous compounds (also including PAHs), are often surrounded by a grain mantle, which consists of water ice containing various molecules.

As far as reactivity on grains is concerned, an additional challenge is the proper definition of the reactive system (*i.e.*, the reacting species plus the water-ice molecules directly involved) to be treated with accurate computational methods. Therefore, to address grain-surface chemistry, a “multilayer” QM/QM’ ONIOM approach (QM standing for quantum mechanics) needs to be employed [131, 132], where the more distant environment part is treated with the less accurate QM’ methods.

## 5. – A “four-pillar approach” for astrochemistry

Experiment or theory? Which one of the two traditional pillars of hard sciences should be exploited to meet a “grand challenge” like the disclosure of the secrets of chemical evolution in space? While such a dispute has characterized all scientific fields for decades, the end of the last century witnessed a reconciliation leading to the widespread employment of integrated experimental-theoretical approaches. Following on from this, fundamental questions arise: how far can such an interplay be pushed? Is it able to solve one of the most intriguing and complex scientific puzzles? Computer-based simulation is currently considered a routine facility in science and represents the third pillar of an integrated experiment/theory/simulation approach. The spontaneous question arising is: what is next? The instinctive answer is data science and machine learning (ML), the growing success of which is related to the availability of an unprecedented amount of high-quality data from a variety of sources. Thanks to the concomitant increase in computing power and the continuous optimization of ML algorithms, they currently represent the fourth pillar of an effective integrated strategy for the investigation of complex phenomena (fig. 5). At the same time, we are witnessing a parallel and fast growth of immersive virtual reality (IVR) technologies with their promise of a more complex human perception overcoming our space-time physiological limitations.

In the framework of an integrated strategy exploiting first principle (bottom up) and data driven (top down) approaches, big data analysis (BDA) tools can be used to perform a critical analysis of the results available in astrochemical kinetic and astronomical spectroscopic databases and for gas-grain network models; then ML algorithms can be employed to proceed in the selection of reactions and molecular species to be studied as well as in deriving models for the rationalization of the chemical evolution. High Performance Computing (HPC) and simulation/visualization (also IVR) facilities can be used to drive and analyze the computational results concerning both spectroscopic and kinetic aspects and to suggest new research routes. Experimental data can be, of course, integrated to computational results in heterogeneous data bases, which can be exploited by new unsupervised workflow managers.

The main possibilities offered by the application of IVR technologies to astrochemical data are: i) the ability to close the gap between human perception and molecular world, this goal being achievable by coupling visual information with proprioception skills and other senses; ii) the possibility to harness human intuition in perturbing and guiding simulations in interactive environments. Therefore, the development of powerful molecular viewers purposely tailored to astrochemical applications could be of particular inter-

