Puzzarini C. & Barone, V. (2020). A never-ending story in the sky: The secrets of chemical evolution. Physics of Life Reviews, 32, pp. 59-94. doi: 10.1016/j.plrev.2019.07.001

Review

A never-ending story in the sky: The secrets of chemical evolution

Cristina Puzzarini [a], Vincenzo Barone [b]

[a] *Dipartimento di Chimica "Giacomo Ciamician", University of Bologna, Via F. Selmi 2, I-40126 Bologna, Italy*
[b] *Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*

Abstract

Cosmic evolution is the tale of progressive transition from simplicity to complexity. The newborn universe started with the simplest atoms formed after the Big Bang and proceeded toward the formation of the so-called 'astronomical complex organic molecules' (aCOMs), most of them showing a clear prebiotic character. Understanding the chemical evolution of the universe is one of the main aims of Astrochemistry, with the starting point being the knowledge whether a molecule is present in the astronomical environment under consideration and, if so, its abundance. However, the interpretation of astronomical detections and the identification of molecules are not at all straightforward. Indeed, the extraterrestrial chemical inventory has been obtained by means of astronomical observations based on spectroscopic signatures determined in laboratory (either experimental or computational) studies. Even though the presence of aCOMs has been known for decades, the processes that lead to their synthesis are still hotly debated or even unknown. It is often assumed that aCOMs are mostly synthesized on grain surfaces during the so-called warm-up phase, when various radicals trapped in the grain mantles acquire mobility and recombine into large molecules. However, recent detections of aCOMs in cold environments have challenged this exclusive role of grain-surface chemistry. Clearly, gas-phase chemistry is at work in cold environments. Moving to Titan's atmosphere, prior to the Cassini-Huygens arrival in the Saturn system, it was generally believed that Earth and interstellar space are the two places where organic molecules are/were synthesized extensively. However, the experimental measurements by the instruments on board the Cassini orbiter spacecraft and the Huygens probe lander have changed this view.

To disclose the "secrets" of chemical evolution across space, the first step is the understanding of how small prebiotic species are formed and how the chemical complexity can further increase. This review indeed addresses the chemical evolution in space, focusing – in particular – on the role played by molecular spectroscopy and quantum-chemical computations. To summarize, in this review we will first of all present how the signatures of molecules can be found in space. Then, we will address, from a computational point of view, the derivation of the molecular spectroscopic features, the investigation of gas-phase formation routes of prebiotic species in the ISM, and the evolution of chemical complexity, from small molecules to haze, in Titan's atmosphere. Finally, an integrated strategy, also involving high-performance computers and virtual reality, will be discussed.

1. Introduction

Particularly fascinating amongst the molecules discovered in the interstellar medium (ISM) are the so-called 'astronomical complex organic molecules' (aCOMs), such as glycoaldehyde [1], acetamide [2] and methyl acetate [3], which present an unusual complexity (in consideration of the environment) that combines multiple functional groups. Probably, even more fascinating is the recent discovery of the first chiral molecule in space, methyloxirane (also known as propylene oxide [4]), here recalling that chirality is the signature of life (at least, as we know it), and of the first aromatic ring in dark clouds, benzonitrile [5], which may be related to the presence of polycyclic aromatic hydrocarbons in the external UV-irradiated regions of clouds. Currently, the census of interstellar molecular species exceeds two hundred [6–8], and the rate of discovery continues at a rapid pace. While it is now clear that there is a diverse and complex chemistry in the ISM, it is not at all clear how these aCOMs were produced. Furthermore, since many of the molecules found in the ISM play a role in the chemistry of life, the question of molecular genesis in the ISM might be related to that of the origin of life itself. This aspect has spurred the search for biomolecular building blocks in interstellar space, with glycine – the simplest amino acid – becoming perhaps the most obvious starting objective (e.g. Refs. [9,10]). Intriguingly, evidence has been presented to suggest that RNA/DNA nucleobases and amino acids have been carried through space on meteorites and exist in other extraterrestrial sources [11,12].

Two alternative theories have been suggested so far on the emergence of life on Earth [13]: (1) exogenous delivery and (2) endogenous synthesis. In the framework of the first theory, prebiotic molecules are postulated to have come from space on comets, asteroids and meteorites. The rationale behind this suggestion is that prebiotic aCOMs have been observed in interstellar clouds, including star-forming regions. The basic idea is that prebiotic molecules were formed in the solar nebula, preserved during the early phases of the solar system formation in the body of comets, asteroids, and meteorites, and finally delivered to Earth by cometary and meteoritic impacts. In the words of Nobel laureate Christian De Duve: "The building blocks of life form naturally in our Galaxy and, most likely, also elsewhere in the cosmos. The chemical seeds of life are universal". In order to establish the veracity of this theory, it is necessary to not only discover prebiotic species in space, but also to understand how they could have been produced in the typically unfavorable conditions (extremely low temperature and density) of the ISM. According to the endogenous theory, 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 Urey-Miller experiment [14,15] revealed the perhaps shocking plausibility of this idea, 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 a laboratory proxy for atmospheric lightning and other energy sources. However, it was later recognized that the complexity of a planet cannot be reproduced in a single laboratory experiment. In the latter context, Titan – the largest moon of Saturn – has been postulated as a model of primitive Earth. Extrapolating, the organic chemistry and the increasing chemical complexity discovered in Titan's atmosphere could be a rich source of information directly relevant to prebiotic organic synthesis in the atmosphere of primitive Earth [16]. Significant in this context is the recent statement of Nobel laureate George Olah [17]: "Of particular interest to us is the remarkable detection of varied carbocations and their similarity with their terrestrial analogues. The proven similarity with our terrestrial studied chemistry provides the first scientific evidence that our Earth is not a unique celestial body for producing the chemical building blocks."

In any event, whether they were delivered to Earth from space or synthetized from simpler building blocks, prebiotic molecules then evolved to form more complex molecules. An important example is formamide (H_2NCHO), a ubiquitous molecule that has been detected in many sources throughout the universe. Formamide is a central compound to connect metabolism (conversion of energy), ruled by proteins, and genetics (passage of information), ruled by DNA and RNA [18]. Formamide can indeed polymerize through biocatalyzed processes to provide all five nucleobases of DNA and RNA as well as carboxylic and amino acids [18,19]. Rightly then, formamide has become one of the most intensively studied prebiotic compounds that are potentially relevant for the origin of life. While the evidence for molecular complexity in the universe is undisputed, there is still much to be understood about the formation of molecules in the typically cold and (largely) collision free environment of the ISM. A similar – or even deeper –

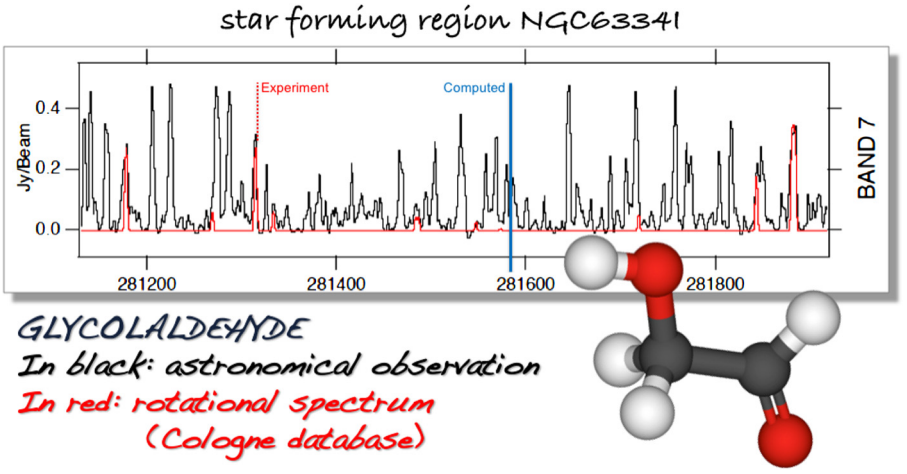

Fig. 1. Simulated spectra of glycolaldehyde ($HC(O)CH_2OH$) plotted in red (data from the Cologne database) over ALMA observations (band 7) of NGC 6334I MM1 plotted in black [29].

lack of information characterizes the chemistry in Titan's atmosphere. Indeed, little is known about how the chemical complexity there evolves [16].

2. Astronomical spectroscopy: from MW to infrared and beyond

In the field of astrochemistry, molecular spectroscopy plays a central role (see e.g. Ref. [20] and references therein). Because of the tremendous distances involved, there is no chance to do direct experiments on astrochemical processes, and detection via interaction of molecules with radiation is the only viable route of investigation [21,22]. Indeed, the astronomical observation of the spectroscopic features of a given molecule provides the definitive, unequivocal proof of its presence in the astronomical environment under consideration (see e.g. Refs. [4,23,24]).

Since the detection of the first molecule in the ISM, namely, methylidyne (CH) [25], the advent of radio astronomy (in the early 1960s) enabled the detection of new molecules with a trend that has continued at a nearly linear rate. However, despite the remarkably steady pace of new molecular detections, the number of unidentified spectral features is still impressive. While most of the detected molecules have been discovered thanks their rotational spectroscopy features, "mysterious unrecognized features" extend across the electromagnetic spectrum. To detail this issue, in the radio regime, broadband spectral line surveys continue to reveal hundreds of features not assignable to transitions of already characterized molecules [26], even if one accounts for transitions of vibrationally excited states or rare isotopic species. The situation has been even worsening in the last decade because of the increased sensitivity of current radiotelescopes (with ALMA being the most powerful tool), thus revealing new spectroscopic features in the so-called confusion limit. Moving to the IR region of the electromagnetic spectrum, the so-denoted unidentified infrared emission bands (UIRs) still continue to elude conclusive molecular identification. The same applies to sharp, distinct emission features ubiquitous in ultraviolet (UV)-irradiated regions in our galaxy and external galaxies. Despite the fact that these emission features are nowadays recognized to be due to polycyclic aromatic compounds (PAHs), no individual molecule has been definitely identified from the analysis of UIR features. Analogous is the case of the diffuse interstellar bands (DIBs) [27]; indeed, these sharp features ranging from the IR to the UV remain nearly completely unassigned, with the only undisputed assignment being that of C_{60}^+ [28].

Radioastronomical observations are those ranging from centimeter-wave to far-infrared. The features observed in this frequency region are those due the rotational motion of molecules. Thanks to the high sensitivity of rotational spectroscopy to molecular structures and isotopic masses, this is the most suitable technique for the identification of molecular species in the gas phase. Fig. 1 provides an example of the assignment of a radioastronomical spectrum: the specific case is the detection of glycolaldehyde in the NGC 6334I star forming region [29]. This example demonstrates the typical complexity of the broadband surveys mentioned above, also highlighting the accuracy requirements in the knowledge of rotational transition frequencies. Despite the fact that state-of-the-art quantum-chemical computations (which will be discussed in the next section) are able to provide predictions for rotational transitions with an accuracy

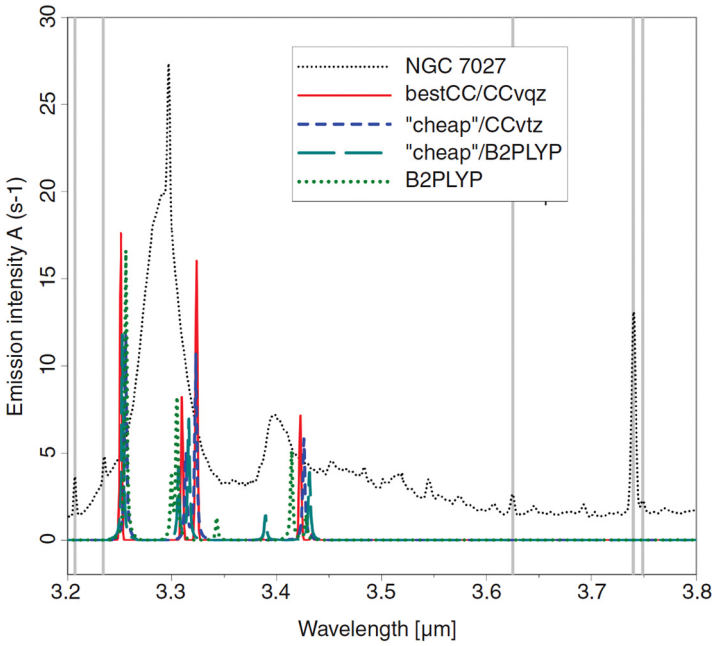

Fig. 2. Comparison between the simulated (color) spectra of oxirane and the observed (black) emission spectra of planetary nebula NGC 7027 [30].

better than 0.1%, this is not sufficient for guiding astronomical searches and/or assignments, thus requiring experimental determinations of the corresponding spectroscopic parameters. Indeed, when dealing with surveys like those in Fig. 1, transition frequencies require to be known with an accuracy possibly better than 100 kHz. While such an accuracy can be easily obtained from experimental studies, quantum-chemical calculations are far from reaching this goal.

Moving to IR astronomy, since the features lying in the region of the electromagnetic spectrum usually involve transitions between vibrational energy levels and since these are substantially higher in energy than pure rotational transitions, the observation of emission from vibrational transitions usually requires exceptionally warm regions or pumping by an external, enhanced UV radiation field. However, this branch of astronomy is a fundamental tool, particularly for the identification of those molecules lacking of permanent dipole moments. Furthermore, IR astronomy allows for observing molecules condensed into interstellar ices, which cannot undergo free rotation. However, the vibrational motions are affected by the condensed-phase environment, which can alter the frequencies of vibrational modes. In this respect, theoretical investigations, possibly supported by selected laboratory experiments, permit to understand how the band frequencies of interest modify because of the presence of other species.

As an example of IR astronomy applied to the identification of molecular species, Fig. 2 shows the comparison between the simulated emission spectra of oxirane (using different levels of theory that will be detailed in the following section) in the 3.2-3.8 μm range (3125-2632 cm^{-1}) [31] and the astronomical observations of the planetary nebula NGC 7027, as reported in a uniform database (2.4-45.4 μm spectra) from the Infrared Space Observatory Short Wavelength Spectrometer (ISO-SWS) [30]. In Fig. 2 all identified line fluxes are marked by vertical gray lines, thus clearly showing the presence of several, still unidentified UIR transitions. We note that the group of oxirane transitions related to the CH_2 stretching fundamental vibrations match well the observed UIR feature at 3.3 μm and that the one at 3.4 μm is well represented by the strong $\nu_2 + \nu_{10}$ combination band. It should be stressed that the quantum-chemical predictions reported in Fig. 2 are reliable and have an accuracy of a few wavenumbers. As will be explained later (theoretical and computational details concerning vibrational spectroscopy will be addressed in the next section), in order to fulfill such an accuracy, fully anharmonic calculations are required. It is furthermore noted that all models, even the least computationally demanding (i.e. that based on the double-hybrid B2PLYP functional [32]), provide very similar results. Nevertheless, this good match cannot be conclusive for the detection of oxirane because the broad features in the astronomical spectrum are most likely resulting from CH stretching vibrations that are present in several aCOMs.

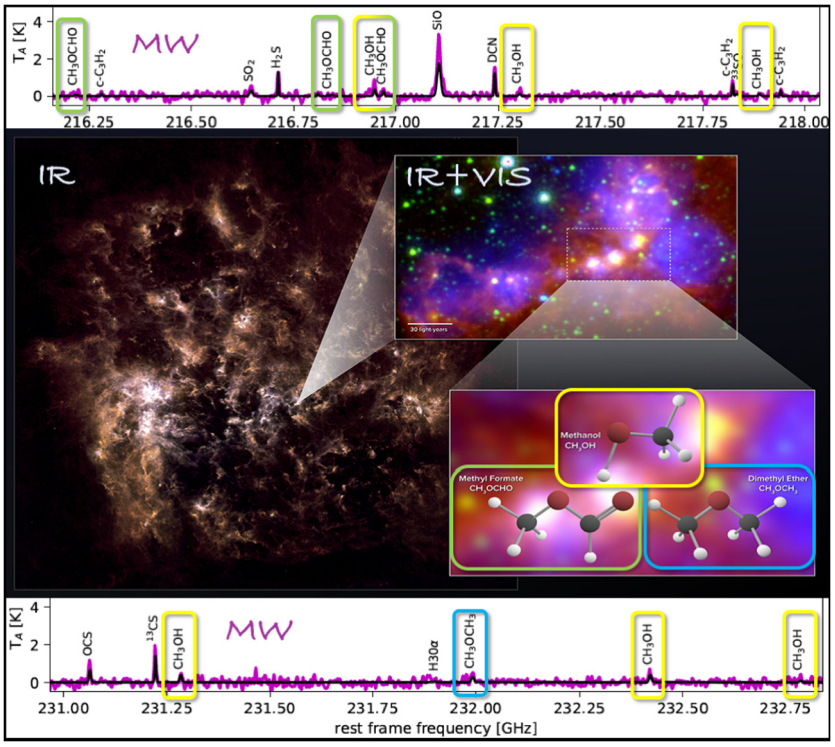

Fig. 3. The Large Magellanic Cloud observed in different regions of the electromagnetic spectrum: on the left, the far-infrared image shows the full galaxy. The zoom-in images on the right are a combination of mid-infrared data from Spitzer and visible (H-alpha) data from the Blanco 4-meter telescope. On the bottom and top panels two Band 6 spectral windows of ALMA are shown: the rotational spectroscopic fingerprints of methanol, dimethyl ether, and methyl formate observed are highlighted [35].

As far as visible and UV astronomy is concerned, as in the case of IR astronomy, no permanent dipole moment is required, thus enabling the detection of species otherwise blind to radioastronomy (see e.g. Ref. [33]). Because of the energy requirements for driving electronic transitions, these typically occur only in extremely energetic environments. To give an example of the role played in discovering new molecules, the near UV fluorescence bands that appear in the spectra of 1P/Halley's inner coma allowed the unequivocally identification of anthracene [34]. Indeed, the cometary spectral features were found to be consistent with the laboratory fluorescence spectrum of anthracene, with a very good agreement observed for the four main peaks at 363, 367.5, 373 and 382.5 nm. In addition to the relevance of the unequivocal identification of a small PAH in a comet environment, this result suggests that comets can help trace the ISM input into the primitive solar nebula.

Fig. 3 shows a nice combination of different astronomical 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 [35]. Fig. 3 therefore gives a flavor of the interplay of different astronomical techniques.

The last topic touched here is related to chiral molecules. As already mentioned, the only chiral molecule detected so far is propylene oxide (see Fig. 4), which was observed in the gas phase in a cold, extended molecular shell around the embedded, massive protostellar clusters in the Sagittarius B2 star-forming region [4]. However, that detection was not able to determine whether a racemic mixture or an enantiomeric excess was present. The most straightforward method to disclose this puzzle is circular dichroism (CD). The preferential absorption of left- or right-handed circularly polarized light (CPL) manifests itself as a change in the difference in the electric field of left versus right-handed CPL. This therefore requires polarization-sensitive astronomical observations. Modern radiotelescopes are capable of highly-accurate, polarization-sensitive observations across wide frequency windows, simultaneously determining the polarization state at each observed frequency. However, for astronomical observations, it is particularly critical to distinguish the observable polarization effects due to CD from other possible effects.

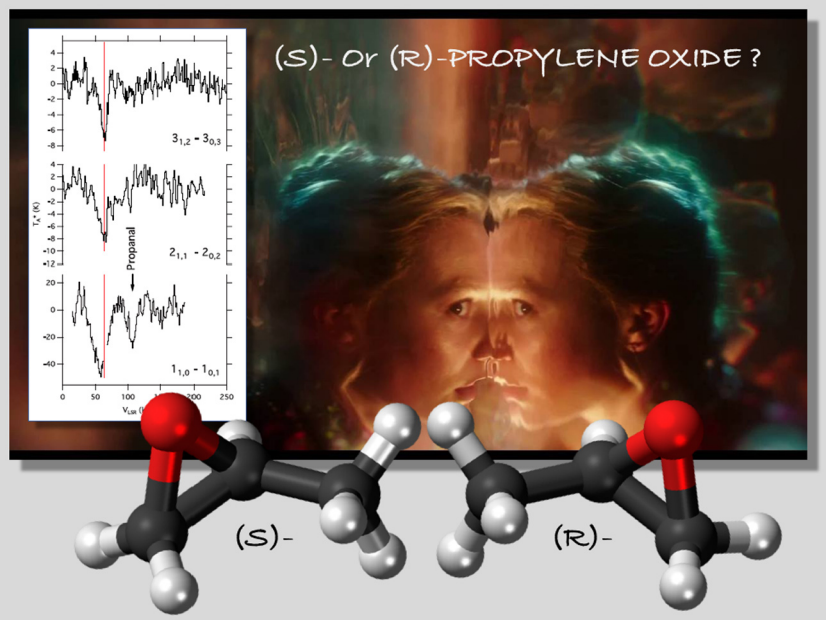

Fig. 4. Observations of the $J = 1_{1,0} \leftarrow 1_{0,1}$, $2_{1,1} \leftarrow 2_{0,2}$, and $3_{1,2} \leftarrow 3_{0,3}$ rotational transitions of propylene oxide, in absorption, toward the Galactic center [4]. The uncertainty about the chiral configuration is highlighted.

