

Quantum Chemistry Meets Spectroscopy for Astrochemistry: Increasing Complexity toward Prebiotic Molecules

Vincenzo Barone,^{*,†} Malgorzata Biczysko,^{*,‡} and Cristina Puzzarini^{*,¶}

[†]Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy

[‡]Consiglio Nazionale delle Ricerche, Istituto di Chimica dei Composti OrganoMetallici (ICCOM-CNR), UOS di Pisa, Area della Ricerca CNR, Via G. Moruzzi 1, I-56124 Pisa, Italy

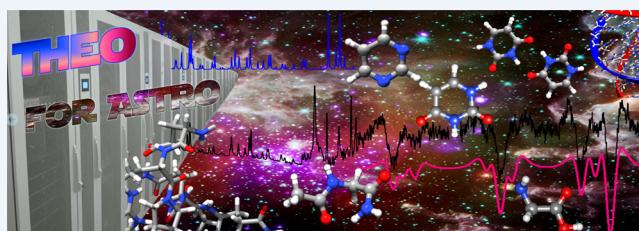
[¶]Dipartimento di Chimica "Giacomo Ciamician", Università di Bologna, Via F. Selmi 2, 40126 Bologna, Italy

CONSPECTUS: For many years, scientists suspected that the interstellar medium was too hostile for organic species and that only a few simple molecules could be formed under such extreme conditions. However, the detection of approximately 180 molecules in interstellar or circumstellar environments in recent decades has changed this view dramatically. A rich chemistry has emerged, and relatively complex molecules such as C_{60} and C_{70} are formed. Recently, researchers have also detected complex organic and potentially prebiotic molecules,

such as amino acids, in meteorites and in other space environments. Those discoveries have further stimulated the debate on the origin of the building blocks of life in the universe. Many efforts continue to focus on the physical, chemical, and astrophysical processes by which prebiotic molecules can be formed in the interstellar dust and dispersed to Earth or to other planets.

Spectroscopic techniques, which are widely used to infer information about molecular structure and dynamics, play a crucial role in the investigation of planetary atmosphere and the interstellar medium. Increasingly these astrochemical investigations are assisted by quantum-mechanical calculations of structures as well as spectroscopic and thermodynamic properties, such as transition frequencies and reaction enthalpies, to guide and support observations, line assignments, and data analysis in these new and chemically complicated situations. However, it has proved challenging to extend accurate quantum-chemical computational approaches to larger systems because of the unfavorable scaling with the number of degrees of freedom (both electronic and nuclear).

In this Account, we show that it is now possible to compute physicochemical properties of building blocks of biomolecules with an accuracy rivaling that of the most sophisticated experimental techniques, and we summarize specific contributions from our groups. As a test case, we present the underlying computational machinery through the investigation of oxirane. We describe how we determine the molecular structure and then how we characterize the rotational and IR spectra, the most important issues for a correct theoretical description and a proper comparison with experiment. Next, we analyze the spectroscopic properties of representative building blocks of DNA bases (uracil and pyrimidine) and of proteins (glycine and glycine dipeptide analogue). Solvation, surface chemistry (dust fraction, adsorption, desorption), and inter- and intramolecular interactions, such as self-organization and self-interaction, are important molecular processes for understanding astrochemistry. Using the specific cases of uracil dimers and glycine adsorbed on silicon grains, we also illustrate approaches in which we treat different regions, interactions, or effects at different levels of sophistication.



■ INTRODUCTION

Astrochemistry is the study of the chemistry that occurs throughout the universe with the aim of understanding its evolution toward molecular complexity. Its birth can be dated back to 1968, when ammonia was first detected by Townes and co-workers toward the center of our galaxy via its inversion transitions.¹ Since then, about 180 molecules have been detected in space, with at least 175 molecules observed unambiguously in the interstellar medium (ISM) or circumstellar shells and more than 50 molecules detected in extragalactic sources (see, for example, the Cologne database²). Our knowledge of the universe chemical inventory has been obtained and continuously updated by means of astronomical observations, which have revealed a

plethora of molecular species.³ Gas-phase species have been mostly discovered via ground-based observations⁴ of their rotational signatures (frequencies going from the millimeter-wave region to far-infrared), with nowadays the unprecedented resolution and sensitivity of the Atacama Large Millimeter/submillimeter Array (ALMA) offering unique opportunities. However, new possibilities are also provided by spectrometers on board satellites (e.g., Herschel) or airborne (SOFIA), thus improving the role played by infrared spectroscopies in retrieving the chemical composition of either planetary atmospheres or the ISM. The dimensions of the detected

Received: September 4, 2014

Published: April 20, 2015

molecules range from diatomics to species containing more than 10 atoms, up to polycyclic aromatic hydrocarbons (PAHs) that were identified to be abundant, for instance, in the atmosphere of Titan, Saturn's largest moon.