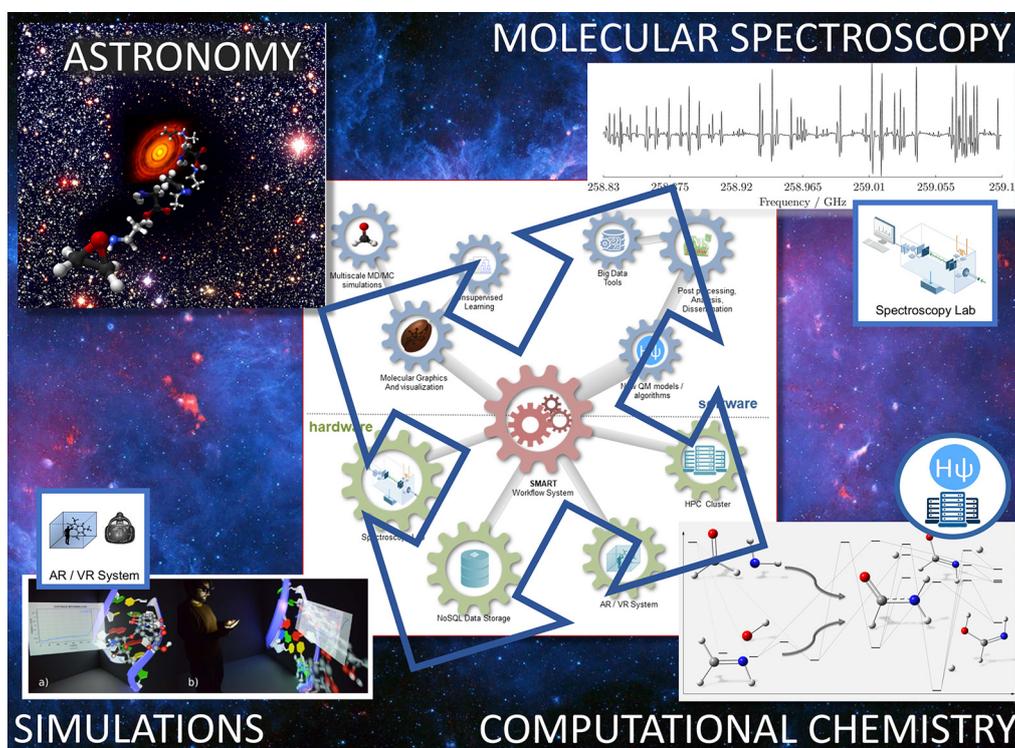


Fig. 5. – Integration of different aspects in astrochemistry: the “four-pillar approach”.

est because, when non-linear relationships and high-dimensional data sets are studied, the qualitative judgment skill unique to human investigators (the so-called “chemical intuition”) may be able to detect trends and patterns difficult to isolate by means of analytical algorithms. Since the importance of such large data sets is growing, it is clear that visualization must cope with this evolution. An example is offered by the Caffeine software, which has been developed to exploit the last-generation IVR technologies [133]. From another point of view, IVR can be effectively used for the analysis of the 4D interferometric data-cubes with the spectral catalogues in order to disentangle the emission due to a large number of prebiotic organic molecules among the forest of lines emerging from the observational data-cubes. To give just one example, a reliable identification of organic species in the neighborhood of protostellar systems requires the analysis of hundreds of spectra (extracted from the 2D interferometric images) and the successive inspection of hundreds of lines (each one with different spread over very large bandwidths). A systematic investigation is nowadays hampered by the limits imposed by analyses performed on screens. More advanced IVR hardware and software driven by powerful high-performance computers can solve the problem in a very effective way and open new exciting possibilities in several fields of astrochemistry.

## 6. – Concluding remarks

The complexity of the astrochemical challenges places demands on the theories of science, stretching the understanding of spectroscopy, kinetics and thermodynamics into areas where large non-ideal systems are hard to understand. The chemistry of the universe is a remarkably rich subject that may have a bearing on the formation of life on Earth and, maybe, elsewhere in the universe. Among the molecules discovered in the ISM, particularly fascinating are the so-called iCOMs that combine multiple functional groups and show a prebiotic character. While the complexity of the chemistry occurring in space is undisputed, it is not at all clear how to answer two fundamental questions: how were and still are these species produced? How do they evolve toward biological building-blocks? Gaining deeper insights into the underlying chemical processes, including the production, reactions and destruction of compounds can help to address some of the most intriguing questions in all scientific disciplines: how did life emerge on primitive Earth? Is it possible that it has also happened elsewhere in the Universe?

In this contribution, the great challenges posed by astrochemistry have been introduced together with the integrated approach envisaged to address them. Being an extremely vast topic, the focus has been mainly placed on the role that the different aspects of computational and theoretical chemistry can play. However, even within this narrow context, we hope to have being able to give a flavor of the fascinating world of astrochemistry.

\* \* \*

This work has been supported by MIUR “PRIN 2015” funds (Grant Number 2015F59J3R) and the Italian Space Agency (ASI; “Life in Space”, project N. 2019-3-U.0).