3. Computational approaches for astrochemistry: from spectroscopic features to formation pathways

As mentioned above, molecular spectroscopy provides the means for detecting molecules in space. Furthermore, because of difficulties in mimicking the extreme conditions that characterize the ISM (but also planetary atmospheres) in the laboratory, accurate state-of-the-art computational approaches play a fundamental role in analyzing feasible reaction mechanisms. At the same time, modern quantum chemistry is an extremely powerful tool that can assist experiment by providing accurate predictions of spectroscopic parameters (e.g. Refs. [20,36,37]) and can also accurately predict the kinetics associated with postulated mechanisms for reactive systems (e.g. Refs. [38–40]).

3.1. Quantum-chemical predictions of relative energies

Organic chemistry in space is nowadays a matter of fact; however, its understanding is still a great challenge for the scientific community. Currently, several C, N and O bearing molecules have been discovered in space, in different regions. In some cases, different isomers of the same molecules 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 more thermodynamically stable molecular isomers are those more favored in the chemical processes occurring [41,42]. This is a first important motivation spurring computational chemists to provide accurate predictions of relative energies for different isomeric forms. This involves state-of-the-art quantum-chemical computations to be performed. However, the role played by accurate energy calculations is not limited to this. Indeed, in order to support molecular spectroscopy, rigorous approaches for the resolution of the nuclear problem should be combined with state-of-the-art electronic structure computations. Furthermore, it should be kept in mind that, in order to investigate reaction kinetics, very accurate energy calculations are required because, at the typical low temperatures of space, rates are exquisitely sensitive to energetics and kinetic barrier heights. For all these reasons, the methodology for the accurate determination of energetics and molecular structures is presented in the following.

To obtain high accuracy in quantum-chemical calculations, it is necessary to reduce as much as possible the errors due to the truncation of both basis set and wavefunction, the so-called one- and N-electron errors, respectively. Focusing on systems that do not have multi-configurational character (i.e., well described by a single-reference wave function), in the last two decades, 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)) [43] has become the

"gold standard" for accurate computations. However, while CCSD(T) provides a good description of the electronic wavefunction, other important contributions – like, 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 putting them together, are required to reach high accuracy (see, e.g. Refs. [44–51]). To recover the errors due to the truncation of both basis set and wavefunction as well as to include other minor contributions that are important when aiming at high accuracy, the composite scheme implemented in the quantum-chemistry program package CFOUR [52], which is at the basis of the so-called HEAT approach [45] for energetics and the so-called "gradient" scheme [46,47] for equilibrium structure determinations, is mainly used in the applications presented in this review. In this scheme the starting point is the CCSD(T) method within the frozen-core approximation. Then, the first important corrections to include 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. Furthermore, the full CC singles, doubles and triples (CCSDT) [53–55] and CC singles, doubles, triples, quadruples (CCSDTQ) [56] models can be 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-ζ quality).

Focusing on equilibrium structure determinations, within the so-called "gradient" scheme, the energy gradient to be minimized in the geometry optimization procedure is set up according to the specific requirements, the target accuracy, and the dimension of the system under consideration. The energy gradient corresponding to what presented above is given by the following expression:

$$\frac{dE}{dx} = \frac{dE^{\infty}(SCF)}{dx} + \frac{d\Delta E^{\infty}(CCSD(T))}{dx} + \frac{d\Delta E(CV)}{dx} + \frac{d\Delta E(fT)}{dx} + \frac{d\Delta E(fQ)}{dx} \,, \tag{1}$$

with the correlation-consistent polarized cc-p(C)VnZ basis sets [57–60] being usually employed in conjunction with this scheme. The first two terms on the right-hand side denote the extrapolation to the CBS limit: Hartree-Fock self consistent field (HF-SCF) extrapolated to the CBS limit according to the three point formula by Feller [61] and CCSD(T) correlation energy extrapolated using the n^{-3} formula [62], two different expressions being employed because of the different convergence behavior. From the available literature (see, for example, Refs. [46,47,63–67] and references therein), the accuracy obtainable with this scheme is better than 0.001 Å for bond lengths and 0.05 degrees for angles. However, it should be noted that inclusion of fT and fQ corrections makes this approach computationally very expensive. On the other hand, if these contributions are ignored, thus leading to the so-called CCSD(T)/CBS+CV approach, a composite scheme affordable also for medium-sized molecules, characterized by a very limited loss of accuracy, is obtained. According to the available literature (see, for example, Refs. [36,46,47,67,68]), the corresponding geometrical parameters are proven to have an accuracy of 0.001–0.002 Å for distances and 0.05–0.1 degrees for angles.

Moving to energetics, the HEAT protocol [45,69,70] allows for achieving high accuracy for thermochemistry, indeed being able to reach the so-called sub-kJ accuracy. The scheme is analogous to that of Eq. (1), also accounting for first-order spin-orbit coupling, diagonal Born-Oppenheimer correction, and scalar relativistic effects. As above, the computational cost can be largely reduced by retaining only the extrapolation to the CBS limit at the CCSD(T) level and the incorporation of the CV corrections. The corresponding scheme (CCSD(T)/CBS+CV) is rather well tested in the literature (see, e.g., Refs. [65,71–73]) and is demonstrated to provide results with an accuracy well within 0.5 kcal/mol. To extend the applicability of composite schemes to larger molecules, an effective solution is provided by the so-called "cheap" approach [74,75], which is obtained by using Møller-Plesset to second-order theory (MP2) [76] 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:

$$E_{cheap} = E(CCSD(T)/VTZ) + \Delta E(MP2/CBS) + \Delta E(MP2/CV) \,, \tag{2}$$

with the second term on the right-hand side accounting for the extrapolation to the CBS limit as in the composite scheme presented above by extrapolating to the CBS limit the HF-SCF and MP2 correlation energies separately. Despite the remarkable reduction of the computational cost, this approach is able to provide accurate results (see, e.g., Refs. [67,74]), also for flexible systems [74] and molecular complexes [77].

By directly applying the additivity approximation to geometrical parameters based on the assumption that they show the same behavior as the energy (see, for example, Refs. [78–81]), the so-called "cheap" composite scheme can

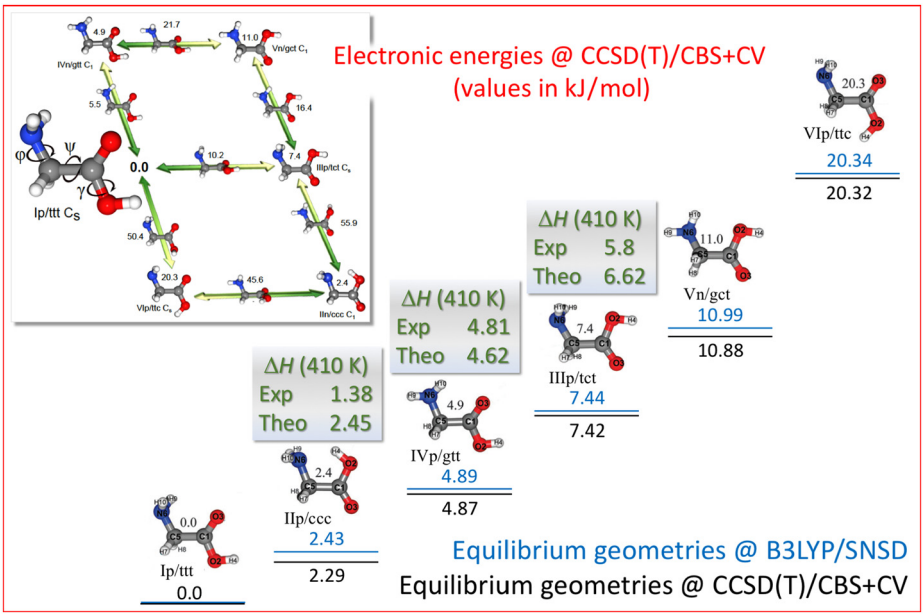

Fig. 5. CCSD(T)/CBS+CV electronic energies of glycine conformers with respect to Ip/ttt either computed the B3LYP/SNSD optimized geometries (in blues) or at the CCSD(T)/CBS+CV optimized geometries (in black). For selected conformers (those for which experimental data are available), enthalpy differences are reported (in green).

also be employed for accurate structural determinations of systems that are too large for composite approaches entirely based on CCSD(T) calculations [68,74,81,82]. To this aim, an important point is the evaluation of the CBS limit by applying the consolidated n^{-3} extrapolation form [62] to the whole parameters without separating the HF-SCF and correlation contributions, the suitability of this approach being demonstrated in Ref. [80]. Such a strategy allows one to avoid geometry optimizations at the HF-SCF level in conjunction with a basis set as large as a quintuple-zeta, thus only requiring equilibrium structures at the MP2/cc-pVTZ and MP2/cc-pVQZ levels to be performed. According to the literature reporting the comparison with best-estimated equilibrium structures obtained using the "gradient" CCSD(T)/CBS+CV scheme as well as with experiment [31,66–68,81], the so-called "cheap" geometry scheme is robust, reliable and accurate [68,74,77]. Indeed, it is noted that for bond distances the maximum absolute deviation with respect to CCSD(T)/CBS+CV is smaller than 0.001 Å. To further extend the dimension of systems amenable to accurate calculations, the double-hybrid B2PLYP functional [32] in conjunction with a triple-zeta quality basis set provides a good alternative in predicting equilibrium structures [66,67,83,84]. Indeed, despite the reduced computational cost, the B2PLYP/(aug-)cc-pVTZ level of theory provides better estimates than CCSD(T)/cc-pVTZ (within the frozen-core approximation), with the accuracy on bond distances ranging from 0.001 Å to 0.003 Å. Therefore, B2PLYP/(aug-)cc-pVTZ represents a very good compromise between computational cost and accuracy, thus being particularly suitable to determine the reference structures for energetic calculations (see, e.g., Refs. [67,72]). Furthermore, it should be noted that the choice of the geometry does not affect significantly the energetics and that methods rooted in the density functional theory (DFT), also employing basis sets as small as the double-zeta ones (see, e.g., Ref. [68]), are suitable for the purpose, thus allowing the application of CCSD(T)-based composite schemes for electronic energy evaluations well beyond the "medium-sized molecule" limit. All DFT computations reported in the present review have been performed with the GAUSSIAN package [85].

To provide an example, the results for the conformational analysis of glycine (in terms of energy differences with respect to the Ip/ttt conformer) are summarized in Fig. 5 (for a full account, the reader is referred to Ref. [68] and references therein). The electronic energies have been computed by means of the CCSD(T)/CBS+CV composite scheme described above at the corresponding optimized structures (CCSD(T)/CBS+CV) as well as using B3LYP/SNSD [86–88] equilibrium geometries. First of all, the very limited effect of the reference geometry is pointed out, the two sets of relative energies differing usually by less than 0.2 kJ/mol. The only exception is noted for the IIn/ccc conformer, which is particularly challenging because of its flat potential energy surface (PES). To address the accuracy

of the CCSD(T)/CBS+CV energies, enthalpy differences are compared to the available experimental data. The thermodynamic functions have been obtained using the Hindered-Rotor Anharmonic Oscillator (HRAO) model [89], thus pointing out an accuracy of 1 kJ/mol for the computed values. The reader is referred to [89] for all methodological details.

3.2. Quantum-chemical predictions of spectroscopic parameters

As already mentioned, the knowledge on the universe chemical inventory has been provided (and is continuously updated) by means of astronomical observations (see, e.g., Ref. [90]) of the molecular spectroscopic features, which can be considered as molecular "fingerprints".

Spectroscopic features are usually accurately obtained in laboratory spectroscopy studies that are increasingly assisted by quantum-chemical calculations to guide and support the spectral recording and analysis (see, e.g., Refs. [23, 91–100]). Indeed, computational tools play an increasing role in molecular spectroscopy not only by actively supporting measurements, but also in designing new experiments, thus permitting the identification of the best experimental conditions according to the facilities available. Therefore, the last decades have witnessed a stronger and stronger integration between experiment and theory in many research areas (see, e.g., Refs. [36,91,94,101–103]). The developments in theoretical methodologies as well as in hardware facilities have led to such improvements that, nowadays, state-of-the-art computational methodologies are able to provide results that are comparable to those delivered by accurate experimental techniques, even if in limited spectroscopic fields and for limited – mostly in size – systems. As a consequence, computational spectroscopy also plays the role of complementing experiment, thus providing the missing information that cannot be obtained by the specific experiment under consideration. In this section, the methodology for accurate quantum-chemical computations of spectroscopic properties is presented and the accuracy obtainable is also discussed.

3.2.1. Rotational spectroscopy

To obtain the rotational energy levels for a given vibrational state, with the ground state usually being the one of interest, the rotational Schrödinger equation needs to be solved. The most important terms of the effective rotational Hamiltonian are the pure rotational and centrifugal-distorsion contributions [104]. The rotational Hamiltonian, within the semi-rigid rotor approximation, can be written as

$$H_{rot} = H_R + H_{qcd} + H_{scd} + \dots , \tag{3}$$

where H_{qcd} and H_{scd} are the quartic and sextic centrifugal terms, respectively, the dots referring to the possibility of including higher-order centrifugal contributions. H_R is the rigid-rotor Hamiltonian:

$$H_R = \sum_{\tau} B_{\tau}^{eq} \mathbf{J}_{\tau}^2 \ , \tag{4}$$

where $\mathbf{J}$ is the rotational angular momentum operator and B_{τ}^{eq} is the rotational constant along the τ intertial axis, and it is defined as follows:

$$B_{\tau}^{eq} = \frac{\hbar^2}{2hcI_{\tau\tau}^{eq}} \ , \tag{5}$$

where $I_{\tau\tau}^{eq}$ is the τth diagonal element of the equilibrium intertia tensor, $\mathbf{I}^{eq}$. From a computational point of view, equilibrium rotational constants are straightforwardly obtained from geometry optimizations, $\mathbf{I}^{eq}$ only depending on the molecular structure and isotopic masses. B_{τ}^{eq} provides the most important contribution to the corresponding rotational constant. However, the effect of molecular vibrations cannot be neglected, the dependence of the rotational constants on the vibrational quantum numbers thus needing to be incorporated. The effect of vibrations on the rotational motion can be conveniently described by means of vibrational perturbation theory (VPT) [104,105]. While there are no corrections at the first order in VPT, at the second order (VPT2) the vibrational dependence of the rotational constant results [106]:

$$B_{\tau}^{v} = B_{\tau}^{eq} - \sum_{i=1}^{N} \alpha_{i,\tau} \left(v_i + \frac{d_i}{2} \right) \ , \tag{6}$$

where the superscript v denotes a specific vibrational state and the sum runs on all fundamental vibrational modes i, with v_i being the corresponding quantum number and d_i its degeneracy order. The $\alpha_{i,\tau}$'s are the so-called vibration-rotation interaction constants and contain three contributions: the first one is a corrective term related to the moment of inertia, the second one is due to the Coriolis interactions, and the last is an anharmonic correction. Therefore, from a computational point of view, anharmonic force field calculations are required to obtain the vibrational corrections to equilibrium rotational constants in order to obtain the rotational constants of the vibrational ground state:

$$B_\tau^0 = B_\tau^{eq} + \Delta B_\tau^0 = B_\tau^{eq} - \frac{1}{2}\sum_{i=1}^{N}\alpha_{i,\tau} \; . \tag{7}$$

Moving to the centrifugal-distorsion Hamiltonian, the quartic terms only depend on the harmonic part of the PES, while the computation of the sextic centrifugal-distortion constants involves harmonic, anharmonic, and Coriolis perturbation terms, thus requiring anharmonic force field computations.

To obtain accurate equilibrium rotational constants, the equilibrium structure should be accurately determined. This can be accomplished by means of the composite schemes previously introduced. According to two recent statistical studies based on molecular species only containing first-row atoms [48] and on molecules also bearing second-row elements [107], 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, in order to keep high accuracy, it is mandatory to include the vibrational corrections, with the latter being efficiently computed already at the DFT level. Even the low B3LYP/SNSD level has been demonstrated to be suitable to meet the required accuracy [108]. Moving to centrifugal distortion constants, as pointed out above, they can be obtained as byproduct of the calculations required to evaluate the vibrational corrections to rotational constants. In Ref. [109], it was shown that different levels of theory, ranging from HF-SCF to CCSD(T), provide very similar results.

According to the literature on this topic (see e.g. Refs. [31,48,107,110]), best-estimated spectroscopic parameters, computed as described above, lead to 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 far-infrared region. This means that 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. An example is provided by Fig. 6, which shows the comparison between experiment and theory for the $J = 12 \leftarrow 11$ rotational transition of two isotopologues ($H^{13}CO^+$ and $HC^{17}O^+$) of the HCO^+ ion, which was the first ionic species detected in space. This figure further points out that state-of-the-art quantum-chemical computations are able to guide astronomical searches and/or assignment only when astronomical spectra are characterized by a limited number of features, very well separated one from the other, thus confirming – as addressed in the previous section – that, while computations in this field are suitable for guiding experiment in the laboratory, they do not have the proper accuracy for assigning broadband surveys like that shown in Fig. 1.