One of the major aims of Astrochemistry is to explain the origin and abundance of the molecules detected. Therefore, astrochemistry spans astronomical observations, modeling, and theoretical or experimental laboratory based investigations. This is outlined in Figure 1, which tries to give a feeling of the

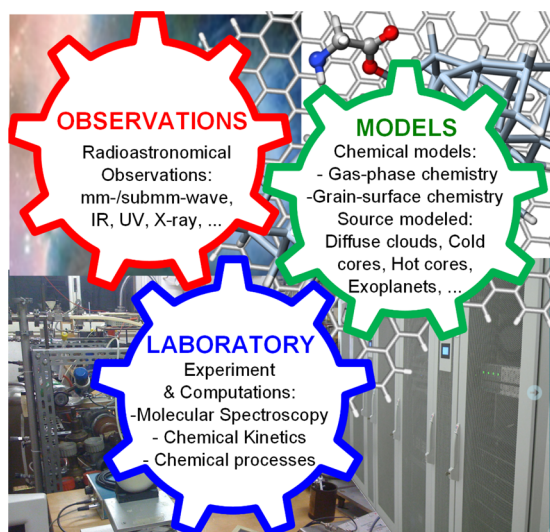


Figure 1. Interplay of observations, laboratory studies, and models in astrochemistry.

interplay of astronomical observations, laboratory studies, and modeling. Astrochemistry is therefore a wide research field that cannot be exhaustively addressed in only one contribution. This is demonstrated by recent dedicated issues in *Chemical Reviews*⁵ and *Physical Chemistry Chemical Physics*.⁶ Inspection of these issues points out an increasing interest for this field but also that important aspects require deeper consideration. In this contribution, we focus on laboratory studies and, in particular, on the role played by quantum-mechanical (QM) calculations in “laboratory” spectroscopic characterizations that, in turn, support astronomical observations. Several quantum chemists have provided remarkable contributions in this connection (among others, we can mention L. Allamandola, C. Bauschlicher, P. Botschwina, R.C. Fortenberry, J. Gauss, T. J. Lee, K. Peterson, H. F. Schaefer, J. Stanton, D. Woon), but this Account is devoted to the specific contributions from our groups to the development of a comprehensive strategy for treating larger molecules with the required accuracy.

Spurred on by the recent detection of complex organic molecules and by the search for prebiotic species in space, we present the computational tools that can support experimental studies using oxirane as a test case. Starting from molecular structure, we show how to compute with the required accuracy the spectroscopic properties at the basis of astronomical observations. Next, the corresponding analysis is presented for representative biomolecule building blocks.

■ TOWARD ACCURATE COMPUTATIONS FOR ASTROCHEMICAL STUDIES: OXIRANE AS A TEST CASE

Spectroscopic signatures provide the unequivocal proof of the presence of chemical species, which is the starting point for the development of astrochemical and astrophysical models and for laying the foundation of astrobiology. Although a great deal of information can be retrieved from laboratory experiments, even for small molecules uncertainties in the assignment of some spectral features may occur. On the other hand, quantum chemistry is nowadays able either to provide accurate spectroscopic information that is not hampered by interaction with environment or to explicitly account for interactions with other molecules (microsolvation) or with solid organic or inorganic supports (the interstellar dust). Our strategy is based on the computation at the highest possible level of equilibrium properties (molecular structure, harmonic frequencies, relative energies, etc.), with anharmonicity and other effects (complexation, molecule–surface interactions) accounted for using sufficiently accurate yet cost-effective models rooted in density functional theory (DFT).