## REFERENCES

- [1] SHAW A. M. (Editor), *Astrochemistry* (John Wiley & Sons) 2006.
- [2] HARTMANN J., *Astrophys. J.*, **19** (1904) 268.
- [3] TRUMPLER R. J., *Publ. Astronom. Soc. Pac.*, **42** (1930) 214.
- [4] MCKELLAR A., *Publ. Astron. Soc. Pac.*, **52** (1940) 187.
- [5] DOUGLAS A. E. and HERZBERG G., *Astrophys. J.*, **94** (1941) 381.
- [6] HOLLIS J. M., VOGEL S. N., SNYDER L. E., JEWELL P. R. and LOVAS F. J., *Astrophys. J.*, **554** (2001) L81.
- [7] HOLLIS J. M., LOVAS F. J., REMIJAN A. J., JEWELL P. R., ILYUSHIN V. V. and KLEINER I., *Astrophys. J.*, **643** (2006) L25.
- [8] TERCERO B., KLEINER I., CERNICHARO J., NGUYEN H. V. L., LÓPEZ A. and CARO G. M. M., *Astrophys. J.*, **770** (2013) L13.
- [9] MCGUIRE B. A., CARROLL P. B., LOOMIS R. A., FINNERAN I. A., JEWELL P. R., REMIJAN A. J. and BLAKE G. A., *Science*, **352** (2016) 1449.
- [10] MCGUIRE B. A., *Astrophys. J. Suppl. Ser.*, **239** (2018) 17.
- [11] YAMAMOTO S. (Editor), *Introduction to Astrochemistry (Chemical Evolution from Interstellar Clouds to Star and Planet Formation)* (Springer) 2017.