3.2.2. Vibrational spectroscopy

To solve the vibrational problem, perturbative methods are particularly effective. The framework of VPT2 is based on Watson Hamiltonian [111] and Taylor expansions of the potential (V) and vibrational (E_v) energies, as well as vibrational wavefunction, up to the second order [112]. The vibrational Hamiltonian (H_{vib}) is defined as follows [111]:

$$\begin{aligned} H_{vib} = & \frac{1}{2}\sum_{i=1}^{N}\omega_i(\mathbf{p}_i{}^2 + \mathbf{q}_i{}^2) + \frac{1}{6}\sum_{i,j,k=1}^{N} k_{ijk}\mathbf{q}_i\mathbf{q}_j\mathbf{q}_k \\ & + \frac{1}{24}\sum_{i,j,k,l=1}^{N} k_{ijkl}\mathbf{q}_i\mathbf{q}_j\mathbf{q}_k\mathbf{q}_l + \sum_\tau \mathbf{B}_\tau^{eq}\sum_{i,j,k,l=1}^{N}\zeta_{ij,\tau}\zeta_{kl,\tau}\sqrt{\frac{\omega_i\omega_k}{\omega_j\omega_l}}\mathbf{q}_i\mathbf{p}_j\mathbf{q}_k\mathbf{p}_l + U \end{aligned} \tag{8}$$

At the VPT2 level, the resolution of the Schrödinger equation associated to the Hamiltonian above leads to the energy levels E_m (in cm^{-1}), with m being a generic vibrational state:

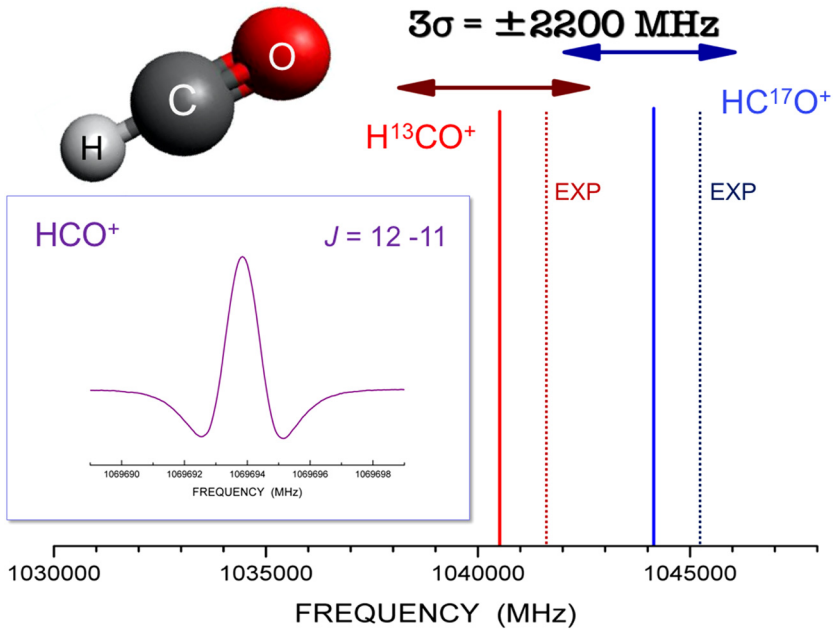

Fig. 6. The $J = 12 \leftarrow 11$ rotational transition of $H^{13}CO^+$ and $HC^{17}O^+$: comparison between computed (CCSD(T)/CBS+CV+fT+fQ data augmented by vibrational corrections at the CCSD(T)/cc-pCVQZ level) and experimental frequencies. For the main isotopic species, the experimentally recorded transition is shown in the inset.

$$E_m = E_0 + \sum_{i=1}^{N} v_i^m \omega_i + \sum_{i,j=1}^{N} \chi_{ij} \left[v_i^m v_j^m + \frac{1}{2}\left(v_i^m + v_j^m\right) \right], \tag{9}$$

where v_i^m is the number of quanta associated to mode i in state m and ω_i the corresponding harmonic wavenumber. E_0 is the zero-point vibrational energy (ZPE):

$$E_0 = \sum_{i=1}^{N} \frac{\omega_i}{2} + \sum_{i,j=1}^{N} \frac{k_{iijj}}{32} - \sum_{i,j,k=1}^{N} \left[\frac{k_{iik} k_{jjk}}{32\omega_k} + \frac{{k_{ijk}}^2}{48(\omega_i + \omega_j + \omega_k)} \right] \tag{10}$$
$$- \sum_{\tau} \frac{\mathbf{B}_\tau^{eq}}{4} \left[1 - \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \{\zeta_{ij,\tau}\}^2 \frac{(\omega_i - \omega_j)^2}{\omega_i \omega_j} \right].$$

In eq. (9), χ is the anharmonicity contributions matrix defined as follows:

$$16\chi_{ii} = k_{iiii} - \frac{5{k_{iii}}^2}{3\omega_i} - \sum_{j=1, j\neq i}^{N} \frac{\left(8{\omega_i}^2 - 3{\omega_j}^2\right){k_{iij}}^2}{\omega_j\left(4{\omega_i}^2 - {\omega_j}^2\right)} \tag{11}$$

$$4\chi_{ij} = k_{iijj} - \frac{2\omega_i {k_{iij}}^2}{\left(4{\omega_i}^2 - {\omega_j}^2\right)} - \frac{2\omega_j {k_{ijj}}^2}{\left(4{\omega_j}^2 - {\omega_i}^2\right)} - \frac{k_{iii}k_{ijj}}{\omega_i} - \frac{k_{jjj}k_{iij}}{\omega_j} \tag{12}$$
$$+ \sum_{k=1, k\neq i,j}^{N} \left[\frac{2\omega_k\left({\omega_i}^2 + {\omega_j}^2 - {\omega_k}^2\right){k_{ijk}}^2}{\Delta_{ijk}} - \frac{k_{iik}k_{jjk}}{\omega_k} \right] + \frac{4\left({\omega_i}^2 + {\omega_j}^2\right)}{\omega_i\omega_j} \sum_{\tau} \mathbf{B}_\tau^{eq} \{\zeta_{ij,\tau}\}^2$$

where

$$\Delta_{ijk} = {\omega_i}^4 + {\omega_j}^4 + {\omega_k}^4 - 2\left({\omega_i}^2{\omega_j}^2 + {\omega_i}^2{\omega_k}^2 + {\omega_j}^2{\omega_k}^2\right). \tag{13}$$

Transition energies from the ground state, ν_m, are then straightforwardly obtained from Eqs. (9) and (10) as $E_m - E_0$ difference.

A critical issue of the VPT2 approach is the potential presence of resonances, called Fermi resonances (FRs), which couple fundamental and overtones or fundamental and 2-mode combinations close in energies. Although Fermi resonances can also be present in the conventional expression of anharmonic ZPE, it is possible to rearrange terms arriving to the resonance-free ZPE expression given in Eq. (1) [113]. Removal of all the resonant terms results in a truncated perturbative treatment, usually referred to as DVPT2 (with D standing for deperturbed). Then, a common correction is to add a subsequent variational treatment, which leads to the so-called GVPT2 approach (with G standing for generalized). In this model, the VPT2 energies devoid of resonances are the diagonal elements of the variational matrix and the discarded terms are reintroduced as off-diagonal coupling terms. New states and energies are obtained from the diagonalization of this matrix. An alternative approach, proposed by Kuhler et al. [114] and denoted degeneracy-corrected perturbation theory (DCPT2), replaces all potentially resonant terms with non-resonant forms. A further generalization leads to a hybrid scheme, referred to as HDCPT2 [89], that combines DCPT2 and VPT2 results for each potentially resonant terms to overcome the shortcomings of the former far from resonances. This is the method of choice for thermodynamic and/or kinetic studies, which require comparable treatments for different stationary points (see the corresponding section).

A second type of resonances (usually referred to as Darling-Dennison resonances (DDRs)) refers to coupling terms missed in the VPT2 vibrational energy due to the truncation of higher-order perturbative contributions. They can be directly included during the variational correction step mentioned earlier as off-diagonal elements of the variational matrix involving either overtones and/or combinations (2-2) or fundamentals (1-1).

Composite schemes can also be applied to the accurate computation of harmonic force fields. According to the literature on this topic (see, for example, Refs. [68,115,116]), composite approaches based on CCSD(T) are able to provide harmonic frequencies with an accuracy of 5-10 cm^{-1}. However, their applicability is limited to small- to medium-sized molecules. As in the case of equilibrium structures, B2PLYP in conjunction with triple-zeta quality basis sets provides an alternative for larger systems and shows an accuracy of 8-15 cm^{-1}. Due to the computational requirements, the anharmonic part of force field calculations, i.e., the computation of cubic and quartic semi-diagonal force constants, is usually evaluated at a lower level of theory. While for small molecules the CCSD(T) method can be employed, when increasing the molecular size global-hybrid (like B3LYP) or double-hybrid functionals are mostly used. Improvements in the accuracy can be achieved 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, while for the anharmonic part MP2 or DFT are used. When harmonic frequencies evaluated by means of composite schemes are combined with anharmonic corrections obtained using either B3LYP in conjunction with double-zeta quality basis sets or B2PLYP with triple-zeta quality sets, fundamentals as well as overtone and combination bands of "standard" (non problematic) molecules are predicted with mean absolute errors (MAEs) of 6-8 cm^{-1} and 5-7 cm^{-1} (see, for example, Ref. [94]), respectively.

Fig. 7 shows the simulated infrared spectra of the phenalenyl cation and anion in the 0-3500 cm^{-1} range, based on the B3LYP/6-31+G(d,p) results from Ref. [117]. The accuracy of the B3LYP/6-31+G(d,p) level of theory for the evaluation of the vibrational transition wavenumbers up to two quanta for aromatic compounds can be analyzed from the comparisons between experimental and computed data available in the literature. For example, in Ref. [115], B3LYP/6-31+G(d,p) data are compared with state-of-the-art computational (using a hybrid force filed as described above) and experimental results for uracil. This comparison points out a mean absolute error of 12 cm^{-1} for B3LYP/6-31+G(d,p) wavenumbers, thus suggesting an accuracy sufficient for guiding astronomical searches of IR signatures; indeed, IR astronomy – unlike radioastronomy – is usually characterized by medium resolution. Even if such an accuracy should be taken with a little caution because of error compensation, some benchmark studies demonstrated the suitability of the B3LYP functional, in conjunction with double-zeta quality basis sets, for reliable predictions of vibrational spectra (see, for example, Refs. [118–120]).

3.2.3. Electronic spectroscopy

The highest-energy transitions of a molecule are those involving excitations of electrons that can be investigated by further increasing the energy of the probe electromagnetic field above the IR region. The energetic range spanned by electronic excitations is much larger than that of the other spectroscopic techniques, indeed going from the visible region, at about 400-800 nm, corresponding to the average excitation energies of valence electrons, to the X-ray region, between 100 and 400 eV, which promotes excitations of core electrons. The simulation of electronic spectra is more challenging than that of IR for several reasons. First of all, the calculation of properties is intrinsically more

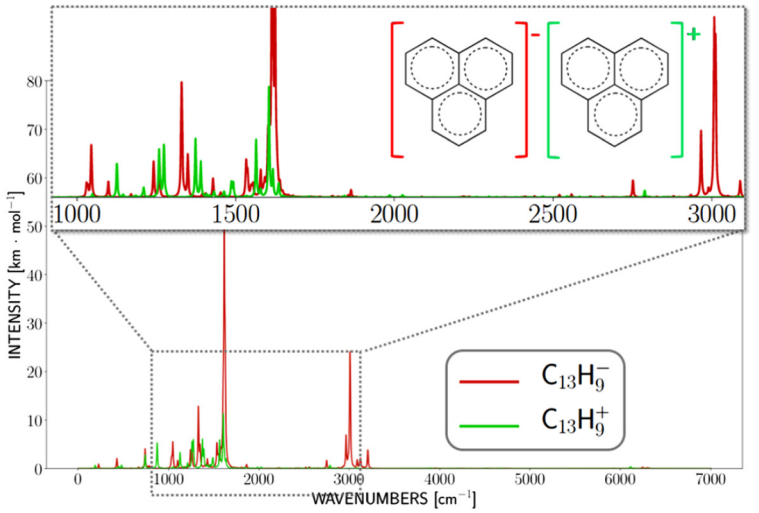

Fig. 7. Simulated anharmonic IR spectra of $C_{13}H_9^+$ and $C_{13}H_9^-$ in the 0-7000 cm^{-1} range computed at the B3LYP/6-31+G(d,p) level of theory obtained by convoluting stick spectra with a Lorentzian line profile (half-width at half-maximum: 5 cm^{-1}). In the inset: the portion of spectra in the 1000-3000 cm^{-1} range (halfwidth at half-maximum: 2 cm^{-1}) is shown.

complex for excited electronic states. Even if significant work has been done in recent years to produce reliable and black-box approaches for the calculation of excited electronic states [121,122], especially in connection with the time-dependent extension of density functional theory (TD-DFT) [123–125], they are still far from being comparable to their counterparts for ground states. Another difficulty arises from the fact that the number of phenomena occurring after an electronic excitation is much larger than that of the other spectroscopies. Furthermore, emission processes subsequent to the primary excitation, such as fluorescence [126], have also to be taken into account. In addition, electronic states are also characterized by their spin and transitions between different spin states are possible. A proper reproduction of these phenomena, which are at the basis – for instance – of phosphorescence spectroscopy [127], requires, from a theoretical point of view, the inclusion of relativistic effects [127], thus increasing the complexity of the electronic structure computations to be performed. All these effects need to be included in spectral simulations in order to obtain results directly comparable to experimental data, and this increases significantly the complexity of modeling electronic spectroscopies.

The resolution of experimental electronic spectra is largely sufficient to reveal a fine structure determined by lower-energy transitions. In particular, transitions between vibrational levels of different electronic states can be singled out in high-resolution spectra and, by further increasing the resolution, even the rotational sub-structure can be evidenced for small molecules [128,129]. Even neglecting rotational effects, the reproduction of vibrational signatures in electronic spectra requires the computation of the PESs of different electronic states, which means – in practice – obtaining the gradients and hessians from the corresponding quantum-chemical methods [130–132]. For this reason, simulations of vibronic spectra at the harmonic level is a challenging task already for medium-sized molecules and the further inclusion of anharmonic effects is possible only for very small systems [133–135].

Another aspect of increasing importance is the ability of chiral systems to interact differently with left- and right-handed circularly polarized light. This effect, known as circular dichroism, is at the basis of all chiroptical measurement instruments [101], which are the primary tools for the assignment of absolute configurations of chiral systems. In the field of electronic spectroscopy, both absorption (ECD) [136], emission (circularly polarized luminescence CPL), and circularly polarized phosphorescence (CPP) [137] techniques are currently in use. The simulation of chiroptical spectroscopies involves additional challenges with respect to their standard counterparts because band intensities are ruled by mixed electric-magnetic transition properties, whose calculation is less straightforward than those for pure electric properties [138,139]. Furthermore, chiroptical spectra are obtained as differences between the spectra recorded with left- and right-handed circularly polarized light. For this reason, an experimental spectrum can display both positive and negative bands and, in general, has a low intensity. This makes the simulation of chiroptical spectra even more sensitive to the quality of the underlying quantum-chemical model. In fact, different contributions to the overall band shape might have different signs and thus they do not necessarily add up, and can even cancel

each other. Therefore, a small change in the individual terms can lead to strong variations in the overall spectrum [140–142].

The need to simulate different spectroscopic techniques calls for the development of a general simulation approach, which – in our group – has been realized through the virtual multifrequency spectrometer (VMS) [94], whose electronic spectroscopy module is in charge of the simulation of all kinds of electronic spectra taking also into account vibrational signatures and chiroptical extensions. Being part of the general VMS tool [94], this module furthermore benefits from powerful pre- and post-processing graphical engines.

As it is well known, in the framework of first-order time dependent perturbation theory, the probability of a transition between the initial state ψ_i and the final state ψ_f (having energies E_i and E_f) is given by

$$p_{if}(t) = \frac{4\pi}{3\hbar^2 c} W(\omega)\, |\langle\psi_f|\mathcal{O}^A_{if}|\psi_i\rangle\langle\psi_f|\mathcal{O}^B_{if}|\psi_i\rangle|\, t\, \delta(\omega_{if} - \omega)\,, \tag{14}$$

where $W(\omega)$ is the intensity of the incident radiation and $\omega_{if} = (E_f - E_i)/\hbar$. The presence of the Dirac δ function shows that the transition occurs only when the frequency of the incident electromagnetic field matches ω_{if}. Thus, the transition probability is determined by the transition moments $\langle\psi_f|\mathcal{O}^A_{if}|\psi_i\rangle = \mathcal{O}_{if}$ (related to electric and, possibly, magnetic components). Within Born-Oppenheimer approximation, the overall wave-function can be factorized into its electronic (ϕ) and nuclear (ψ) contributions, the latter corresponding – in the following – only to vibrations. As a consequence, for orthonormal states, one can write

$$\langle\mathcal{O}_{if}\rangle \simeq \langle\psi_{r(f)}|\mathcal{O}^e_{if}(R)|\psi_{s(i)}\rangle\,, \tag{15}$$

where $\mathcal{O}^e_{if}(R)$ is the electronic transition moment as a function of the nuclear coordinates R. Since the analytical dependence of the electronic transition moment on the nuclear coordinates is not known, it is developed as a Taylor expansion around the equilibrium geometry $\mathbf{R}_i^{eq}$ of the initial state. In this connection, it is useful to refer to the normal modes of the initial state Q, so that

$$\mathcal{O}^e_{if}(R) = \mathcal{O}^e_{if}(\mathbf{R}_i^{eq}) + \sum_{a=1}^{N_{vib}} \left(\frac{\partial\mathcal{O}_{if}}{\partial Q_a}\right) Q_a\,, \tag{16}$$

where the derivatives are computed at the equilibrium geometry of the initial state ($\mathbf{R}_i^{eq}$), where all Q_a's vanish. The first term is known as Franck-Condon (FC) contribution and the second term as Herzberg-Teller (HT) contribution. Then, the expectation value of the transition property can be rewritten as

$$\langle\mathcal{O}_{if}\rangle \simeq \mathcal{O}^e_{if}(\mathbf{R}_i^{eq})\langle\psi_{r(f)}|\psi_{s(i)}\rangle + \sum_{a=1}^{N_{vib}} \left(\frac{\partial\mathcal{O}_{if}}{\partial Q_a}\right)\langle\psi_{r(f)}|Q_a|\psi_{s(i)}\rangle\,. \tag{17}$$

The integrals in the second term can be expressed as combinations of overlap integrals. Furthermore, the normal coordinates of the final state $\bar{\mathbf{Q}}$ can be expressed as

$$\bar{\mathbf{Q}} = \mathbf{QJ} + \mathbf{K}\,, \tag{18}$$

where $\mathbf{J}$ is a $N_{vib} \times N_{vib}$ matrix, denoted as Duschinsky matrix, while $\mathbf{K}$ is a N_{vib}-dimensional vector, the so-denoted shift vector. In practice, $\mathbf{K}$ represents the shift between the equilibrium geometries of the two electronic states involved in the transition, while $\mathbf{J}$ expresses the modes of the final electronic state as linear combinations of those of the initial state.

Owing to the limitations of the harmonic approximation, two different representations are usually employed. In the first one, referred to as adiabatic, the PES of the final state is expanded around its own equilibrium geometry (adiabatic Hessian, AH), whereas in the vertical approximation, it is expanded around the equilibrium geometry of the initial state (vertical Hessian, VH). In both cases, simplified models, where mode-mixing as well as frequency change effects are neglected, can be derived (adiabatic shift, AS, and vertical gradient, VG, respectively).

By using those approximations, the sum given in Eq. (17) can be expressed in terms of the overlap integrals between the vibrational levels of the two electronic states, usually referred to as FC integrals. In the so-called time independent (TI) approach, these integrals are calculated using either analytical or recursive formulae, the second approach being better suited for a general implementation.

It is possible to gather several spectroscopic techniques (OPA, OPE, ECD and CPL) in a unique framework by adopting a general sum-over-states formula

$$\Im(\omega) = \alpha\omega^{\beta} \sum_{i} \sum_{f} \rho_i \mathcal{O}_{if}^{A,e} \mathcal{O}_{if}^{B,e*} \, \delta(\omega_{if} - \omega) \, , \tag{19}$$

where ρ is the Boltzmann population and the parameters α, β, $\mathcal{O}_{if}^{A,e}$, and $\mathcal{O}_{if}^{B,e}$ depend on the specific spectroscopic technique (see Ref. [143] for further details). The main limitation of this model relies on the number of FC integrals to compute, which increases significantly with the size of the target system. In order to make computations feasible also for medium- and large-sized systems, prescreening strategies can be profitably employed, to select a priori the FC integrals giving the largest contributions to the final spectrum. In particular, a class-based prescreening system has been demonstrated to perform well also for large-sized semi-rigid systems. In the alternative time-dependent (TD) framework, the vibronic spectrum is calculated from the Fourier transform of the autocorrelation function, as follows:

$$\Im(\omega) = \alpha\omega^{\beta} \int_{-\infty}^{+\infty} dt \, \chi(t) e^{-i(\omega - \omega_{if})t} \, , \tag{20}$$

where $\chi(t)$ is the transition moment autocorrelation function, which has an analytical (albeit complicated) expression within the approximations outlined above. This equation can be used to calculate the vibronic spectrum by evaluating the Fourier integral using a numerical integration algorithm. In the TD formulation (at variance with respect to the TI one), infinite summations are no longer present, so that the computational cost of the overall procedure is independent on the temperature.

Both TI [141–144] and TD [145–148] approaches have been implemented in the framework of a general platform including all the spectroscopic techniques mentioned above. The more standard TI approach, which is based on a sum-over-states formulation, allows the assignment of individual vibronic transitions. However, in spite of remarkable efficiency improvements, convergence of the results can be sometimes problematic and the inclusion of temperature effects is quite cumbersome. On the other hand, TD algorithms provide fully-converged band-shapes at low computational cost, also when temperature effects are included. However, the assignment of individual transitions is no longer possible. Therefore, the combined use of TD and TI approaches provides a general and robust tool in order to reconcile accuracy and interpretability.

To target large-sized, potentially flexible systems, thus showing large-amplitude motions (LAMs), harmonic vibronic models in Cartesian coordinates have been generalized to support internal coordinates [149]. Such an approach improves significantly the results because it allows to strongly reduce mode-couplings. In practice, this results in much more accurate band-shapes already at the harmonic level, which can be further improved by treating, the leading LAMs by quasi-variational procedures based on the so-called discrete-variable representation (DVR) [150]. At the same time, the leading anharmonic contributions on the other small amplitude motions (SAMs) can be taken into account by VPT2, also thanks to the recent development of effective analytical hessians for TD-DFT [151]. The same algorithms can also be applied to simulate the vibrational signatures of near-edge X-ray (NEXAFS) and X-ray photoelectron (XPS) spectra.

As it is well known, a wide range of phenomena can occur following an electronic excitation, including both radiative and non-radiative decays. In particular, emission spectroscopic techniques, such as fluorescence and phosphorescence, are strongly influenced by competitive non-radiative decay processes. In order to build a complete framework, all these phenomena are properly taken into account.

The potentialities of the above general strategy for the computation of electronic spectra will be now illustrated by a few examples. As an application of the theoretical-computational framework described above, we have selected the OPA and OPE spectra of trans-2,2′-bithiophene, which is planar in the ground state, but significantly twisted (SCCS dihedral angle of about 160 degrees) in the excited state. The TD approach has been chosen because the presence of strong distortions between initial and final states renders the convergence of TI approaches problematic. All the computations have been performed at the TD-DFT level using the CAM-B3LYP-D3BJ range-separated functional [152] including dispersion corrections (by means of the D3BJ model of Grimme [153,154]) and the SNSD basis set [155] because this model chemistry offers the best reliability/cost compromise. Comparison of the AH|FC spectrum computed using Cartesian coordinates with its experimental counterpart is quite disappointing for both position of

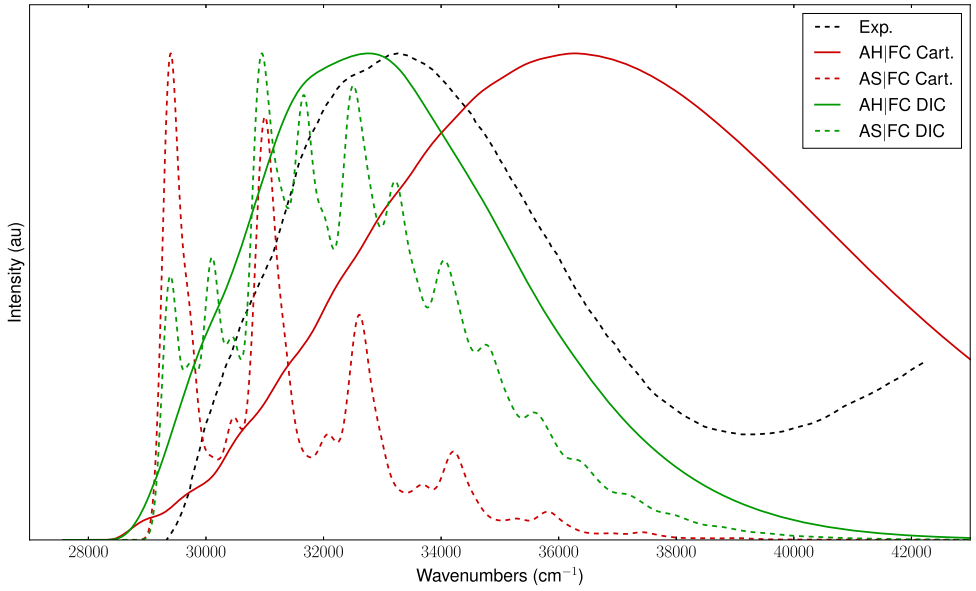

Fig. 8. Comparison between the experimental OPA spectrum of trans-2,2′-bithiophene and the corresponding simulations obtained in the framework of the FC approximation employing AH or adiabatic shift AS models using Cartesian (Cart.) or DICs.

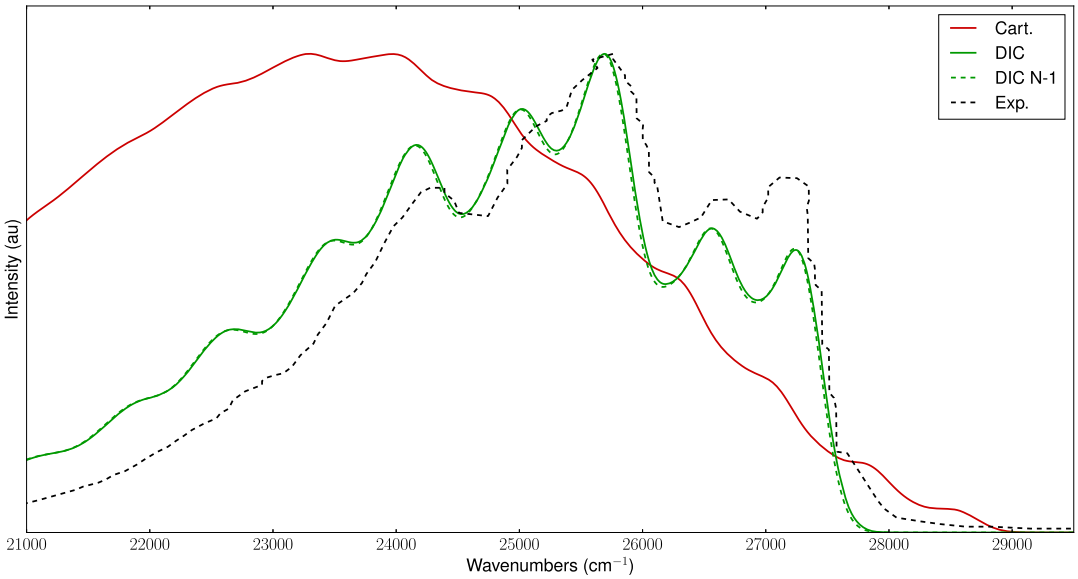

Fig. 9. Comparison between the experimental high-resolution OPE spectrum of trans-2,2′-bithiophene and the corresponding simulations obtained in the framework of the FC approximation employing the AH model using Cartesian (Cart.) or DICs. In the latter case, the inter-ring torsion has also been treated by the quasi-variational DVR approach (DIC N-1).

absorption maxima and overall band-shape. As a matter of fact, significant distortions along the inter-ring dihedral angle lead to strong unphysical couplings in the **J** matrix due to the rectilinear nature of Cartesian coordinates. As shown in Fig. 8, the use of delocalized internal coordinates (DICs) lead to significantly better results. This trend is even more evident in the OPE spectrum (see Fig. 9) due to the very high resolution of the experimental laser induced fluorescence (LIF) spectrum [156]. In such circumstances, only the use of DICs allows for correctly reproducing the experiment, while their combination with the explicit DVR treatment of the large-amplitude torsion leads to a negligible improvement.

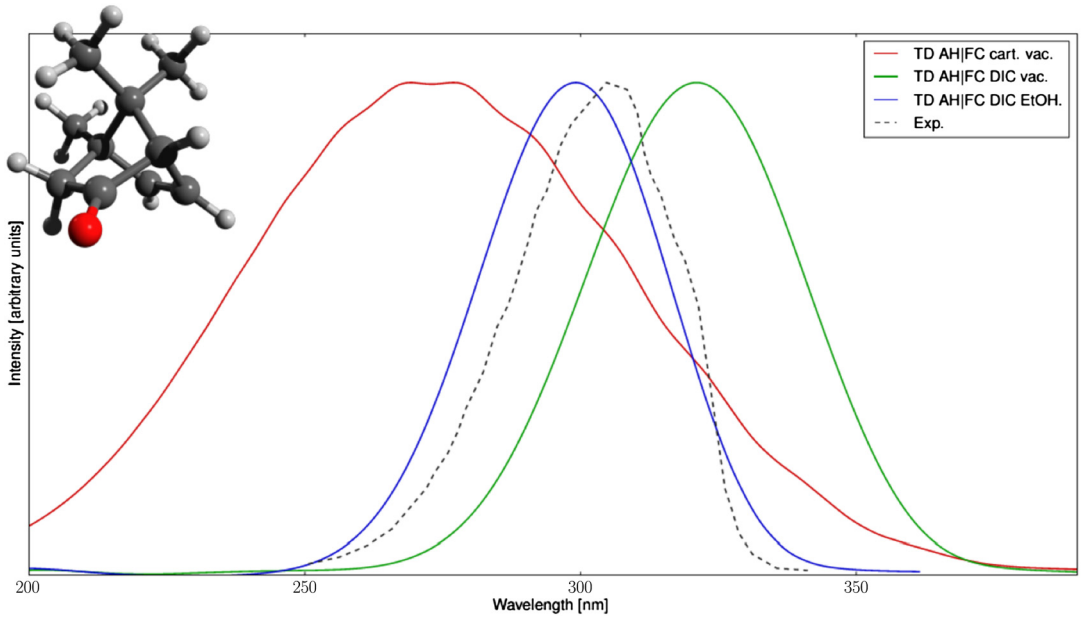

Fig. 10. Comparison between the experimental ECD spectrum for the $S_1 \leftarrow S_0$ transition of (1S)-dehydro-epicamphore [157] and the theoretical results in Cartesian and DICs coordinates computed at the TD AH|FC level in vacuum and in solution (ethanol). In the latter case also TD AH|FCHT results are shown. Gaussian distribution functions with a half-width at half-maximum of 100 cm^{-1} have been used to reproduce broadening effects.

As an example of chiroptical spectroscopy, let us consider the ECD spectrum of (1S)-dehydro-epicamphore (Fig. 10). The definition of a non-redundant set of internal coordinates for this system, as well as for all bicyclic compounds is not straightforward. On the other hand, DICs can be easily built starting from the topology of the molecule. From an experimental point of view, an extensive study of ECD and CPL spectra for several camphore derivatives has been recently performed. The ECD spectrum of (1S)-dehydro-epicamphore is particularly challenging because the $S_1 \leftarrow S_0$ ($\pi \rightarrow \pi^*$) transition involves a significant out-of-plane distortion of the carbonyl group, which rules the sign of the ECD spectrum [157] according to the so-called octant rule [158]. The ECD spectra have been simulated starting from DFT (for the ground state) and TD-DFT (for the excited state) B3LYP/SNSD computations. Next, vibrational modulation effects have been computed by the TD approach described above at the AH|FC level and using different coordinate sets. The spectra collected in Fig. 10 show once again that, for electronic excitations involving significant structural rearrangements, internal coordinates lead to results in agreement with experiment, whereas Cartesian coordinates face against strong difficulties related to the overestimation of the couplings among several normal modes (C=O wagging and C=O stretching in the present case). As a matter of fact, even if the resolution of the experimental spectrum is low, and single vibronic peaks cannot be detected, the reproduction of the overall band shape remains rather poor in Cartesian coordinates because the convolution of all the vibronic transitions gives an excessively broadened spectrum. By employing internal coordinates, the agreement between experiment and simulations becomes satisfactory, even if a constant shift between the two spectra is present. Since the experimental spectrum has been recorded in ethanol solution, computations have been redone by employing the polarizable continuum model (PCM) [159] to take solvent effects into account. At this level, the absolute position of the spectrum is reproduced satisfactorily and the bandshape does not change significantly. The relevance of the good results of the last example in an astrochemical context is the ability of quantum chemistry in modeling the spectroscopic features also in condensed phases.

3.3. Thermochemistry and kinetics

A central quantity in the framework of thermochemistry and kinetics is the molecular rovibrational density of states (DOS) $\rho(E) = \Delta N(E)/\Delta E$, i.e. the number $\Delta N(E)$ of rovibrational states per energy interval ΔE within the range $[E, E + \Delta E]$ of the involved molecular species. The rotational and vibrational motions of the molecule are usually

assumed to be approximately factorized (in other words, the energy of the molecule, E, is approximately the sum of the rotational, E_r, and vibrational, E_v, energies). As a result, the rovibrational DOS can be written as a convolution of its rotational and vibrational contributions. While a classical expression can be used for the rotational DOS, a quantum anharmonic treatment is recommended for the vibrational motions.

Direct count of harmonic modes can be efficiently performed employing the Beyer-Swinehart [160] algorithm and this can be extended to general separable degrees of freedom thanks to the Stein-Rabinovitch [161] adaptation. However, the direct count of coupled anharmonic degrees of freedom becomes rapidly unaffordable from a computational point of view when increasing the size of the system under consideration. A viable alternative is represented by the Wang-Landau sampling method [162,163] as extended to the quantum anharmonic vibrational problem by Basire et al. [164] and refined by Nguyen and Barker [165]. In a general implementation of the Wang-Landau scheme [166], it is assumed that a set of N anharmonic oscillators described by a corresponding N-uplet of quantum numbers has an energy level manifold described by a polynomial form. For instance, as described above, VPT2 leads to the quadratic form of Eqs. (9) and (10), but more general expansions can be used by fitting the vibrational levels obtained by, e.g., variational methods. Even more generally, the Wang-Landau scheme can be used with any sequence assigning an energy value to a running integer which represents a quantum number. Thus, for instance, some energy levels and the corresponding quantum numbers can be stored in an array, whereas all the other levels can be obtained from an analytical expression issuing from fitting and/or the VPT2 expression. This paves the way to use of the Wang-Landau scheme for flexible systems and thus calculating the DOS for a hindered rotor along with a set of anharmonic oscillators treated at the VPT2 level (thus avoiding the need of computing convolution functions for the overall density function).

To build up the density of states, an initial N-uplet of vibrational quantum numbers is chosen randomly and, subsequently, the vibrational quantum numbers are varied in a random fashion until the whole manifold has been adequately explored. In all cases, the derivative of the energy with respect to the quantum number is also obtained in order to ensure that the dissociation limit is not exceeded (where the turning point of the polynomial expression is reached). Dividing the energy range into small intervals, the number of levels within each state gives the density of states $\rho(E)$. Subsequently, a Laplace transform yields the partition function $Z(T)$:

$$Z(T) = \int \rho(E) e^{-\beta E} dE \,, \tag{21}$$

where $\beta = 1/(k_B T)$, with k_B being the Boltzmann constant and T the absolute temperature. The translational center of mass is rigorously separable. At not too high temperatures, it is a good approximation to treat the rotations around the principal axes as separable as well, so that

$$Z(T) \cong Z_t(T) \times Z_r(T) \times Z_v(T) \,. \tag{22}$$

Since analytical expressions are available for the translational (Z_t) and rotational (Z_r) contributions, we will concentrate in the following on the vibrational contribution (Z_v). On these grounds, a general strategy to calculate DOSs and partition functions for molecules involving anharmonic vibrations and/or hindered rotations can be summarized as follows [166]:

- Vibrational densities of states are calculated by the Wang-Landau scheme. The scheme can include both harmonic and anharmonic vibrational motions, the latter being treated by the VPT2 approach. As an option, some of the vibrational modes of the molecule can be designated as harmonic. In this case, their density of states can be easily convoluted with the rest through a Beyer-Swinehart scheme.
- Hindered rotor energy levels (obtained variationally by a one-dimensional DVR approach) are fitted to a function of the quantum number n and subsequently included in the overall Wang-Landau scheme. Together with molecules containing single hindered rotations, the approach is expected to work equally well for multiple hindered rotations, at least when the coupling between the hindered rotors is negligible. An analogous approach can be used for other large-amplitude motions (e.g. ring puckerings or inversions) provided that they are decoupled from the other SAMs [150].
- Transition states (TSs) can be modeled in the same way provided that the transition vector is properly taken into account.

Coming to reaction rates, we make explicit reference to the transition state theory (TST) and its most recent extensions both in its canonical or microcanonical versions. Noted is that for unimolecular reactions, the microcanonical variant corresponds to the Rice-Ramsperger-Kassel-Marcus (RRKM) theory. In absence of tunneling, the canonical rate constant can be written as:

$$k(T) = \frac{1}{T}\frac{Z(T)^{\neq}}{Z(T)} e^{-\beta V^{\neq}} , \tag{23}$$

where $V^{\neq}$ is the bare potential barrier, the superscript $\neq$ referring to the TS. It has to be noted that a reaction path degeneracy (i.e. the existence of equivalent paths) is rigorously incorporated in reaction rates by introducing in the right-hand term the multiplicative factor $\Im^{\neq}\Xi/\Im\Xi^{\neq}$, where $\Im$ is the number of optical isomers and Ξ is the symmetry number [167].

Within the harmonic approximation, the reactant vibrational partition function has the following analytical expression:

$$Z_V(T) = \frac{e^{-\beta V_0}}{\prod_i (1 - e^{-\beta \Delta_i})} , \tag{24}$$

where V_0 is the ground state energy and Δ_i the lowest excitation energy of mode i. To introduce the leading anharmonic effects, a simple and reliable approximation (known as simple perturbation theory, SPT [168]) is to compute V_0 and Δ_i by means of VPT2. If tunneling effects are neglected, $Z_V(T)^{\neq}$ can also be evaluated using Eq. (24).

However, tunneling introduces the problem of non-separability. Hernandez and Miller [169] have shown that, in the framework of semiclassical approximation (in which quantum numbers corresponds to actions and motion along the reaction coordinate is governed by the imaginary action θ), $Z_V(T)^{\neq}$ can be evaluated in an effective way:

$$Z_V(T)^{\neq} = \int_{-\infty}^{+\infty} d\theta \frac{1}{2} sech^2(\theta) Q(\beta, \theta)^{\neq} , \tag{25}$$

where

$$Q(\beta, \theta)^{\neq} = \sum_n e^{-\beta E(n^{\neq}, \theta)} . \tag{26}$$

Within VPT2, the energy is a quadratic function of θ, so that it can be inverted to obtain θ as an explicit function of the energy [170,171], but the correct treatment of resonances in the GVPT2 model precludes this possibility. The equation above does not imply any inversion and can be used also in connection with GVPT2 or more advanced models, but at the price of a direct count of vibrational states.

An alternative is offered by the SPT expression of the partition function [168] or by the HDCPT2 model [89], which effectively removes resonances and allows for the use of the Wang-Landau algorithm to compute the DOS and of the Laplace transform to obtain the canonical rate constant. Therefore, being the microcanonical rate constant given by

$$k(E) = \frac{G_v^{\neq}(E_v)}{h\rho(E)} , \tag{27}$$

the canonical rate constant becomes

$$k(T) = \frac{Z_t(T)^{\neq} Z_r(T)^{\neq}}{h Z_t(T) Z_r(T) Z_v(T)} \int_{-\infty}^{+\infty} G_v^{\neq}(E_v) e^{-\beta E_v} dE_v , \tag{28}$$

where $G_v^{\neq}(E_v)$ is the cumulative reaction probability:

$$G_v^{\neq}(E_v) = \sum_{n_1} \sum_{n_2} \cdots \sum_{n_{F-1}=0} P_n(E) , \tag{29}$$

with the semiclassical tunneling probability $P_n(E)$ expressed as

$$P_n(E) = \frac{1}{1 + e^{2\theta(\mathbf{n},E)}} \,. \tag{30}$$

$\theta(\mathbf{n}, E)$ is the barrier penetration integral, which – using VPT2 or HDCPT2 – is given by the following explicit function of the total energy E and of the $(F-1)$ quantum numbers of the bound vibrations [170]:

$$\theta(\mathbf{n}, E) = \pi \Delta\epsilon \frac{2}{1 + \sqrt{1 + 4\Delta\epsilon\Delta\chi}} \tag{31}$$

where

$$\Delta\epsilon = \frac{\Delta E}{\Omega_F} \quad ; \quad \Delta\chi = \frac{\chi_{FF}}{\Omega_F} \tag{32}$$

$$\Delta E = \Delta V_{ag} - E + \sum_{k=1}^{F-1} \omega_k n_k \chi_{kl} \left[n_k n_l + \frac{1}{2}(n_k + n_l) \right] \tag{33}$$

$$\Omega_F = \frac{\omega_F}{i} + \sum_{k=1}^{F-1} \frac{\chi_{kF}}{i} \left(n_k + \frac{1}{2} \right) \,. \tag{34}$$

ΔV_{ag} is the vibrationally adiabatic ground-state potential energy difference, i.e. the difference between the ZPE of the TS and that of the reactant(s). Note that only the contributions of the bound vibrations are included in the ZPE of the TS. The ω_F and χ_{kF} terms are imaginary; therefore, Ω_F is always a real quantity. The standard harmonic approximation is recovered when all terms vanish and the so-called ground state approximation is obtained by retaining only the term with all the quantum numbers equal to 0. It is noteworthy that the inclusion of anharmonicity can either increase or reduce the reaction rates computed at the harmonic level. The origin of this behavior can be better analyzed by expanding Eq. (31) in a power series around $\Delta E = 0$:

$$\theta(\mathbf{n}, E) = \pi \Delta\epsilon [1 - \Delta\epsilon\Delta\chi + \cdots] \tag{35}$$

and recalling that a reduction of $\theta(\mathbf{n}, E)$ increases the tunneling probability. The following rules of thumb can then be derived:

- Asymptotic values of anharmonic rate constants are always larger than the corresponding harmonic ones due to the increased density of bound vibrational states at saddle points.
- Positive (negative) values of χ_{FF} increase (decrease) the tunneling probability. Note that this cannot be directly expressed in terms of diagonal anharmonic force constants because χ_{FF} includes also contributions from bound modes.
- The modes Q_k for which the real $i\chi_{kF}$ couplings are negative increase Ω which, in turn, increases the tunneling probability, and this conversely decreases for bound modes with positive values of $i\chi_{kF}$. The effect is enhanced by a progressive accumulation of energy (i.e. increase of the quantum number) in the involved modes.