- [12] BARONE V., BICZYSKO M. and PUZZARINI C., *Acc. Chem. Res.*, **48** (2015) 1413.
- [13] CALLAHAN M. P., SMITH K. E., CLEAVES H. J., RUZICKA J., STERN J. C., GLAVIN D. P., HOUSE C. H. and DWORKIN J. P., *Proc. Natl. Acad. Sci. U.S.A.*, **108** (2011) 13995.
- [14] CHYBA C. and SAGAN C., *Nature*, **355** (1992) 125.
- [15] MILLER S. L., *Science*, **117** (1953) 528.
- [16] MILLER S. L. and UREY H. C., *Science*, **130** (1959) 245.
- [17] TENNYSON J. (Editor), *Astronomical Spectroscopy* (Imperial College Press) 2005.
- [18] ZALESKI D. P., SEIFERT N. A., STEBER A. L., MUCKLE M. T., LOOMIS R. A., CORBY J. F., MARTINEZ O., 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. and PATE B. H., *Astrophys. J.*, **765** (2013) L10.
- [19] BELLOCHE A., GARROD R. T., MÜLLER H. S. P. and MENTEN K. M., *Science*, **345** (2014) 1584.
- [20] CERNICHARO J., DANIEL F., CASTRO-CARRIZO A., AGUNDEZ M., MARCELINO N., JOBLIN C., GOICOEHEA J. R. and GUÉLIN M., *Astrophys. J.*, **778** (2013) L25.
- [21] MCCALL B. J. and GRIFFIN R. E., *Proc. R. Soc. A*, **469** (2013) 20120604.
- [22] CAMPBELL E. K., HOLZ M., GERLICH D. and MAIER J. P., *Nature*, **523** (2015) 322.
- [23] SEWILO M., INDEBETOUW R., CHARNLEY S. B., ZAHORECZ S., OLIVEIRA J. M., VAN LOON J. T., WARD J. L., CHEN C.-H. R., WISEMAN J., FUKUI Y., KAWAMURA A., MEIXNER M., ONISHI T. and SCHILKE P., *Astrophys. J.*, **853** (2018) L19.
- [24] PUZZARINI C., STANTON J. F. and GAUSS J., *Int. Rev. Phys. Chem.*, **29** (2010) 273.
- [25] BICZYSKO M., BLOINO J. and PUZZARINI C., *WIREs Comput. Mol. Sci.*, **8** (2018) e1349.
- [26] VAZART F., LATOUCHE C., SKOUTERIS D., BALUCANI N. and BARONE V., *Astrophys. J.*, **810** (2015) 111.
- [27] VAZART F., CALDERINI D., PUZZARINI C., SKOUTERIS D. and BARONE V., *J. Chem. Theory Comput.*, **12** (2016) 5385.
- [28] SKOUTERIS D., BALUCANI N., CECCARELLI C., VAZART F., PUZZARINI C., BARONE V., CODELLA C. and LEFLOCH B., *Astrophys. J.*, **854** (2018) 135.
- [29] LATTELAIS M., PAUZAT F., ELLINGER Y. and CECCARELLI C., *Astrophys. J.*, **696** (2009) L133.
- [30] LATTELAIS M., PAUZAT F., ELLINGER Y. and CECCARELLI C., *Astron. Astrophys.*, **519** (2010) A30.
- [31] LOOMIS R. A., MCGUIRE B. A., SHINGLEDECKER C., JOHNSON C. H., BLAIR S., ROBERTSON A. and REMIJAN A. J., *Astrophys. J.*, **799** (2015) 34.
- [32] LOISON J.-C., AGUNDEZ M., MARCELINO N., WAKELAM V., HICKSON K. M., CERNICHARO J., GERIN M., ROUEFF E. and GULIN M., *Mon. Not. R. Astron. Soc.*, **456** (2016) 4101.
- [33] RAGHAVACHARI K., TRUCKS G. W., POPLE J. A. and HEAD-GORDON M., *Chem. Phys. Lett.*, **157** (1989) 479.
- [34] CSÁSZÁR A. G., ALLEN W. D. and SCHAEFER III H. F., *J. Chem. Phys.*, **108** (1998) 9751.
- [35] TAJTI A., SZALAY P. G., CSÁSZÁR A. G., KÁLLAY M., GAUSS J., VALEEV E. F., FLOWERS B. A., VÁZQUEZ J. and STANTON J. F., *J. Chem. Phys.*, **121** (2004) 11599.
- [36] HECKERT M., KÁLLAY M. and GAUSS J., *Mol. Phys.*, **103** (2005) 2109.
- [37] HECKERT M., KÁLLAY M., TEW D. P., KLOPPER W. and GAUSS J., *J. Chem. Phys.*, **125** (2006) 044108.
- [38] PUZZARINI C., HECKERT J. and GAUSS J., *J. Chem. Phys.*, **128** (2008) 194108.
- [39] BOESE A. D., OREN M., ATASOYLU O., MARTIN J. M. L., KÁLLAY M. and GAUSS J., *J. Chem. Phys.*, **120** (2004) 4129.