From a practical point of view, the Wang-Landau algorithm discussed above can be effectively parallelized and generalized for the computation of reaction rates by simply computing (and storing) Ω_F and ΔE whenever a set of quantum numbers is accepted in an energy bin and then averaging their values over the number of visits in each bin to obtain $\langle \Omega_F \rangle$ and $\langle \Delta E \rangle$ to be used in the evaluation of $\langle P(E) \rangle$ [172]. However, the theory breaks down whenever $4\Delta\epsilon\Delta\chi < -1$. In such circumstances, the second order perturbative treatment must be replaced by more sophisticate models. More generally, VPT2 becomes unreliable in the so-called deep tunneling region, i.e. at energies well below the potential energy barrier, or, equivalently, at low temperatures. Some promising improvements have been proposed in this connection based on higher order perturbation theory [173] or on improved representation of the potential energy profile up to products and reactants [174]. From another point of view, several tests have shown that remarkably accurate results can be obtained retaining only the anharmonic contributions involving the transition vector (i.e. χ_{FF} and χ_{kF}). These can be determined by just two Hessian matrices computed at small positive and negative displacements from the TS along the transition vector, provided that all missing terms of χ_{kF}, i.e. $\frac{k_{kkj}k_{jFF}}{\omega_j}$ with $k, j \neq F$ (see Eq. (12)), are negligible. More generally, reduced dimensionality models can be employed, which include the explicit

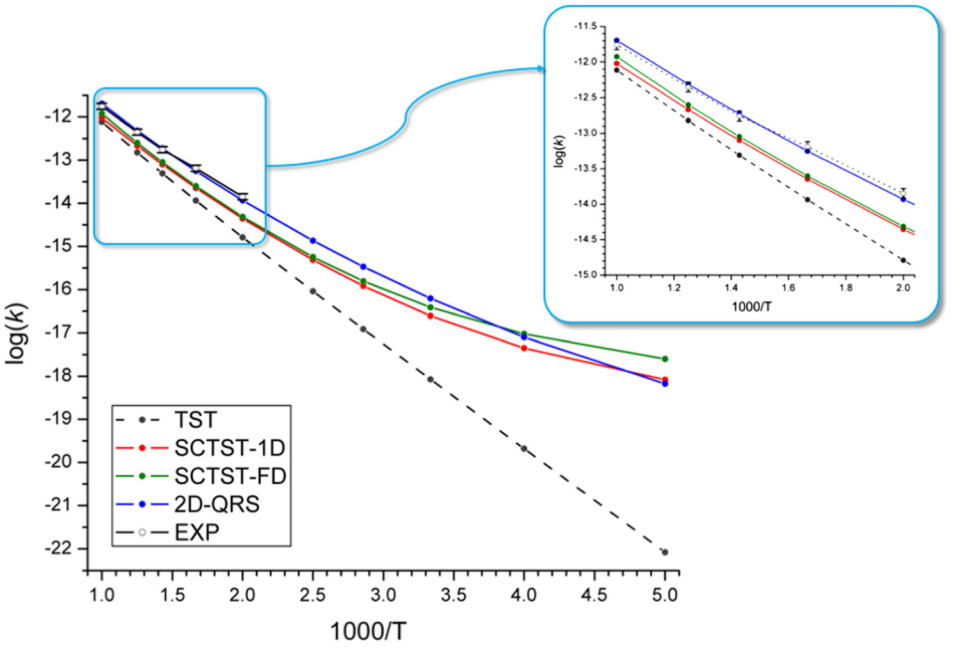

Fig. 11. Rate constant (cm^3 $molecule^{-1}$ s^{-1}; logarithmic scale) as a function of temperature (1000/T) for the reaction $H + C_2H_5 \rightarrow H_2 + C_2H_6$ obtained by conventional transition state theory (TST) and its semi-classical extension (SCTST) both in one-dimensional (1D) and full-dimensional (FD) implementations [178]. Two dimensional Quantum Reactive Scattering (2D-QRS) computations are taken from Ref. [179]; experimental data from Ref. [180].

computation of anharmonic contributions for a reduced number of bound modes strongly coupled to the transition vector [175].

As an example of the reliability of these models, Fig. 11 compares different flavors of TS-based methods to very accurate two-dimensional quantum scattering computations and to experiment for the reaction $H + C_2H_6 \rightarrow H_2 + C_2H_5$ (noted is that equilibrium geometries and vibrational frequencies at the MP2/cc-pVTZ level are combined with CCSD(T)/aug-cc-pVTZ electronic energies). It is quite apparent that tunneling plays a significant role and that the conventional TST is unreliable at low and intermediate temperatures. On the other hand, both one-dimensional (1D) and full dimensional (FD) semi-classical extensions of TST (SCTST) results are in semi-quantitative agreement with experiment and accurate computations. The effectivity of SCTST based on VPT2 anharmonic computations and Wang-Landau evaluation of DOS paves the route toward the investigation of quite complex reactions. Furthermore, hybrid and double-hybrid functionals provide sufficiently accurate geometries and frequencies, which can be employed together with accurate electronic energies issuing from the composite schemes previously described. Moreover, both VPT2 and Wang-Landau algorithms have been very effectively parallelized.

Reactions with very small (or vanishing) energy barriers must be treated in a different way, namely by means of the variational transition state theory (VTST) [176] or, more simply, by capture theory [177]. In the latter case, calculations are performed at various long-range distances of the reactants, and the energies obtained are fitted to a $1/R^6$ functional form (both for the London dispersion forces and the rotating dipole ones), which is then used to perform the capture calculation, with the assumption that each successful capture leads to the intermediate. Although VTST leads, of course, to more accurate results, when the potential energy curve is monotonic and inner TSs are not found, the much less computationally expensive scheme provided by capture theory is expected to give reliable results.

As far as dissociation back to reactants is concerned, a detailed balance argument can be used, whereby the unimolecular rate constant for back-dissociation, k_{back}, is given by the equation:

$$k_{back}(E) = k_{capt}(E)\frac{\rho_R(E)}{\rho_I(E)}\,, \tag{36}$$

where k_{capt} is the capture rate constant, ρ_R is the density of states per volume unit of the reactants, and ρ_I is the density of states for the intermediate.

For typical interstellar space and atmospheric reactions, the phenomenological time evolution for an arbitrarily interconnected kinetic system of molecular species can be usually described using a coupled set of differential equations in which the population of each species is a term of the so-called population vector [181,182]. Assuming that N species make up the kinetic scheme, the coupled set of differential equations (denoted as master equation, ME) may be written in terms of an $N \times N$ rate coefficient matrix $\mathbf{K}$ representing N coupled first-order or pseudo-first-order differential equations

$$\frac{d}{dt}\mathbf{p} = \mathbf{K}\mathbf{p}\,, \tag{37}$$

where a generic matrix element K_{ab} of the $\mathbf{K}$ matrix is the rate coefficient $k_{b\to a}$(T,P) for all possible reactions, and $\mathbf{p}$ is a vector of species populations. Since the sum of the populations is a constant, the sum of the elements of each column of the $\mathbf{K}$ matrix must be zero. This allows one to define the diagonal elements of $\mathbf{K}$ as $\mathrm{K}_{aa} = -\sum_{(b\neq a)} \mathrm{K}_{ba}$, so that $\mathrm{K}_{aa}\ \mathrm{p}_a = -\sum_{(b\neq a)} \mathrm{K}_{ba}\ \mathrm{p}_a$.

Diagonalization of this rate matrix yields a solution in terms of N eigenvalues and N eigenvectors. The first eigenvalue, often referred to as λ_0, is equal to zero, and the corresponding eigenvector gives the equilibrium Boltzmann distribution of all the species entering the general reaction network [182].

The principal hypothesis behind this model is that collisional relaxation occurs on time scales much shorter than those that characterize phenomenological kinetics [183]. Thus, separation between the internal energy relaxation eigenvalues (IEREs) and chemically significant eigenvalues (CSEs) ensures that the CSEs obtained from the diagonalization of a more general matrix (which explicitly includes collisional relaxation) are identical to those that could be obtained from diagonalization of $\mathbf{K}$ [184]. This model provides a global description of the time-dependent kinetics in terms of $N \times N$ rate coefficients, and in many cases, the phenomenological rate coefficient is the quantity of interest. However, when CSEs and IEREs are not well separated (in practice if they do not differ by more than an order of magnitude), the system cannot be represented by a set of first-order rate equations linking the populations of the N species in the vector $\mathbf{p}$. This more general situation (usually referred to as energy grained master equation, EGME) leads to very large matrices and other numerical issues. Under such circumstances, alternative approaches like, e.g. kinetic Montecarlo can be employed [185,186]. Until now, we referred to the one-dimensional ME, wherein the total rovibrational energy of the system, E, is the independent variable. Indeed, a more thorough treatment would not only consider the time-dependent evolution of the system with respect to the total energy E, but also the angular momentum, J, as a constant of motion [187]. Although the resulting two-dimensional ME has been usually considered too demanding, except for simple systems, recent works have produced the first computer codes that can be generally used at least on high-performance computers [188].

3.4. Exogenous delivery: prebiotic molecules in the interstellar medium

Even though the presence of aCOMs has been known for decades, the processes that lead to their synthesis are still unclear and often hotly debated. After the failure of gas-phase chemistry in explaining the detected abundances of aCOMS, the gas-grain chemistry was introduced and took over. This assumes that aCOMs 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. [189,190]). However, recent detections of aCOMs in cold environments (see, e.g., Refs. [191–195]) have casted doubt upon 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. [196–198]). It is therefore well accepted that gas-phase reactions have been overlooked. A significant example in this context is the formation of formamide (NH_2CHO) in the ISM. Kahane et al. [199] were the first to propose that formamide can be formed in the gas phase by the reaction of formaldehyde (H_2CO) and amidogen (NH_2). Subsequently, Barone et al. [200], Vazart et al. [39], and Skouteris et al. [201] demonstrated that, by means of accurate quantum-chemical calculations, this reaction can efficiently occur at low temperatures and can explain the available astronomical observations. Indeed, the observational study carried out in Ref. [202] with the IRAM-NOEMA interferometer supported the gas-phase formation of formamide. Furthermore, observations of its deuterated forms by Coutens et al. [203] were found in good agreement with the theoretical prediction that deuterated amidogen or formaldehyde lead to deuterated formamide [201].

3.4.1. Gas-phase chemistry

The starting point for the complete characterization of gas-phase formation pathways is the design of a feasible and accessible reactive PES leading to the aCOM of interest. The potential precursors should be suitably selected among the molecular species already detected in space. The following step is the investigation of the reactive PES itself with the identification of all stationary points (minima and transition states) along the path, with a cost-effective computational model being used at this stage. Usually, different routes toward the sought product can be derived. However, only those that can be feasible in the typical conditions of the astronomical environment under consideration will be further investigated. For instance, 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^3). In such extreme conditions, accessible chemical routes are those for which all energy barriers lie below the energy of the reactants. Subsequently, for the selected reaction schemes, an effective computational strategy requires the accurate computation of structural, energetic, and vibrational features of all the intermediates and transition states involved.

In the last years, we have developed and validated an integrated approach employing last-generation hybrid and double hybrid models rooted in the DFT to obtain reliable molecular structures at which accurate energies are evaluated by means of the composite schemes previously introduced. In parallel, DFT can be used to accurately evaluate ZPE and anharmonic vibrational frequencies within VPT2. This strategy leads to the accurate thermochemical characterization of the formation pathway, which then requires the integration of accurate kinetic calculations. As explained in the previous section, exact dynamics is presently unfeasible for systems with more than five atoms; therefore, kinetic calculations should be carried out as detailed above.

To go more in detail, the integrated approach mentioned above is based on a preliminary investigation of the full reactive PES at the B3LYP/SNSD level. All stationary points are then better investigated (and thus re-optimized) using the double-hybrid B2PLYP functional in conjunction with a triple-zeta quality basis-set augmented by diffuse functions, also including the D3BJ empirical dispersion correction. It is noted that the nature of all stationary points found on the PES is checked by computing the corresponding Hessian. Cubic and semidiagonal quartic force constants are subsequently computed in order to obtain anharmonic frequencies within VPT2, as described above. For all stationary points, improved electronic energies are then obtained by means of the so-called CCSD(T)/CBS+CV composite approach:

$$E_{CBS+CV} = E^{\infty}(SCF) + \Delta E^{\infty}(CCSD(T)) + \Delta E(CV)\,, \tag{38}$$

where the three terms in the right-hand side have been already introduced. 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 (fT) and quadruple excitations (fQ) can be accounted for in an analogous manner to what done in Eq. (1).

The impact of the level of theory on the computed reaction rate is exemplified by Fig. 12, which shows that different computational schemes can lead to rather different results. In the specific case considered, it is evident that the CBS-QB3 composite approach [204,205] provides a reaction rate which is two order of magnitude larger than that obtained using the CCSD(T)/CBS+CV scheme. A further improvement of the latter by inclusion of the fT and fQ contributions leads to a small variation in the rate constant, thus supporting the suitability and accuracy of the CCSD(T)/CBS+CV composite scheme.

Fig. 13 shows the abundance of glycolaldehyde plotted against the abundance of ethanol for five different astrochemical objects. It is seen that astronomical data follow closely the theoretical predictions based on the model investigated in Ref. [40], which is the same applied to the previous example. The three red curves correspond to different branching ratios of the ethanol radicals on hydrogen abstraction by the OH radical; for further details the reader is referred to Ref. [40].

3.4.2. Grain and gas-grain chemistry

Interstellar matter is composed of molecular and atomic gas and dust grains. In molecular clouds, the grain cores, mainly consisting of nonvolatile silicate and carbonaceous compounds (also including PAHs), are often surrounded by a grain mantle, which consists of water ice containing various molecules. Dust grains 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. Dust grains absorbing molecules and atoms on their surface provide the opportunity for chemical reactions to take place. Indeed, molecules can be synthetized not only via gas-phase chemistry, but also on bare dust grains or in ice mantles (grain-surface chemistry) [207]. Adsorbed

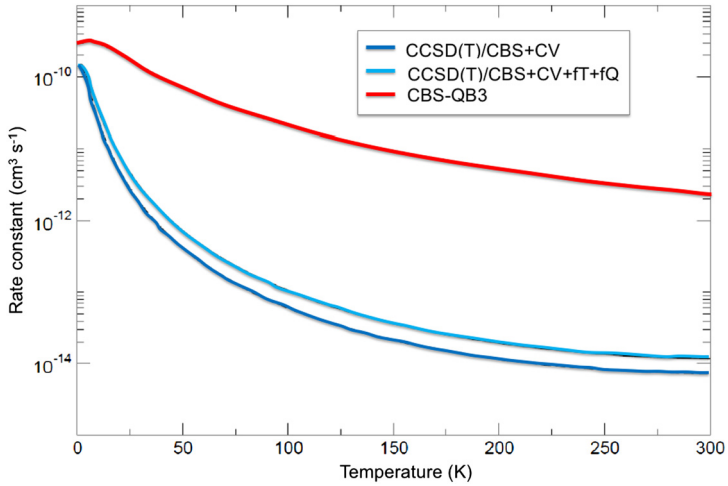

Fig. 12. Rate constant (cm^3 $molecule^{-1}$ s^{-1}; logarithmic scale) as a function of temperature for the $NH_2 + H_2CO \rightarrow HCONH_2 + H$ reaction.

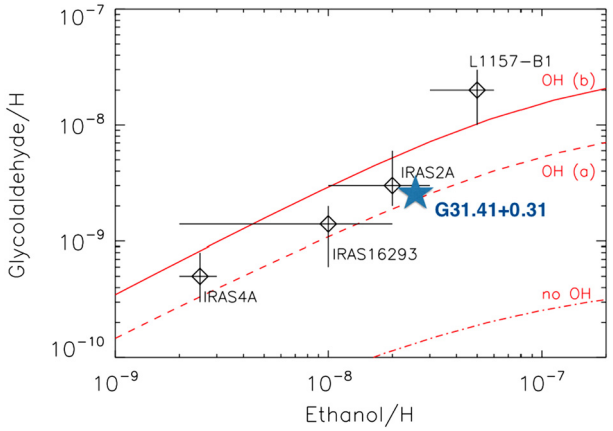

Fig. 13. Abundance of glycolaldehyde plotted against the abundance of ethanol for five different astrochemical objects. The most recent detection is highlighted by a blue star [206].

molecules can be desorbed thermally, this process being known as depletion of molecules. Once depleted, molecules do not stay on grains, but can lead to reactions either in the gas-phase or with a species still adsorbed (gas-grain chemistry). Gas-grain chemical models have been developed to explicitly treat the chemical kinetics of both the gas phase and the grain/ice surface (see, for example, Ref. [208]). Even if our focus is mainly on gas-phase reactions, a brief outline of the methodology for reactions involving dust grains is provided.

To react on grains, molecules and atoms need to move on the surface to encounter a partner to react. The model largely employed to describe species on the grain surface is the Langmuir-Hinshelwood mechanism [209]; in a simple way, it can be said that the dust grain plays the role of a third body in the reaction, thus acting as a catalyst, with an unoccupied site on the surface being treated as an effective chemical reacting system. To study possible gas-grain formation mechanisms, the starting point is the correct and reliable description of the reacting partners, i.e. the molecule (or radical) adsorbed on the grain interacting with the molecule (or radical) in the gas phase. In order to assure accurate and reliable results, a "multilayer" QM/QM′ (QM standing for quantum mechanics) ONIOM approach has to be employed [210,211]: the reactive system (i.e. the reacting species plus the water molecules of the grain mantle directly involved) is treated with an accurate and expensive QM method, while the remaining environment part is treated with a less accurate QM′ method. To further improve the description of environmental effects, in particular to incorporate those due to the grain-mantle molecules not included in the ONIOM treatment, the polarizable continuum model (PCM) [212] can be used, which provides a good compromise between accuracy and computational cost.

Subsequently, the energy of the stationary points located on the reactive PES might be further improved (with respect to the QM method initially considered) using composite schemes, according to the dimension of the system under consideration. Finally, reaction rate constants can be calculated using *ad hoc* kinetic codes, with the rate constants of the elementary steps being calculated as described above.

A recent example concerns the formation of formamide, which – as already mentioned – is currently debated. In Ref. [213], the possible formation routes in water-rich amorphous ices, simulated by considering a 33-water molecule cluster, has been presented. In addition to the well-established scenario of the radical-radical association reaction between NH_2 and HCO, the authors also considered the reaction of the CN radical with the water molecules of the ice mantle, which can lead through multiple steps to the formation of formamide, thus suggesting that the H_2O molecules of the ice mantle, usually considered as an inert support, can instead react with active radicals to form more complex species and can also act as catalysts by helping H transfer processes. Interestingly, in the 10-50 K temperature range considered, the authors obtained a unimolecular rate constant of $\sim 1.9 \times 10^9$ s^{-1} for the $H_2O + CN \rightarrow NH_2CO \rightarrow NH_2CHO$ reaction with an almost complete lack of temperature dependence. Such a rate coefficient implies an average reactive time of 0.5 ns, which is quite smaller than the "typical" ones for gas-phase processes (in the ps and sub-ps range). However, the calculated rate coefficient was found model-dependent, thus underlying the importance of the choice of the kinetic model to apply.

3.5. Endogenous theory: prebiotic chemistry in Titan's atmosphere

Titan has long been a subject of interest because it provides an excellent example of abiotic processing of organic material. In addition to the interest as a planetary subject, the organic chemistry on Titan has attracted the interest as a possible analog of the early Earth [214,215]. Therefore, the investigation of its atmosphere might shed light on the organic evolution occurred in the atmosphere of early Earth.

The chemistry taking place in Titan's atmosphere is rich and varied. The ingredients are N_2 (~98%), methane (~2%) and other hydrocarbons, like ethane, that are subjected in the upper atmosphere to UV radiation and energetic particles. The result is the production of small radicals, like e.g. N and CH_3, and non-classical carbocations, like e.g. CH_4^+ and CH_5^+. These radicals and cations then further react with neutral species, like hydrocarbons.

To give an example, according to Ref. [216], the $N(^2D)$ radical can react with ethane leading to the formation of the intermediate CH_3CH_2NH, which is quite stable with respect to both reactants and dissociation products. However, the reaction further proceeds via either a direct fission to the products or isomerization to other intermediates, which, in turn, fragment into other products. The fragmentation of a bond of the initial intermediate leads to methanimine (CH_2NH) and the CH_3 radical, which can give rise to other reactions, or $CH_2NH + CH_3$ if the C=N double bond is not formed. If one of the pre-existing C–H bonds of the CH_2 group not directly involved in the insertion process breaks apart, then ethanimine ($CH_3CH{=}NH$) and H are formed. Another possibility is offered if one of the pre-existing C–H bonds of the CH_3 group, also not directly involved in the insertion process, breaks apart, with the CH_2CH_2NH biradical being formed. Finally, if one of the new bonds, formed by the $N(^2D)$ insertion into a C-H bond, breaks apart, the formation of $NH + CH_3CH_2$ or $CH_3CH_2N + H$ occurs. All these possibilities were found to be both thermochemically and kinetically allowed.

Titan's atmosphere also contains unsaturated hydrocarbons, like acetylene, ethylene and butadiene. These are key species for the formation of PAHs and nitrogen-containing PAHs (NPAHs). The latter are important prebiotic species because they can be considered the missing link between nitrogen bearing acyclic molecules and prebiological systems, like nucleobases. Recently, Parker and Kaiser [217] reviewed different reaction pathways leading to their formation via gas phase radical mediated aromatization reactions. To give an example, pyridine can be formed from the reaction of vinyl cyanide, which has been unambiguously detected in Titan atmosphere [218], and its radical or via cyano radicals reacting with 1,3-butadiene. 1,4-dihydro(iso)quinoline and (iso)quinoline, which can be considered two examples of small NPAHs, can be synthesized through reaction of pyridyl radicals with 1,3-butadiene or sequentially with two acetylene molecules, respectively. Parker and Kaiser experimentally demonstrated the validity of aromatization reactions via neutral-neutral reactions to explain the chemical evolution from acyclic systems to polycyclic aromatic hydrocarbons and expanded this concept to the synthesis of NPAHs, also showing that these are *de facto* barrierless reactions and thus feasible at low temperatures [217].

The Ion Neutral Mass Spectrometer (INMS) and the Cassini Plasma Spectrometer (CAPS) instruments onboard the Cassini spacecraft discovered the presence of a large variety of carbocations and carbanions in Titan's upper

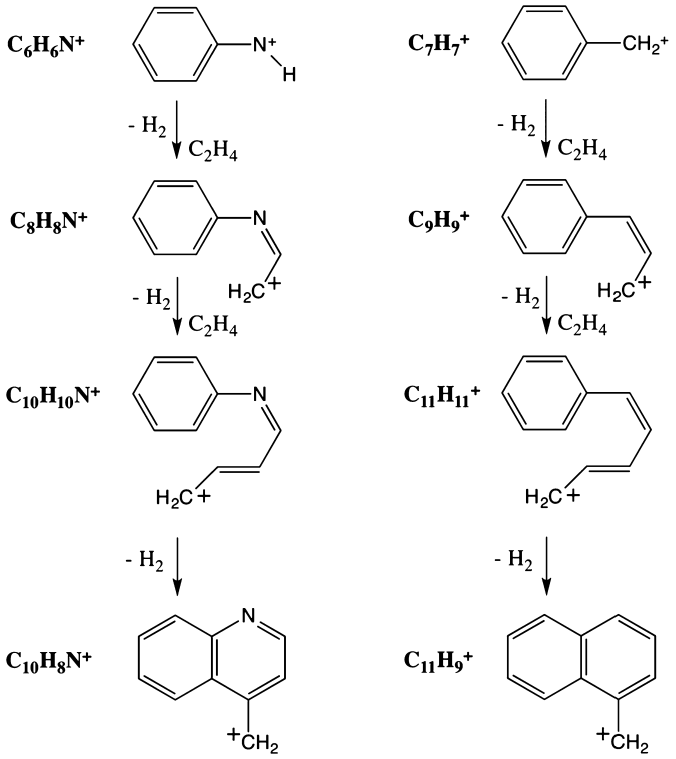

Fig. 14. Schemes displaying the progression of the chain reactions from $C_6H_5NH^+$ (left) and $C_6H_5CH_2^+$ (right) to $C_{10}H_8N^+$ and $C_{11}H_9^+$, respectively.

atmosphere, ranging in size from small species to heavy positive (up to 350 amu) and negative (up to 10,000 amu) ions (see, Ref. [16] and references therein). These numerous large carbocations and carbanions are postulated to play an important role in the chemical evolution observed there. In Ref. [16], the composition detected by the Cassini mass spectrometers was investigated and explained in terms of molecular structures and plausible reaction routes from simple molecules to complex polycyclic hydrocarbons, thus theorizing the formation of complex macromolecules. In view of understanding the chemical growth in Titan's atmosphere and the formation of its characteristic haze, particularly important are the reactions between aromatic cations, like $C_6H_5CH_2^+$ and $C_6H_5NH^+$ (C_6H_5 being the phenyl group; see Fig. 14), and hydrocarbons. The key steps of their reactions with ethylene molecules toward the condensation of a second aromatic ring are shown in Fig. 14, thus demonstrating how PAHs and NPAHs grow in size not only via neutral-neutral reactions, but also via ion-neutral reactions. The reaction of $C_6H_5NH^+$ (which is abundant in Titan's atmosphere [219]) with unsaturated hydrocarbons, such as ethylene and propene, introduces the imino R–C=N–R′ moiety into macromolecules, thus leading to the formation of prebiotic systems characterized by a certain degree of complexity. Even if the chemical composition of Titan's haze is still poorly understood [220], it is largely accepted that the macromolecules formed by organic chain reactions are the precursors of the nanoscale polymers responsible of it. The condensed-phase atmospheric macromolecular material then precipitates on Titan's surface and further evolves through the solute-solvent interaction chemistry.

The "chemical life" on Titan can be summarized as follows: formation of prebiotic species in Titan's upper atmosphere, further evolution in macromolecules and their assembling in a haze layer, precipitation and sedimentation of complex organics on its surface, and organic catalysis occurring in hydrocarbon lakes. The question that arises is then: had this sequence of endogenous steps any connection with what happened on the primitive Earth? Titan's atmosphere is relatively poor of oxygen compared to terrestrial planets, however the organic chemistry taking place on Titan and starting in its upper atmosphere is reminiscent of the groundbreaking Miller-Urey's experiment [14,15] that demonstrated the necessity of a reducing atmosphere for gas-phase organic synthesis to occur. Therefore, the subsequent question to address is whether photochemical production of complex molecules containing C, N, O, and H is possible in Titan's upper atmosphere. The oxygen chemistry could not be constrained by the instruments onboard the

Cassini spacecraft, this implying that the CO molecule cannot be mass resolved from the major species N_2, nor can atomic oxygen be resolved from the second most abundant species, CH_4. Similarly, for ionic species, H_2O^+ cannot be distinguished from NH_4^+. On the other hand, the Ion Mass Spectrometer (IMS of the CAPS) measurements allowed for the definitive detection of O^+ in Titan's upper atmosphere [221], with recent photochemical models suggesting that oxygen ions are then incorporated into CO and CO_2 [222]. Although Titan's physical-chemical conditions, such as temperature, pressure and UV radiation flux, cannot be entirely reproduced in the laboratory simulation experiments, Hörst et al. [223] demonstrated that important prebiotic species could be formed in Titan's atmosphere at high altitudes: aerosols produced in an experiment simulating Titan's atmosphere with an enhanced amount of CO (i.e. $N_2/CH_4/CO$ gas mixtures of 96.2%/2.0%/1.8% and 93.2%/5.0%/1.8%) were found to contain amino acids and nucleotide bases (like cytosine, uracil, thymine, guanine, glycine, and alanine). On a general ground, protonated oxides of alkenes in Titan's upper atmosphere are of great interest because they may initiate a completely new chemical pathway for complex prebiotic chemistry.

4. Computer simulations, big data and virtual reality as new tools for astrochemistry

Large dimensional and complex data sets are likely to be generated in astrochemical studies. A detailed yet compact representation of molecular structures and the inclusion of their properties in formulas and graphs have always been at the heart of chemistry. Since its inception, computer molecular graphics has demonstrated its power as an investigation tool in molecular sciences, constituting itself as an extension of chemical language. In addition, researchers make an increasingly extensive use of computational methods to model and forecast the properties of a great variety of systems over a wide range of space and time domains. However, without proper graphical representations (ball-and-stick cartoons, ribbons, molecular orbitals), such simulations would provide very little scientific insight to the users, these being just a sheer amount of numbers. On the other hand, the massive increase growth of computer graphics technologies for three-dimensional Immersive Virtual Reality (IVR) makes it possible to achieve a further evolution.

The main possibilities created by the application IVR technologies for astrochemical data are: (i) the ability to close the gap between human perception and molecular world: this goal can be achieved 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. VR tools include a vast array: from cheap consumer grade ones (interactive sensors, immersive helmets or force-feedback devices) to very costly specialized hardware (such as CAVE3D installations) [224].

Therefore, the development of powerful molecular viewers purposely tailored to astrochemical applications could be of particular interest 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. Several well-known molecular graphics editors have been provided recently with experimental plug-ins to be used with immersive or interactive technologies. However, due to limits in the standardization of the underlying hardware and application program interfaces, dominant paradigms have yet to emerge in the exploitation of such technologies. The choice to develop new unified environments, specifically engineered for new technologies, or extend current ones is very important and is likely to affect the real success of such applications once the hardware becomes standardized. An example is offered by the Caffeine software, which is being developed to exploit current generation IVR technologies [225]. It allows to create, model and save molecular structures and trajectories, loading data from the most diffuse file formats or dedicated databases and is particularly suited to be used in CAVE3D installations, interactively and by means of a mobile user interface for tablets/smartphones.

Finally, it has to be stressed that IVR can have a revolutionary and fruitful application 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. 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 line profiles) spread over very large bandwidths (e.g. more than 50 GHz at the IRAM telescopes). A systematic analysis 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.

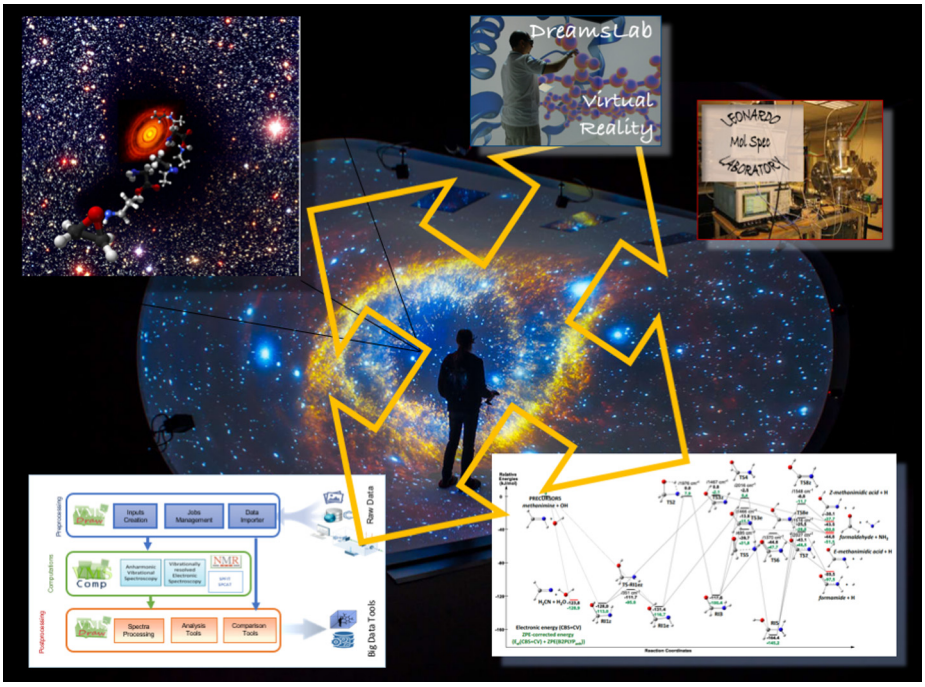

Fig. 15. Virtual reality promises a new perception and integration of different aspects of astrochemistry including astronomical observations, as well as in vitro and in silico experiments.

Along these lines Fig. 15 sketches a possible scenario in which IVR plays a central role in integrating astronomical observations, laboratory experiments and in-silico simulations toward a comprehensive analysis of key problems in astrochemistry.

5. A four-pillar approach to disclose the secrets of the interstellar space

The traditional "legs" (or "pillars") of the scientific methodology were theory and experiment. This paradigm evolved at the end of the 20th century and the beginning of the 21st century, which witnessed the birth of the "third pillar" of scientific inquiry (together with theory and experimentation): computational science. This has been enabling researchers to build and test models of complex phenomena like multi-century climate shifts, multidimensional flight stresses on aircraft, and stellar explosions. Recently, this "leg" has been complemented by a "fourth paradigm": the usage of advanced computing capabilities to manipulate and explore massive datasets (Fig. 16). For example, the decoding of the human genome in 2001 was a triumph of large-scale data analysis. Nowadays, we can state that science has four legs, and two of them are computational. It is in this general context that a four-pillar approach to astrochemistry and astrobiology can be envisaged with the main objective of providing the scientific community with an effective cyber-infrastructure managed by dedicated workflow systems. By integrating computational simulations, experimental data, external heterogeneous sources and augmented reality/virtual reality (AR/VR) facilities, the four-pillar approach has the final aim of guiding users in finding answers to still open key questions and, being flexible enough, to new ones. To make the computational part of the platform as general as possible (new spectroscopic signatures, more reliable results for large systems, effective evaluation of multi-channel reactions, environmental effects, etc.), several theoretical/computational developments are still needed. At the same time, specialized data centers, for high-performance computing (HPC), and integrated laboratories, for experimental molecular spectroscopy, must be developed and fully integrated into general platforms, also by means of AR/VR facilities.

The HPC facilities required for computer simulations, ranging from multi-scale approaches to molecular sciences, are nowadays accessible in various ways together with a massive amount of data. While automatic data analysis procedures are commonly employed to filter the relevant information and to derive the target quantities, new opportunities for a deeper understanding of the underlying physical-chemical properties are provided by VR. As mentioned

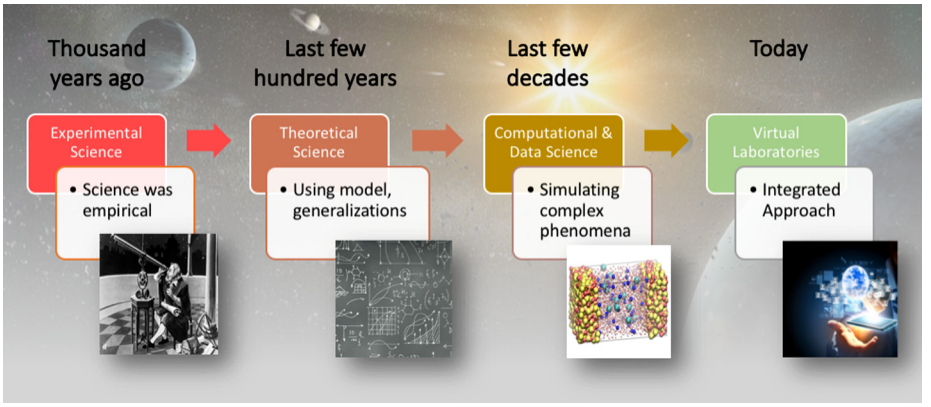

Fig. 16. Ages of scientific research.

in the previous section, the latter allows one to exploit the capabilities of the human visual system to readily identify structures, patterns and anomalies, so as to stimulate the mental process of insights and understanding about large sets of data and complex phenomena. Thanks to the development of accessible and powerful hardware and to the increased maturity and flexibility of the related software, immersive VR technologies can be profitably exploited also for scientific purposes.

A crucial element of virtual laboratories is a workflow system built on top of the middleware layer which provides a visual programming front-end. The above background shows that time is now ripe to take a step further, and to move from a passive computational machinery to a pro-active tool able to assist scientists in performing simulations and making decisions. In the context of this review, this goal can be achieved by cyber-infrastructures dedicated to astrochemistry and astrobiology.

In our opinion, it is currently possible to develop integrated platforms for astrochemistry with the aim of: (*i*) developing new theoretical methods for spectroscopic, kinetic and astronomical applications, (*ii*) integrating them with top notch experimental data in all applications mentioned above, and (*iii*) coupling theory and experiment into novel big-data approaches in an effective cyber-infrastructure. A first integrated tool in this direction is the, already mentioned, VMS software, which has been developed in the framework of molecular spectroscopy [94,226,227]. However, its capabilities need to be enhanced and extended to kinetic aspects integrated with new developments in computational spectroscopy, multi-scale methods, and machine learning procedures. Data visualization and manipulation can also be made available through AR/VR tools, beyond the proof of concept stage, usable on a daily basis and able to tap into human pattern recognition skills. To this end, it is necessary to implement a set of 3D interactive applications embedded within the workflow system, also employing state-of-the-art molecular graphics techniques.

6. Concluding remarks

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 and toward the origin of life elsewhere in space. The complexity of the problem 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. To take an important step forward, it is necessary to explore the molecular universe with an understanding of all of the local molecular environments and disclose possible chemical reactions using the concepts of physical chemistry. The chemistry of the universe is a remarkably rich subject that may have a bearing on the formation of life on Earth and elsewhere. Particularly fascinating amongst the molecules discovered in the interstellar medium are the aCOMs that combines multiple functional groups illustrative of the rich chemistry of the ISM. While it is now widely recognized that there is a diverse and complex chemistry occurring in space, it is not at all clear how to answer two fundamental questions: how is it that these species are produced? How do they evolve toward biological building-blocks? Indeed, these issues are of remarkable significance, especially since many of the molecules detected in space play a role in the chemistry of life. This aspect of astrochemistry has attracted intense interest in the general scientific community and spurred the search for biomolecular building blocks in interstellar space. Intriguingly, evidence has been presented to suggest that nucle-

obases (the fundamental elements of RNA and DNA) and amino acids have been carried through space on meteorites and exist in other extraterrestrial sources. In addition to the identification of prebiotic species in space, understanding the underlying chemical processes, including the production, reactions and destruction of compounds, is one of the main aims of astrochemistry. Gaining deeper insights into these issues can help to address some of the most intriguing questions in all of science – how did life emerge on primitive Earth? Is it possible that it has also happened elsewhere in the Universe? – as these surely help to set the stage for the emergence of life to occur, both on Earth and elsewhere in the cosmos.

The previous paragraphs tried to set the scene of the complex and still mysterious chemical evolution occurring in the universe and explained which spirit and which aims are at the basis of this review. The chemical evolution across space has been addressed from different points of view: from the detection of chemical species in the ISM and planetary environments to their formation pathways, from spectroscopy to thermochemistry and kinetics, from quantum-chemical computations to astronomical studies. Indeed, the aim of this review is to provide some insight into the secrets of chemical evolution in space and to present an integrated approach to disclose them. The envisaged strategy involves different branches of sciences: it requires a strong interplay between experiment and theory, between chemistry and physics (with a flavor of biology), also exploiting the great potentialities of high-performance computers and virtual reality for a reconciliation of accuracy and interpretability of complex phenomena [228].

Acknowledgements

This work has been supported by MIUR “PRIN 2015” funds (Grant Number 2015F59J3R).