- [40] KARTON A., RABINOVICH E., MARTIN J. M. L. and RUSCIC B., *J. Chem. Phys.*, **125** (2006) 144108.
- [41] DIXON D. A., FELLER D. and PETERSON K. A., *Ann. Rep. Comput. Chem.*, **8** (2012) 1.
- [42] NOGA J. and BARTLETT R. J., *J. Chem. Phys.*, **86** (1987) 7041.
- [43] SCUSERIA G. E. and SCHAEFER III H. F., *Chem. Phys. Lett.*, **152** (1988) 382.
- [44] WATTS J. D. and BARTLETT R. J., *J. Chem. Phys.*, **93** (1990) 6104.
- [45] KÁLLAY M. and SURJÁN P. R., *J. Chem. Phys.*, **115** (2001) 2945.
- [46] PUZZARINI C., *Phys. Chem. Chem. Phys.*, **13** (2011) 21319.
- [47] BARONE V., BICZYSKO M., BLOINO J. and PUZZARINI C., *Phys. Chem. Chem. Phys.*, **15** (2013) 1358.
- [48] PUZZARINI C., PENOCCHIO E., BICZYSKO M. and BARONE V., *J. Phys. Chem. A*, **118** (2014) 6648.
- [49] PUZZARINI C., *J. Phys. Chem. A*, **119** (2015) 11614.
- [50] PUZZARINI C., BICZYSKO M., BARONE V., LARGO L., PEÑA I., CABEZAS C. and ALONSO J. L., *J. Phys. Chem. Lett.*, **5** (2014) 534.
- [51] PUZZARINI C. and BICZYSKO M., *J. Phys. Chem. A*, **119** (2015) 5386.
- [52] MØLLER C. and PLESSET M. S., *Phys. Rev.*, **46** (1934) 618.
- [53] BARONE V., BICZYSKO M., BLOINO J., CIMINO P., PENOCCHIO E. and PUZZARINI C., *J. Chem. Theory Comput.*, **11** (2015) 4342.
- [54] SPADA L., TASINATO N., VAZART F., BARONE V., CAMINATI W. and PUZZARINI C., *Chem. Eur. J.*, **23** (2017) 4876.
- [55] BARONE V., *Theor. Chim. Acta*, **91** (1995) 113.
- [56] PUZZARINI C., BICZYSKO M. and BARONE V., *J. Chem. Theory Comput.*, **6** (2010) 828.
- [57] BARONE V., BICZYSKO M., BLOINO J., EGIDI F. and PUZZARINI C., *J. Chem. Phys.*, **138** (2013) 234303.
- [58] STANTON J. F., GAUSS J., HARDING M. E. and SZALAY P. G., *CFour: a quantum chemical program package*, with contributions from AUER A. A., BARTLETT R. J., BENEDIKT U., BERGER C., BERNHOLDT D. E., BOMBLE Y. J., CHRISTIANSEN O., ENGEL F., HECKERT M., HEUN O., HUBER C., JAGAU T.-C., JONSSON D., JUSÉLIUS J., KLEIN K., LAUDERDALE W. J., LIPPARINI F., MATTHEWS D., METZROTH T., MÜCK L. A., O'NEILL D. P., PRICE D. R., PROCHNOW E., PUZZARINI C., RUUD K., SCHIFFMANN F., SCHWALBACH W., STOPKOWICZ S., TAJTI A., VÁZQUEZ J., WANG F., WATTS J. D. and the integral packages MOLECULE (ALMLÓF J. and TAYLOR P. R.), PROPS (TAYLOR P. R.), ABACUS (HELGAKER T., JENSEN H. J. AA., JØRGENSEN P. and OLSEN J.), and ECP routines by MITIN A. V. and VAN WÜLLEN C. For the current version, see <http://www.cfour.de> (2016).
- [59] PUZZARINI C., *Phys. Chem. Chem. Phys.*, **15** (2013) 6595.
- [60] THORWIRTH S., MÜCK L. A., GAUSS J., TAMASSIA F., LATTANZI V. and MCCARTHY M. C., *J. Phys. Chem. Lett.*, **2** (2011) 1228.
- [61] BARONE V., BICZYSKO M., BLOINO J. and PUZZARINI C., *Phys. Chem. Chem. Phys.*, **15** (2013) 10094.
- [62] MARTIN J. M. and TAYLOR P. R., *Chem. Phys. Lett.*, **248** (1996) 336.
- [63] DEMAISON J., MARGULÈS L. and BOGGS J. E., *Struct. Chem.*, **14** (2003) 159.
- [64] PUZZARINI C., *J. Phys. Chem. A*, **113** (2009) 14530.
- [65] PUZZARINI C. and BARONE V., *Phys. Chem. Chem. Phys.*, **13** (2011) 7189.
- [66] PUZZARINI C., BICZYSKO M., BARONE V., PEÑA I., CABEZAS C. and ALONSO J. L., *Phys. Chem. Chem. Phys.*, **15** (2013) 16965.
- [67] GRIMME S., *J. Chem. Phys.*, **124** (2006) 034108.
- [68] BICZYSKO M., PANEK P., SCALMANI G., BLOINO J. and BARONE V., *J. Chem. Theory Comput.*, **6** (2010) 2115.