References

[1] Hollis JM, Vogel SN, Snyder LE, Jewell PR, Lovas FJ. The spatial scale of glycolaldehyde in the Galactic center. Astrophys J 2001;554:L81–5.
[2] Hollis JM, Lovas FJ, Remijan AJ, Jewell PR, Ilyushin VV, Kleiner I. Detection of acetamide (CH_3CONH_2): the largest interstellar molecule with a peptide bond. Astrophys J 2006;643:L25–8.
[3] Tercero B, Kleiner I, Cernicharo J, Nguyen HVL, López A, Caro GMM. Discovery of methyl acetate and gauche ethyl formate in Orion. Astrophys J 2013;770:L13.
[4] McGuire BA, Carroll PB, Loomis RA, Finneran IA, Jewell PR, Remijan AJ, et al. Discovery of the interstellar chiral molecule propylene oxide (CH_3CHCH_2O). Science 2016;352:1449–52.
[5] McGuire BA, Burkhardt AM, Kalenskii S, Shingledecker CN, Remijan AJ, Herbst E, et al. Detection of the aromatic molecule benzonitrile (c-C_6H_5CN) in the interstellar medium. Science 2018;359:202–5.
[6] Müller HSP, Thorwirth S, Roth DA, Winnewisser G. The Cologne Database for Molecular Spectroscopy, CDMS. Astron Astrophys 2001;370:L49–52.
[7] Müller HSP, Schlöder F, Stutzki J, Winnewisser G. The Cologne Database for Molecular Spectroscopy, CDMS: a useful tool for astronomers and spectroscopists. J Mol Struct 2005;742:215–27.
[8] Endres CP, Schlemmer S, Schilke P, Stutzki J, Müller HS. The Cologne Database for Molecular Spectroscopy, CDMS, in the virtual atomic and molecular data centre, VAMDC. J Mol Spectrosc 2016;327:95–104.
[9] Garrod RT. A three-phase chemical model of hot cores: the formation of glycine. Astrophys J 2013;765:60.
[10] Jimenez-Serra I, Testi L, Caselli P, Viti S. Detectability of glycine in solar-type system precursors. Astrophys J Lett 2014;787:L33.
[11] Callahan MP, Smith KE, Cleaves HJ, Ruzicka J, Stern JC, Glavin DP, et al. Carbonaceous meteorites contain a wide range of extraterrestrial nucleobases. Proc Natl Acad Sci 2011;108:13995–8.
[12] Elsila JE, Glavin DP, Dworkin JP. Cometary glycine detected in samples returned by Stardust. Meteorit Planet Sci 2009;44:1323–30.
[13] Chyba C, Sagan C. Endogenous production, exogenous delivery and impact-shock synthesis of organic molecules: an inventory for the origins of life. Nature 1992;355:125.
[14] Miller SL, Urey HC. Organic compound synthesis on the primitive Earth. Science 1959;130:245.
[15] Miller SL. Production of amino acids under possible primitive Earth conditions. Science 1953;117:528.
[16] Ali A, Sittler Jr E, Chornay D, Rowe B, Puzzarini C. Organic chemistry in Titan’s upper atmosphere and its astrobiological consequences, I: views towards Cassini plasma spectrometer (CAPS) and ion neutral mass spectrometer (INMS) experiments in space. Planet Space Sci 2015;109–110:46–63.
[17] Olah GA, Mathew T, Prakash GKS. Relevance and significance of extraterrestrial abiological hydrocarbon chemistry. J Am Chem Soc 2016;138:6905–11.
[18] Saladino R, Botta G, Pino S, Costanzo G, Di Mauro E. Genetics first or metabolism first? The formamide clue. Chem Soc Rev 2012;41:5526–65.
[19] Saladino R, Carota E, Botta G, Kapralov M, Timoshenko GN, Rozanov AY, et al. Meteorite-catalyzed syntheses of nucleosides and of other prebiotic compounds from formamide under proton irradiation. Proc Natl Acad Sci 2015;112:E2746–55.

[20] Barone V, Biczysko M, Puzzarini C. Quantum chemistry meets spectroscopy for astrochemistry: increasing complexity toward prebiotic molecules. Acc Chem Res 2015;48:1413–22.

[21] Tennyson J, editor. Astronomical spectroscopy. Imperial College Press; 2005.

[22] Yamamoto S, editor. Introduction to astrochemistry: chemical evolution from interstellar clouds to star and planet formation. Springer; 2017.

[23] Zaleski DP, Seifert NA, Steber AL, Muckle MT, Loomis RA, Corby JF, et al. Detection of E-cyanomethanimine toward Sagittarius B2(N) in the Green Bank Telescope PRIMOS survey. Astrophys J 2013;765:L10.

[24] Belloche A, Garrod RT, Müller HSP, Menten KM. Detection of a branched alkyl molecule in the interstellar medium: iso-propyl cyanide. Science 2014;345:1584–7.

[25] McKellar A. Evidence for the molecular origin of some hitherto unidentified interstellar lines. Publ Astron Soc Pac 1940;52:187.

[26] Cernicharo J, Daniel F, Castro-Carrizo A, Agundez M, Marcelino N, Joblin C, et al. Unveiling the dust nucleation zone of IRC+10216 with ALMA. Astrophys J 2013;778:L25.

[27] McCall BJ, Griffin RE. On the discovery of the diffuse interstellar bands. Proc R Soc A 2013;469:20120604.

[28] Campbell EK, Holz M, Gerlich D, Maier JP. Laboratory confirmation of C_{60}^+ as the carrier of two diffuse interstellar bands. Nature 2015;523:322.

[29] McGuire BA, Shingledecker CN, Willis ER, Burkhardt AM, El-Abd S, Motiyenko RA, et al. ALMA detection of interstellar methoxymethanol (CH_3OCH_2OH). Astrophys J 2017;851:L46.

[30] Sloan GC, Kraemer KE, Price SD, Shipman RF. A uniform database of 2.4–45.4 micron spectra from the Infrared Space Observatory Short Wavelength Spectrometer. Astrophys J Suppl Ser 2003;147:379–401.

[31] 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;785:107.

[32] Grimme S. Semiempirical hybrid density functional with perturbative second-order correlation. J Chem Phys 2006;124:034108.

[33] Knauth DC, Andersson B-G, McCandliss SR, Warren Moos H. The interstellar N_2 abundance towards HD 124314 from far-ultraviolet observations. Nature 2004;429:636.

[34] Clairemidi J, Moreels G, Mousis O, Bréchignac P. Identification of anthracene in Comet 1P/Halley. Astron Astrophys 2008;492:245–50.

[35] Sewiło M, Indebetouw R, Charnley SB, Zahorecz S, Oliveira JM, van Loon JT, et al. The detection of hot cores and complex organic molecules in the large Magellanic Cloud. Astrophys J 2018;853:L19.

[36] Puzzarini C, Stanton JF, Gauss J. Quantum-chemical calculation of spectroscopic parameters for rotational spectroscopy. Int Rev Phys Chem 2010;29:273–367.

[37] Biczysko M, Bloino J, Puzzarini C. Computational challenges in astrochemistry. WIREs Comput Mol Sci 2018;8:e1349.

[38] 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;810:111.

[39] 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;12:5385–97.

[40] Skouteris D, Balucani N, Ceccarelli C, Vazart F, Puzzarini C, Barone V, et al. The genealogical tree of ethanol: gas-phase formation of glycolaldehyde, acetic acid, and formic acid. Astrophys J 2018;854:135.

[41] Lattelais M, Pauzat F, Ellinger Y, Ceccarelli C. Interstellar complex organic molecules and the minimum energy principle. Astrophys J 2009;696:L133–6.

[42] Lattelais M, Pauzat F, Ellinger Y, Ceccarelli C. A new weapon for the interstellar complex organic molecule hunt: the minimum energy principle. Astron Astrophys 2010;519:A30.

[43] Raghavachari K, Trucks GW, Pople JA, Head-Gordon M. A fifth-order perturbation comparison of electron correlation theories. Chem Phys Lett 1989;157:479–83.

[44] Császár AG, Allen WD, Schaefer III HF. In pursuit of the ab initio limit for conformational energy prototypes. J Chem Phys 1998;108:9751–64.

[45] Tajti A, Szalay PG, Császár AG, Kállay M, Gauss J, Valeev EF, et al. Heat: high accuracy extrapolated ab initio thermochemistry. J Chem Phys 2004;121:11599–613.

[46] Heckert M, Kállay M, Gauss J. Molecular equilibrium geometries based on coupled-cluster calculations including quadruple excitations. Mol Phys 2005;103:2109.

[47] Heckert M, Kállay M, Tew DP, Klopper W, Gauss J. Basis-set extrapolation techniques for the accurate calculation of molecular equilibrium geometries using coupled-cluster theory. J Chem Phys 2006;125:044108.

[48] Puzzarini C, Heckert J, Gauss J. The accuracy of rotational constants predicted by high-level quantum-chemical calculations, I: molecules containing first-row atoms. J Chem Phys 2008;128:194108.

[49] Boese AD, Oren M, Atasoylu O, Martin JML, Kállay M, Gauss J. W3 theory: robust computational thermochemistry in the kJ/mol accuracy range. J Chem Phys 2004;120:4129–41.

[50] Karton A, Rabinovich E, Martin JML, Ruscic B. W4 theory for computational thermochemistry: in pursuit of confident sub-kJ/mol predictions. J Chem Phys 2006;125:144108.

[51] Dixon DA, Feller D, Peterson KA. Chapter one – A practical guide to reliable first principles computational thermochemistry predictions across the periodic table. Annu Rep Comput Chem 2012;8:1–28.

[52] Stanton JF, Gauss J, Harding ME, Szalay PG. CFOUR: a quantum chemical program package. with contributions from A.A. Auer, R.J. Bartlett, U. Benedikt, C. Berger, D.E. Bernholdt, Y.J. Bomble, O. Christiansen, F. Engel, M. Heckert, O. Heun, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, K. Klein, W.J. Lauderdale, F. Lipparini, D. Matthews, T. Metzroth, L.A. Mück, D.P. O'Neill, D.R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J.D. Watts and the integral packages MOLECULE (J. Almlöf and P.R. Taylor), PROPS (P.R. Taylor), ABACUS (T. Helgaker, H.J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A.V. Mitin and C. van Wüllen. For the current version, see http://www.cfour.de, 2016.

[53] Noga J, Bartlett RJ. The full CCSDT model for molecular electronic structure. J Chem Phys 1987;86:7041–50.
[54] Scuseria GE, Schaefer III HF. A new implementation of the full CCSDT model for molecular electronic-structure. Chem Phys Lett 1988;152:382–6.
[55] Watts JD, Bartlett RJ. The Coupled-Cluster single, double, and triple excitation model for open-shell single reference functions. J Chem Phys 1990;93:6104–5.
[56] Kállay M, Surján PR. Higher excitations in Coupled-Cluster theory. J Chem Phys 2001;115:2945–54.
[57] Dunning Jr TH. Gaussian basis sets for use in correlated molecular calculations, I: the atoms boron through neon and hydrogen. J Chem Phys 1989;90:1007–23.
[58] Woon DE, Dunning Jr TH. Gaussian basis sets for use in correlated molecular calculations, V: core-valence basis sets for boron through neon. J Chem Phys 1995;103:4572.
[59] Wilson AK, van Mourik T, Dunning Jr TH. Gaussian basis sets for use in correlated molecular calculations, VI: sextuple zeta correlation consistent basis sets for boron through neon. J Mol Struct, Theochem 1996;388:339–49.
[60] Peterson KA, Dunning Jr TH. 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;117:10548–60.
[61] 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;98:7059.
[62] Helgaker T, Klopper W, Koch H, Noga J. Basis-set convergence of correlated calculations on water. J Chem Phys 1997;106:9639.
[63] Puzzarini C. Rotational spectroscopy meets theory. Phys Chem Chem Phys 2013;15:6595–607.
[64] Thorwirth S, Mück LA, Gauss J, Tamassia F, Lattanzi V, McCarthy MC. Silicon oxysulfide, OSiS: rotational spectrum, quantum-chemical calculations, and equilibrium structure. J Phys Chem Lett 2011;2:1228.
[65] Barone V, Biczysko M, Bloino J, Puzzarini C. Glycine conformers: a never-ending story? Phys Chem Chem Phys 2013;15:1358–63.
[66] Barone V, Biczysko M, Bloino J, Puzzarini C. Accurate molecular structures and infrared spectra of trans-2,3-dideuterooxirane, methyloxirane, and trans-2,3-dimethyloxirane. J Chem Phys 2014;141:034107.
[67] 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;11:4342–63.
[68] Barone V, Biczysko M, Bloino J, Puzzarini C. Accurate structure, thermodynamic and spectroscopic parameters from CC and CC/DFT schemes: the challenge of the conformational equilibrium in glycine. Phys Chem Chem Phys 2013;15:10094–111.
[69] Bomble YJ, Vázquez J, Kállay M, Michauk C, Szalay PG, Császár AG, et al. High-accuracy extrapolated ab initio thermochemistry, II: minor improvements to the protocol and a vital simplification. J Chem Phys 2006;125:064108.
[70] Harding ME, Vázquez J, Ruscic B, Wilson AK, Gauss J, Stanton JF. High-accuracy extrapolated ab initio thermochemistry, III: additional improvements and overview. J Chem Phys 2008;128:114111.
[71] Puzzarini C. Accurate thermochemistry and spectroscopy of the oxygen-protonated sulfur dioxide isomers. Phys Chem Chem Phys 2011;13:21319–27.
[72] Puzzarini C, Penocchio E, Biczysko M, Barone V. Molecular structure and spectroscopic signatures of acrolein: theory meets experiment. J Phys Chem A 2014;118:6648.
[73] Puzzarini C. Isomerism of cyanomethanimine: accurate structural, energetic, and spectroscopic characterization. J Phys Chem A 2015;119:11614–22.
[74] Puzzarini C, Biczysko M, Barone V, Largo L, Peña I, Cabezas C, et al. 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;5:534–40.
[75] Puzzarini C, Biczysko M. Microsolvation of 2-thiouracil: molecular structure and spectroscopic parameters of the thiouracil-water complex. J Phys Chem A 2015;119:5386–95.
[76] Møller C, Plesset MS. Note on an approximation treatment for many-electron systems. Phys Rev 1934;46:618–22.
[77] Spada L, Tasinato N, Vazart F, Barone V, Caminati W, Puzzarini C. Noncovalent interactions and internal dynamics in pyridine-ammonia: a combined quantum-chemical and microwave spectroscopy study. Chem Eur J 2017;23:4876–83.
[78] Martin JM, Taylor PR. The geometry, vibrational frequencies, and total atomization energy of ethylene: a calibration study. Chem Phys Lett 1996;248:336–44.
[79] Demaison J, Margulès L, Boggs JE. The equilibrium C–Cl, C–Br, and C–I bond lengths from ab initio calculations, microwave and infrared spectroscopies, and empirical correlations. Struct Chem 2003;14:159–74.
[80] Puzzarini C. Extrapolation to the complete basis set limit of structural parameters: comparison of different approaches. J Phys Chem A 2009;113:14530–5.
[81] Puzzarini C, Barone V. Extending the molecular size in accurate quantum-chemical calculations: the equilibrium structure and spectroscopic properties of uracil. Phys Chem Chem Phys 2011;13:7189–97.
[82] Puzzarini C, Biczysko M, Barone V, Peña I, Cabezas C, Alonso JL. Accurate molecular structure and spectroscopic properties of nucleobases: a combined computational-microwave investigation of 2-thiouracil as a case study. Phys Chem Chem Phys 2013;15:16965–75.
[83] Puzzarini C, Barone V. Diving for accurate structures in the ocean of molecular systems with the help of spectroscopy and quantum chemistry. Acc Chem Res 2018;51:548–56.
[84] Puzzarini C, Tasinato N, Bloino J, Spada L, Barone V. State-of-the-art computation of the rotational and IR spectra of the methyl-cyclopropyl cation: hints on its detection in space. Phys Chem Chem Phys 2019;21:3432.
[85] Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X, Caricato M, Marenich AV, Bloino J, Janesko BG, Gomperts R, Mennucci B, Hratchian HP, Ortiz JV, Izmaylov AF, Sonnenberg JL, Williams-Young D, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski VG, Gao J, Rega N, Zheng G, Liang W, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark MJ, Heyd JJ, Brothers EN, Kudin KN, Staroverov VN, Keith TA, Kobayashi R, Normand J,

Raghavachari K, Rendell AP, Burant JC, Iyengar SS, Tomasi J, Cossi M, Millam JM, Klene M, Adamo C, Cammi R, Ochterski JW, Martin RL, Morokuma K, Farkas O, Foresman JB, Fox DJ. Gaussian 16 revision A.03. Wallingford, CT: Gaussian Inc.; 2016.
[86] Becke AD. Density–functional thermochemistry, III: the role of exact exchange. J Chem Phys 1993;98:5648–52.
[87] Lee C, Yang W, Parr RG. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys Rev B 1988;37:785–9.
[88] Double triple-ζ basis sets of the SNS family are available in the download section. last visited: 3 February 2019 http://smart.sns.it/.
[89] Bloino J, Biczysko M, Barone V. General perturbative approach for spectroscopy, thermodynamics, and kinetics: methodological background and benchmark studies. J Chem Theory Comput 2012;8:1015–36.
[90] Tielens AGGM. The molecular universe. Rev Mod Phys 2013;85:1021–81.
[91] Barone V, editor. Computational strategies for spectroscopy: from small molecules to nano systems. John Wiley & Sons, Inc.; 2011.
[92] Cazzoli G, Lattanzi V, Kirsch T, Gauss J, Tercero B, Cernicharo J, et al. Laboratory measurements and astronomical search for the HSO radical. Astron Astrophys 2016;591:A126.
[93] Obenchain DA, Spada L, Alessandrini S, Rampino S, Herbers S, Tasinato N, et al. Unveiling the sulfur-sulfur bridge: accurate structural and energetic characterization of a homo chalcogen inter-molecular bond. Angew Chem, Int Ed Engl 2018;57:15822–6.
[94] Barone V. The Virtual Multifrequency Spectrometer: a new paradigm for spectroscopy. WIREs Comput Mol Sci 2016;6:86–110.
[95] Fortenberry RC. Quantum astrochemical spectroscopy. Int J Quant Chem 2017;117:81–91.
[96] Fornaro T, Brucato JR, Feuillie C, Sverjensky DA, Hazen RM, Brunetto R, et al. Binding of nucleic acid components to the serpentinite-hosted hydrothermal mineral brucite. Astrobiology 2018;18:989–1007.
[97] Najbauer EE, Bazso G, Apostolo R, Fausto R, Biczysko M, Barone V, et al. Identification of serine conformers by matrix-isolation IR spectroscopy aided by near-infrared laser-induced conformational change, 2D correlation analysis, and quantum mechanical anharmonic computations. J Phys Chem B 2015;119:10496–510.
[98] Barone V, Bellina F, Biczysko M, Bloino J, Fornaro T, Latouche C, et al. Toward the design of alkynylimidazole fluorophores: computational and experimental characterization of spectroscopic features in solution and in poly(methyl methacrylate). Phys Chem Chem Phys 2015;17:26710–23.
[99] Muniz-Miranda F, Pedone A, Battistelli G, Montalti M, Bloino J, Barone V. Benchmarking TD-DFT against vibrationally resolved absorption spectra at room temperature: 7-aminocoumarins as test cases. J Chem Theory Comput 2015;11:5371–84.
[100] Patti A, Pedotti S, Mazzeo G, Longhi G, Abbate S, Paoloni L, et al. Ferrocenes with simple chiral substituents: an in-depth theoretical and experimental VCD and ECD study. Phys Chem Chem Phys 2019;21:9419–32.
[101] Berova N, Polavarapu PL, Nakanishi K, Woody RW, editors. Comprehensive chiroptical spectroscopy: instrumentation, methodologies, and theoretical simulations. Hoboken, New Jersey: John Wiley & Sons, Inc.; 2012.
[102] Allen WD, East ALL, Császár AG. Structures and conformations of non-rigid molecules. Dordrecht: Kluwer; 1993.
[103] Wang W, Donini O, Reyes CM, Kollman PA. Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions. Annu Rev Biophys Biomol Struct 2001;30:211–43.
[104] Aliev MR, Watson JKG. Chapter 1 – Higher-order effects in the vibration–rotation spectra of semirigid molecules. In: Rao KN, editor. Molecular spectroscopy: modern research, vol. 3. Academic Press; 1985. p. 1–67.
[105] Watson JKG. Determination of centrifugal distortion coefficients of asymmetric-top molecules, III: sextic coefficients. J Chem Phys 1968;48:4517–24.
[106] Mills IM. Vibration-rotation structure in asymmetric- and symmetric-top molecules. In: Rao KN, Mathews CW, editors. Molecular spectroscopy: modern research. New York: Academic Press; 1972. p. 115–40. Ch. 3.2.
[107] Alessandrini S, Gauss J, Puzzarini C. Accuracy of rotational parameters predicted by high-level quantum-chemical calculations: case study of sulfur-containing molecules of astrochemical interest. J Chem Theory Comput 2018;14:5360–71.
[108] Piccardo M, Penocchio E, Puzzarini C, Biczysko M, Barone V. Semi-experimental equilibrium structure determinations by employing B3LYP/SNSD anharmonic force fields: validation and application to semirigid organic molecules. J Phys Chem A 2015;119:2058–82.
[109] Puzzarini C, Cazzoli G, Gauss J. The rotational spectra of $HD^{17}O$ and $D_2^{17}O$: experiment and quantum-chemical calculations. J Chem Phys 2012;137:154311.
[110] Melli A, Melosso M, Tasinato N, Bosi G, Spada L, Bloino J, et al. Rotational and infrared spectroscopy of ethanimine: a route toward its astrophysical and planetary detection. Astrophys J 2018;855:123.
[111] Watson JK. Simplification of the molecular vibration-rotation hamiltonian. Mol Phys 1968;15:479–90.
[112] Nielsen HH. The vibration-rotation energies of molecules. Rev Mod Phys 1951;23:90–136.
[113] Schuurman MS, Allen WD, von Ragué Schleyer P, Schaefer III HF. The highly anharmonic BH_5 potential energy surface characterized in the ab initio limit. J Chem Phys 2005;122:104302.
[114] Kuhler KM, Truhlar DG, Isaacson AD. General method for removing resonance singularities in quantum mechanical perturbation theory. J Chem Phys 1996;104:4664–70.
[115] Puzzarini C, Biczysko M, Barone V. Accurate anharmonic vibrational frequencies for uracil: the performance of composite schemes and hybrid CC/DFT model. J Chem Theory Comput 2011;7:3702–10.
[116] 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;6:828–38.
[117] Puzzarini C, Baiardi A, Bloino J, Barone V, Murphy TE, Drew HD, et al. Spectroscopic characterization of key aromatic and heterocyclic molecules: a route toward the origin of life. Astron J 2017;154:82.
[118] 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;16:1759–87.
[119] Bloino J. A VPT2 route to near-infrared spectroscopy: the role of mechanical and electrical anharmonicity. J Phys Chem A 2015;119:5269–87.

[120] Bloino J, Baiardi A, Biczysko M. Aiming at an accurate prediction of vibrational and electronic spectra for medium-to-large molecules: an overview. Int J Quant Chem 2016;116:1543–74.

[121] Krylov AI. Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: the Hitchhiker's guide to Fock space. Annu Rev Phys Chem 2008;59:433–62.

[122] Knecht S, Hedegard ED, Keller S, Kovyrshin A, Ma AMY, Stein CJ, et al. New approaches for ab initio calculations of molecules with strong electron correlation. Chimia, Int J Chem 2016;70:244–51.

[123] Burke K, Werschnik J, Gross EKU. Time-dependent density functional theory: past, present, and future. J Chem Phys 2005;123:062206.

[124] Grimme S, Neese F. Double-hybrid density functional theory for excited electronic states of molecules. J Chem Phys 2007;127:154116.

[125] Laurent AD, Jacquemin D. TD-DFT benchmarks: a review. Int J Quant Chem 2013;113:2019–39.

[126] Lakowicz JR, editor. Principles of fluorescence spectroscopy. Springer Verlag; 2006.

[127] Baryshnikov G, Minaev B, Ågren H. Theory and calculation of the phosphorescence phenomenon. Chem Rev 2017;117:6500–37.

[128] Carter S, Handy NC, Rosmus P, Chambaud G. A variational method for the calculation of spin-rovibronic levels of Renner-Teller triatomic molecules. Mol Phys 1990;71:605–22.

[129] Boudon V, Champion J-P, Gabard T, Loëte M, Michelot F, Pierre G, et al. Symmetry-adapted tensorial formalism to model rovibrational and rovibronic spectra of molecules pertaining to various point groups. J Mol Spectrosc 2004;228:620–34.

[130] Furche F, Ahlrichs R. Adiabatic time-dependent density functional methods for excited state properties. J Chem Phys 2002;117:7433–47.

[131] Scalmani G, Frisch MJ, Mennucci B, Tomasi J, Cammi R, Barone V. Geometries and properties of excited states in the gas phase and in solution: theory and application of a time-dependent density functional theory polarizable continuum model. J Chem Phys 2006;124:094107.

[132] Liu J, Liang W. Analytical approach for the excited-state hessian in time-dependent density functional theory: formalism, implementation, and performance. J Chem Phys 2011;135:184111.

[133] Luis JM, Bishop DM, Kirtman B. A different approach for calculating Franck-Condon factors including anharmonicity. J Chem Phys 2004;120:813–22.

[134] Luis JM, Kirtman B, Christiansen O. A variational approach for calculating Franck-Condon factors including mode-mode anharmonic coupling. J Chem Phys 2006;125:154114.

[135] Rodriguez-Garcia V, Yagi K, Hirao K, Iwata S, Hirata S. Franck-Condon factors based on anharmonic vibrational wave functions of polyatomic molecules. J Chem Phys 2006;125:014109.

[136] Berova N, Nakanishi K, Woody RW, editors. Circular dichroism: principles and applications. 2nd edition. John Wiley & Sons, Inc; 2000.

[137] Riehl JP, Richardson FS. Circularly polarized luminescence spectroscopy. Chem Rev 1986;86:1–16.

[138] Ciofini I, Daul CA. DFT calculations of molecular magnetic properties of coordination compounds. Coord Chem Rev 2003;238–239:187–209. Theoretical and Computational Chemistry.

[139] Helgaker T, Coriani S, Jørgensen P, Kristensen K, Olsen J, Ruud K. Recent advances in wave function-based methods of molecular-property calculations. Chem Rev 2012;112:543–631.

[140] Bloino J, Biczysko M, Barone V. Anharmonic effects on vibrational spectra intensities: infrared, Raman, vibrational circular dichroism, and Raman optical activity. J Phys Chem A 2015;119:11862–74.

[141] Lin N, Santoro F, Rizzo A, Luo Y, Zhao X, Barone V. Theory for vibrationally resolved two-photon circular dichroism spectra: application to (R)-(+)-3-methylcyclopentanone. J Phys Chem A 2009;113:4198–207.

[142] Santoro F, Barone V. Computational approach to the study of the lineshape of absorption and electronic circular dichroism spectra. Int J Quant Chem 2010;110:476–86.

[143] Bloino J, Biczysko M, Santoro F, Barone V. General approach to compute vibrationally resolved one-photon electronic spectra. J Chem Theory Comput 2010;6:1256–74.

[144] Santoro F, Cappelli C, Barone V. Effective time-independent calculations of vibrational resonance Raman spectra of isolated and solvated molecules including Duschinsky and Herzberg-Teller effects. J Chem Theory Comput 2011;7:1824–39.

[145] Baiardi A, Bloino J, Barone V. General time dependent approach to vibronic spectroscopy including Franck–Condon, Herzberg–Teller, and Duschinsky effects. J Chem Theory Comput 2013;9:4097–115.

[146] Baiardi A, Bloino J, Barone V. A general time-dependent route to resonance-Raman spectroscopy including Franck–Condon, Herzberg–Teller and Duschinsky effects. J Chem Phys 2014;141:114108.

[147] Baiardi A, Bloino J, Barone V. Time-dependent formulation of resonance Raman optical activity spectroscopy. J Chem Theory Comput 2018;14:6370–90.

[148] Egidi F, Fusè M, Baiardi A, Bloino J, Li X, Barone V. Computational simulation of vibrationally resolved spectra for spin-forbidden transitions. Chirality 2018;30:850–65.

[149] Baiardi A, Bloino J, Barone V. General formulation of vibronic spectroscopy in internal coordinates. J Chem Phys 2016;144:084114.

[150] Baiardi A, Bloino J, Barone V. Simulation of vibronic spectra of flexible systems: hybrid DVR-harmonic approaches. J Chem Theory Comput 2017;13:2804–22.

[151] Egidi F, Williams-Young DB, Baiardi A, Bloino J, Scalmani G, Frisch MJ, et al. Effective inclusion of mechanical and electrical anharmonicity in excited electronic states: VPT2-TDDFT route. J Chem Theory Comput 2017;13:2789–803.

[152] Yanai T, Tew DP, Handy NC. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). Chem Phys Lett 2004;393:51–7.

[153] Grimme S, Antony J, Ehrlich S, Krieg H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys 2010;132:154104.

[154] Grimme S. Density functional theory with London dispersion corrections. WIREs Comput Mol Sci 2011;1:211–28.

[155] Barone V, Cimino P. Accurate and feasible computations of structural and magnetic properties of large free radicals: the PBE0/N07D model. Chem Phys Lett 2008;454:139–43.

[156] Takayanagi M, Gejo T, Hanazaki I. Geometry and torsional potential of 2,2′-bithiophene in a supersonic jet. J Phys Chem 1994;98:12893–8.
[157] Longhi G, Castiglioni E, Abbate S, Lebon F, Lightner DA. Experimental and calculated cpl spectra and related spectroscopic data of camphor and other simple chiral bicyclic ketones. Chirality 2013;25:589–99.
[158] Rodger A, Norden B, editors. Circular dichroism and linear dichroism. Oxford University Press; 1997.
[159] Mennucci B. Polarizable continuum model. WIREs Comput Mol Sci 2012;2:386–404.
[160] Beyer T, Swinehart DF. Algorithm 448: number of multiply-restricted partitions. Commun ACM 1973;16:379.
[161] Stein SE, Rabinovitch BS. Accurate evaluation of internal energy level sums and densities including anharmonic oscillators and hindered rotors. J Chem Phys 1973;58:2438–45. https://doi.org/10.1063/1.1679522.
[162] Wang F, Landau DP. Efficient, multiple-range random walk algorithm to calculate the density of states. Phys Rev Lett 2001;86:2050–3.
[163] Zhou C, Bhatt RN. Understanding and improving the Wang-Landau algorithm. Phys Rev E 2005;72:025701.
[164] Basire M, Parneix P, Calvo F. Quantum anharmonic densities of states using the Wang–Landau method. J Chem Phys 2008;129:081101.
[165] Nguyen TL, Barker JR. Sums and densities of fully coupled anharmonic vibrational states: a comparison of three practical methods. J Phys Chem A 2010;114:3718–30.
[166] Skouteris D, Calderini D, Barone V. Methods for calculating partition functions of molecules involving large amplitude and/or anharmonic motions. J Chem Theory Comput 2016;12:1011–8.
[167] Karas AJ, Gilbert RG, Collins MA. Rigorous derivation of reaction path degeneracy in transition state theory. Chem Phys Lett 1992;193:181–4.
[168] Truhlar DG, Isaacson AD. Simple perturbation theory estimates of equilibrium constants from force fields. J Chem Phys 1991;94:357–9.
[169] Hernandez R, Miller WH. Semiclassical transition state theory: a new perspective. Chem Phys Lett 1993;214:129–36.
[170] Miller WH, Hernandez R, Handy NC, Jayatilaka D, Willets A. Ab initio calculation of anharmonic constants for a transition state, with application to semiclassical transition state tunneling probabilities. Chem Phys Lett 1990;172:62–8.
[171] Nguyen TL, Stanton JF, Barker JR. A practical implementation of semi-classical transition state theory for polyatomics. Chem Phys Lett 2010;499:9–15.
[172] Aieta C, Gabas F, Ceotto M. Parallel implementation of semiclassical transition state theory. J Chem Theory Comput 2019;15:2142–53.
[173] Goel P, Stanton JF. Semiclassical transition state theory based on fourth order vibrational perturbation theory: model system studies beyond symmetric Eckart barrier. J Chem Phys 2018;149:134109.
[174] Wagner AF. Improved multidimensional semiclassical tunneling theory. J Phys Chem A 2013;117:13089–100.
[175] Shan X, Burd TAH, Clary DC. New developments in semiclassical transition-state theory. J Phys Chem A 2019;123:4639–57.
[176] Bao JL, Truhlar DG. Variational transition state theory: theoretical framework and recent developments. Chem Soc Rev 2017;46:7548–96.
[177] Weston HE, Schwartz RE, editors. Chemical kinetics. Englewood Cliffs, NJ: Prentice-Hall Inc.; 1972.
[178] Greene SM, Shan X, Clary DC. Rate constants of chemical reactions from semiclassical transition state theory in full and one dimension. J Chem Phys 2016;144:244116.
[179] von Horsten HF, Banks ST, Clary DC. An efficient route to thermal rate constants in reduced dimensional quantum scattering simulations: applications to the abstraction of hydrogen from alkanes. J Chem Phys 2011;135:094311.
[180] Bryukov MG, Slagle IR, Knyazev VD. Kinetics of reactions of H atoms with ethane and chlorinated ethanes. J Phys Chem A 2001;105:6900–9.
[181] Miller JA, Klippenstein SJ. Master equation methods in gas phase chemical kinetics. J Phys Chem A 2006;110:10528–44.
[182] Robertson SH, Pilling MJ, Jitariu LC, Hillier IH. Master equation methods for multiple well systems: application to the 1-,2-pentyl system. Phys Chem Chem Phys 2007;9:4085–97.
[183] Barker JR, Weston RE. Collisional energy transfer probability densities P(E, J; E′, J′) for monatomics colliding with large molecules. J Phys Chem A 2010;114:10619–33.
[184] Glowacki DR, Liang C-H, Morley C, Pilling MJ, Robertson SH. MESMER: an open-source master equation solver for multi-energy well reactions. J Phys Chem A 2012;116:9545–60.
[185] Gillespie DT. Stochastic simulation of chemical kinetics. Annu Rev Phys Chem 2007;58:35–55.
[186] Barker JR. Multiple-well, multiple-path unimolecular reaction systems, I: MultiWell computer program suite. Int J Chem Kinet 2001;33:232–45.
[187] Robertson SH, Pilling MJ, Gates KE, Smith SC. Application of inverse iteration to 2-dimensional master equations. J Comput Chem 1997;18:1004–10.
[188] Nguyen TL, Stanton JF. A steady-state approximation to the two-dimensional master equation for chemical kinetics calculations. J Phys Chem A 2015;119:7627–36.
[189] Garrod RT, Herbst E. Formation of methyl formate and other organic species in the warm-up phase of hot molecular cores. Astron Astrophys 2006;457:927–36.
[190] Garrod RT, Weaver SLW, Herbst E. Complex chemistry in star-forming regions: an expanded gas-grain warm-up chemical model. Astrophys J 2008;682:283–302.
[191] Bacmann A, Taquet V, Faure A, Kahane C, Ceccarelli C. Detection of complex organic molecules in a prestellar core: a new challenge for astrochemical models. Astron Astrophys 2012;541:L12.
[192] Cernicharo J, Marcelino N, Roueff E, Gerin M, Jiménez-Escobar A, Caro GMM. Discovery of the methoxy radical, CH_3O, toward B1: dust grain and gas-phase chemistry in cold dark clouds. Astrophys J 2012;759:L43.
[193] Jaber AA, Ceccarelli C, Kahane C, Caux E. The census of complex organic molecules in the solar-type protostar IRAS16293-2422. Astrophys J 2014;791:29.
[194] Vastel C, Ceccarelli C, Lefloch B, Bachiller R. The origin of complex organic molecules in prestellar cores. Astrophys J 2014;795:L2.
[195] Jiménez-Serra I, Vasyunin AI, Caselli P, Marcelino N, Billot N, Viti S, et al. The spatial distribution of complex organic molecules in the L1544 pre-stellar core. Astrophys J 2016;830:L6.

[196] Vasyunin AI, Herbst E. Reactive desorption and radiative association as possible drivers of complex molecule formation in the cold interstellar medium. Astrophys J 2013;769:34.

[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 2015;449:L16–20.

[198] Vasyunin AI, Caselli P, Dulieu F, Jiménez-Serra I. Formation of complex molecules in prestellar cores: a multilayer approach. Astrophys J 2017;842:33.

[199] 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.

[200] Barone V, Latouche C, Skouteris D, Vazart F, Balucani N, Ceccarelli C, et al. Gas-phase formation of the prebiotic molecule formamide: insights from new quantum computations. Mon Not R Astron Soc Lett 2015;453:L31.

[201] Skouteris D, Vazart F, Ceccarelli C, Balucani N, Puzzarini C, Barone V. New quantum chemical computations of formamide deuteration support gas-phase formation of this prebiotic molecule. Mon Not R Astron Soc Lett 2017;468:L1.

[202] Codella C, Ceccarelli C, Caselli P, Balucani N, Barone V, Fontani F, et al. Seeds of Life in Space (SOLIS), II: formamide in protostellar shocks: evidence for gas-phase formation. Astron Astrophys 2017;605:L3.

[203] Coutens A, Jørgensen JK, van der Wiel MHD, Müller HSP, Lykke JM, Bjerkeli P, et al. The ALMA-PILS survey: first detections of deuterated formamide and deuterated isocyanic acid in the interstellar medium. Astron Astrophys 2016;590:L6.

[204] Montgomery Jr JA, Frisch MJ, Ochterski JW, Petersson GA. A complete basis set model chemistry, VI: use of density functional geometries and frequencies. J Chem Phys 1999;110:2822–7.

[205] Montgomery Jr JA, Frisch MJ, Ochterski JW, Petersson GA. A complete basis set model chemistry, VII: use of the minimum population localization method. J Chem Phys 2000;112:6532–42.

[206] Rivilla VM, Beltrán MT, Cesaroni R, Fontani F, Codella C, Zhang Q. Formation of ethylene glycol and other complex organic molecules in star-forming regions. Astron Astrophys 2017;598:A59.

[207] Arumainayagam CR, Garrod RT, Boyer MC, Hay AK, Bao ST, Campbell JS, et al. Extraterrestrial prebiotic molecules: photochemistry vs. radiation chemistry of interstellar ices. Chem Soc Rev 2019;48:2293–314.

[208] Garrod RT, Widicus Weaver SL. Simulations of hot-core chemistry. Chem Rev 2013;113:8939–60.

[209] Atkins P, De Paula J. Physical chemistry. Oxford University Press; 2014.

[210] Chung LW, Sameera WMC, Ramozzi R, Page AJ, Hatanaka M, Petrova GP, et al. The ONIOM method and its applications. Chem Rev 2015;115:5678–796.

[211] Re S, Morokuma K. ONIOM study of chemical reactions in microsolvation clusters: $(H_2O)_n CH_3Cl + OH^-(H_2O)_m$ $(n + m = 1$ and 2). J Phys Chem A 2001;105:7185–97.

[212] Tomasi J, Mennucci B, Cammi R. Quantum mechanical continuum solvation models. Chem Rev 2005;105:2999–3094.

[213] Rimola A, Skouteris D, Balucani N, Ceccarelli C, Enrique-Romero J, Taquet V, et al. Can formamide be formed on interstellar ice? An atomistic perspective. ACS Earth Space Chem 2018;2:720–34.

[214] Clarke D, Ferris JP. Chemical evolution on titan: comparisons to the prebiotic Earth. Orig Life Evol Biosph 1997;27:225–48.

[215] Raulin F, McKay C, Lunine J, Owen T, editors. Titan from Cassini-Huygens. Dordrecht: Springer; 2009.

[216] Balucani N, Leonori F, Petrucci R, Stazi M, Skouteris D, Rosi M, et al. Formation of nitriles and imines in the atmosphere of Titan: combined crossed-beam and theoretical studies on the reaction dynamics of excited nitrogen atoms $N(^2D)$ with ethane. Faraday Discuss 2010;147:189–216.

[217] Parker DSN, Kaiser RI. On the formation of nitrogen-substituted polycyclic aromatic hydrocarbons (NPAHs) in circumstellar and interstellar environments. Chem Soc Rev 2017;46:452–63.

[218] Palmer MY, Cordiner MA, Nixon CA, Charnley SB, Teanby NA, Kisiel Z, et al. ALMA detection and astrobiological potential of vinyl cyanide on Titan. Sci Adv 2017;3:e1700022.

[219] Vuitton V, Yelle RV, Lavvas P. Composition and chemistry of Titan's thermosphere and ionosphere. Philos Trans R Soc A 2009;367:729–41.

[220] Lavvas P, Yelle RV, Koskinen T, Bazin A, Vuitton V, Vigren E, et al. Aerosol growth in titan's ionosphere. Proc Natl Acad Sci 2013;110:2729–34.

[221] Hartle RE, Sittler Jr EC, Neubauer FM, Johnson RE, Smith HT, Crary F, et al. Preliminary interpretation of Titan plasma interaction as observed by the Cassini Plasma Spectrometer: comparisons with Voyager 1. Geophys Res Lett 2006;33:L08201.

[222] Hörst S, Vuitton V, Yelle R. Origin of oxygen species in Titan's atmosphere. J Geophys Res 2008;113:E10006.

[223] Hörst S, Yelle R, Buch A, Carrasco N, Cernogora G, Dutuit O, et al. Formation of amino acids and nucleotide bases in a Titan atmosphere simulation experiment. Astrobiology 2012;12:809–17.

[224] Licari D, Mancini G, Brogni A, Salvadori A, Barone V. The SMART cyberinfrastructure: space-time multiscale approaches for research and technology. CRC Press; 2017.

[225] Salvadori A, Del Frate G, Pagliai M, Mancini G, Barone V. Immersive virtual reality in computational chemistry: applications to the analysis of QM and MM data. Int J Quant Chem 2016;116:1731–46.

[226] Licari D, Tasinato N, Spada L, Puzzarini C, Barone V. VMS-ROT: a new module of the Virtual Multifrequency Spectrometer for simulation, interpretation, and fitting of rotational spectra. J Chem Theory Comput 2017;13:4382–96.

[227] Licari D, Fusè M, Salvadori A, Tasinato N, Mendolicchio M, Mancini G, et al. Towards the SMART workflow system for computational spectroscopy. Phys Chem Chem Phys 2018;20:26034–52.

[228] Puzzarini C, Bloino J, Tasinato N, Barone V. Accuracy and interpretability: the devil and the holy grail. New routes across old boundaries in computational spectroscopy. Chem Rev 2019;119:8131–91.