- [69] BARONE V., BICZYNSKO M., BLOINO J. and PUZZARINI C., *J. Chem. Phys.*, **141** (2014) 034107.
- [70] PUZZARINI C. and BARONE V., *Acc. Chem. Res.*, **51** (2018) 548.
- [71] PUZZARINI C., TASINATO N., BLOINO J., SPADA L. and BARONE V., *Phys. Chem. Chem. Phys.*, **21** (2019) 3432.
- [72] TIELENS A. G. G. M., *Rev. Mod. Phys.*, **85** (2013) 1021.
- [73] BARONE V. (Editor), *Computational Strategies for Spectroscopy: from Small Molecules to Nano Systems* (John Wiley & Sons, Inc.) 2011.
- [74] CAZZOLI G., LATTANZI V., KIRSCH T., GAUSS J., TERCERO B., CERNICCHARO J. and PUZZARINI C., *Astron. Astrophys.*, **591** (2016) A126.
- [75] OBENCHAIN D. A., SPADA L., ALESSANDRINI S., RAMPINO S., HERBERS S., TASINATO N., MENDOLICCHIO M., KRAUS P., GAUSS J., PUZZARINI C., GRABOW J.-U. and BARONE V., *Angew. Chem. Int. Ed.*, **57** (2018) 15822.
- [76] BARONE V., *WIREs Comput. Mol. Sci.*, **6** (2016) 86.
- [77] FORTENBERRY R. C., *Int. J. Quantum Chem.*, **117** (2017) 81.
- [78] FORNARO T., BRUCATO J. R., FEULLIE C., SVERJENSKY D. A., HAZEN R. M., BRUNETTO R., D'AMORE M. and BARONE V., *Astrobiology*, **18** (2018) 989.
- [79] NAJBAUER E. E., BAZSO G., APOSTOLO R., FAUSTO R., BICZYNSKO M., BARONE V. and TARCZAY G., *J. Phys. Chem. B*, **119** (2015) 10496.
- [80] BARONE V., BELLINA F., BICZYNSKO M., BLOINO J., FORNARO T., LATOUCHE C., LESSI M., MARIANETTI G., MINEI P., PANATTONI A. and PUCCI A., *Phys. Chem. Chem. Phys.*, **17** (2015) 26710.
- [81] MUNIZ-MIRANDA F., PEDONE A., BATTISTELLI G., MONTALTI M., BLOINO J. and BARONE V., *J. Chem. Theory Comput.*, **11** (2015) 5371.
- [82] PATTI A., PEDOTTI S., MAZZEO G., LONGHI G., ABBATE S., PAOLONI L., BLOINO J., RAMPINO S. and BARONE V., *Phys. Chem. Chem. Phys.*, **21** (2019) 9419.
- [83] ALIEV M. R. and WATSON J. K. G., *Higher-order effects in the vibration-rotation spectra of semirigid molecules*, in *Molecular Spectroscopy: Modern Research*, edited by RAO K. N., Vol. **3** (Academic Press) 1985, Chapt. 1, pp. 1-67.
- [84] GORDY W. and COOK R. L., *Microwave Molecular Spectra*, 3rd edition (Wiley, New York) 1984.
- [85] WATSON J. K. G., *J. Chem. Phys.*, **48** (1968) 4517.
- [86] ALESSANDRINI S., GAUSS J. and PUZZARINI C., *J. Chem. Theory Comput.*, **14** (2018) 5360.
- [87] BECKE A. D., *J. Chem. Phys.*, **98** (1993) 5648.
- [88] LEE C., YANG W. and PARR R. G., *Phys. Rev. B*, **37** (1988) 785.
- [89] Double and triple- $\zeta$  basis sets of the SNS family are available in the download section, URL: <http://smart.sns.it/> (last visited: 3 February 2019).
- [90] PICCARDO M., PENOCCHIO E., PUZZARINI C., BICZYNSKO M. and BARONE V., *J. Phys. Chem. A*, **119** (2015) 2058.
- [91] PUZZARINI C., CAZZOLI G. and GAUSS J., *J. Chem. Phys.*, **137** (2012) 154311.
- [92] PUZZARINI C., BICZYNSKO M., BLOINO J. and BARONE V., *Astrophys. J.*, **785** (2014) 107.
- [93] MELLI A., MELOSSO M., TASINATO N., BOSI G., SPADA L., BLOINO J., MENDOLICCHIO M., DORE L., BARONE V. and PUZZARINI C., *Astrophys. J.*, **855** (2018) 123.
- [94] WATSON J. K., *Mol. Phys.*, **15** (1968) 479.
- [95] NIELSEN H. H., *Rev. Mod. Phys.*, **23** (1951) 90.
- [96] PUZZARINI C., BLOINO J., TASINATO N. and BARONE V., *Chem. Rev.*, **119** (2019) 8131.
- [97] BARONE V., BLOINO J., GUIDO C. and LIPPARINI F., *Chem. Phys. Lett.*, **496** (2010) 157.

- [98] BLOINO J., BICZYNSKO M. and BARONE V., *J. Phys. Chem. A*, **119** (2015) 11862.
- [99] BAIARDI A., BLOINO J. and BARONE V., *J. Chem. Theory Comput.*, **13** (2017) 2804.
- [100] BAIARDI A., BLOINO J. and BARONE V., *J. Chem. Theory Comput.*, **11** (2015) 3267.
- [101] KUHLER K. M., TRUHLAR D. G. and ISAACSON A. D., *J. Chem. Phys.*, **104** (1996) 4664.
- [102] BLOINO J., BICZYNSKO M. and BARONE V., *J. Chem. Theory Comput.*, **8** (2012) 1015.
- [103] SCHURMAN M. S., ALLEN W. D., VON RAGUÉ SCHLEYER P. and SCHAEFER III H. F., *J. Chem. Phys.*, **122** (2005) 104302.
- [104] PUZZARINI C., BICZYNSKO M. and BARONE V., *J. Chem. Theory Comput.*, **7** (2011) 3702.
- [105] TRUHLAR D. G. and ISAACSON A. D., *J. Chem. Phys.*, **94** (1991) 357.
- [106] SHAN X., BURD T. A. H. and CLARY D. C., *J. Phys. Chem. A*, **123** (2019) 4639.
- [107] HERNANDEZ R. and MILLER W. H., *Chem. Phys. Lett.*, **214** (1993) 129.
- [108] STEINFELD J. I., FRANCISCO J. S. and HASE W. L. (Editors), *Chemical Kinetics and Dynamics* (Prentice Hall) 1998.
- [109] BAO J. L. and TRUHLAR D. G., *Chem. Soc. Rev.*, **46** (2017) 7548.
- [110] WESTON R. E. and SCHWARTZ H. E. (Editors), *Chemical Kinetics* (Prentice-Hall Inc., Englewood Cliffs, NJ) 1972.
- [111] MILLER J. A. and KLIPPENSTEIN S. J., *J. Phys. Chem. A*, **110** (2006) 10528.
- [112] ROBERTSON S. H., PILLING M. J., JITARIU L. C. and HILLIER I. H., *Phys. Chem. Chem. Phys.*, **9** (2007) 4085.
- [113] GARROD R. T. and HERBST E., *Astron. Astrophys.*, **457** (2006) 927.
- [114] GARROD R. T., WEAVER S. L. W. and HERBST E., *Astrophys. J.*, **682** (2008) 283.
- [115] BACMANN A., TAQUET V., FAURE A., KAHANE C. and CECCARELLI C., *Astron. Astrophys.*, **541** (2012) L12.
- [116] CERNICHARO J., MARCELINO N., ROUEFF E., GERIN M., JIMÉNEZ-ESCOBAR A. and CARO G. M. M., *Astrophys. J.*, **759** (2012) L43.
- [117] JABER A. A., CECCARELLI C., KAHANE C. and CAUX E., *Astrophys. J.*, **791** (2014) 29.
- [118] VASTEL C., CECCARELLI C., LEFLOCH B. and BACHILLER R., *Astrophys. J.*, **795** (2014) L2.
- [119] JIMÉNEZ-SERRA I., VASYUNIN A. I., CASELLI P., MARCELINO N., BILLOT N., VITI S., TESTI L., VASTEL C., LEFLOCH B. and BACHILLER R., *Astrophys. J.*, **830** (2016) L6.
- [120] VASYUNIN A. I. and HERBST E., *Astrophys. J.*, **769** (2013) 34.
- [121] BALUCANI N., CECCARELLI C. and TAQUET V., *Mon. Not. R. Astron. Soc.*, **449** (2015) L16.
- [122] VASYUNIN A. I., CASELLI P., DULIEU F. and JIMÉNEZ-SERRA I., *Astrophys. J.*, **842** (2017) 33.
- [123] KAHANE C., CECCARELLI C., FAURE A. and CAUX E., *Astrophys. J.*, **763** (2013) L38.
- [124] BARONE V., LATOUCHE C., SKOUTERIS D., VAZART F., BALUCANI N., CECCARELLI C. and LEFLOCH B., *Mon. Not. R. Astron. Soc. Lett.*, **453** (2015) L31.
- [125] SKOUTERIS D., VAZART F., CECCARELLI C., BALUCANI N., PUZZARINI C. and BARONE V., *Mon. Not. R. Astron. Soc. Lett.*, **468** (2017) L1.
- [126] 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. I., 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., QUENARD D., RIMOLA A., SAKAI N., SKOUTERIS D., TAQUET V., TESTI L., THEULÉ P., UGLIENGO P., VASTEL C., VAZART F., WIESENFELD L. and YAMAMOTO S., *Astron. Astrophys.*, **605** (2017) L3.

- [127] LICARI D., TASINATO N., SPADA L., PUZZARINI C. and BARONE V., *J. Chem. Theory Comput.*, **13** (2017) 4382.
- [128] GRIMME S., ANTONY J., EHRLICH S. and KRIEG H., *J. Chem. Phys.*, **132** (2010) 154104.
- [129] GRIMME S., *WIREs Comput. Mol. Sci.*, **1** (2011) 211.
- [130] ARUMAINAYAGAM C. R., GARROD R. T., BOYER M. C., HAY A. K., BAO S. T., CAMPBELL J. S., WANG J., NOWAK C. M., ARUMAINAYAGAM M. R. and HODGE P. J., *Chem. Soc. Rev.*, **48** (2019) 2293.
- [131] RE S. and MOROKUMA K., *J. Phys. Chem. A*, **105** (2001) 7185.
- [132] CHUNG L. W., SAMEERA W. M. C., RAMOZZI R., PAGE A. J., HATANAKA M., PETROVA G. P., HARRIS T. V., LI X., KE Z., LIU F., LI H.-B., DING L. and MOROKUMA K., *Chem. Rev.*, **115** (2015) 5678.
- [133] SALVADORI A., DEL FRATE G., PAGLIAI M., MANCINI G. and BARONE V., *Int. J. Quantum Chem.*, **116** (2016) 1731.