Molecular Structure

The best choice to report structural information is the equilibrium structure since it is well-defined (minimum of the Born–Oppenheimer (BO) potential energy surface (PES)), it excludes vibrational effects in a rigorous manner, and within the BO approximation, it is independent of the considered isotopic species. In this context, an important outcome has been provided by the introduction of the so-called semiexperimental (SE) equilibrium geometry (r_e^{SE}),^{7–9} which is obtained by a least-squares fit of experimental rotational constants of different isotopologues computationally corrected for vibrational contributions. The corresponding corrections, evaluated by means of second-order vibrational perturbation theory (VPT2),¹⁰ are the computational bottleneck of the SE protocol, since they require the calculation of anharmonic force fields at a sufficiently accurate level of theory. It has been shown that anharmonic force fields evaluated by hybrid density functionals and medium-sized basis sets are sufficiently accurate.^{8,9,11,12} They can be routinely obtained for molecules containing about 50 atoms thanks to the favorable scaling of DFT with the number of degrees of freedom (better than N^3n^3 , where N and n are the numbers of atoms and electronic basis functions, respectively). As a matter of fact, this approach permits the determination of structural parameters accurate within 0.001 Å and 0.1° for bond lengths and angles, respectively,^{8,9,13} thus representing a suitable reference for QM calculations.

Concerning, instead, theoretical structures, for systems not showing excessive multireference character, coupled-cluster (CC) techniques represent the methods of choice, and full inclusion of single and double excitations augmented by a perturbative treatment of triple excitations (the so-called CCSD(T) model)¹⁴ has become the gold standard for energy evaluations and geometry optimizations. Despite the quite unfavorable scaling of this method with the number of electronic basis functions (n^6 iterative and n^7 for the perturbative treatment), the ongoing developments of software and hardware have already made computations for some hundreds of basis functions feasible. Unfortunately, to determine accurate equilibrium structures, it is necessary to account simultaneously for the principal error sources, that is, basis-set and wave function truncation effects, and for core-

correlation contributions. The most successful route toward this fulfilment is provided by composite methods, where the various contributions are evaluated separately at the highest possible level and then combined together relying on the theoretically derived or empirically validated additive approximations.^{11,15–19} In particular, the model based on additivity applied at an energy gradient level and involving CCSD(T) computations with different basis sets (denoted as best-CC in the literature^{11,20}) is both theoretically justified,^{15,16} and remarkably accurate.^{13,15,16} The size of molecular systems amenable to accurate analysis can be significantly extended by estimating the complete basis-set limit and core-correlation effects by second-order Møller–Plesset perturbation theory (MP2),²¹ on top of CCSD(T) calculations with a medium-sized basis set. Furthermore, this approach (denoted in literature as best-cheap²⁰) relies on the empirical approximation that additive schemes can be applied to geometry parameters rather than to energy gradients. In general terms, the best-CC method can be used for semirigid molecules containing at most six first-row atoms (in addition to hydrogens and no more than one second-row nucleus) and provides an accuracy of 0.001 Å and 0.1° for bond lengths and angles, respectively. On the other hand, molecules containing up to 15 non-hydrogen atoms can be treated by means of the best-cheap approach with an estimated accuracy of 0.0015 Å and 0.15° for bond lengths and angles, respectively.^{11,19,22} Figure 2 reports these error estimates for the specific case of oxirane, whereas Figure 3 confirms the accuracy of the best-cheap model for larger systems (glycine and uracil).^{11,19}

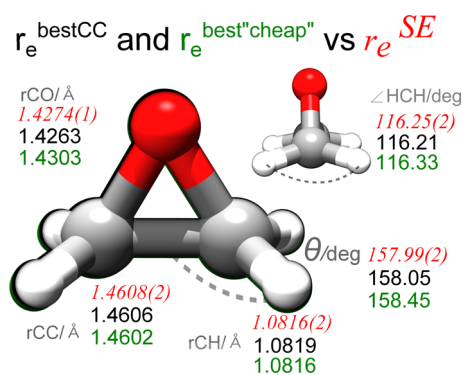


Figure 2. Molecular structure of oxirane: equilibrium (r_e^{theo}) geometries obtained by means of the best-CC (black) and best-cheap (green) composite approaches compared with the semi-experimental (r_e^{SE}) equilibrium structure (red and italic, uncertainty on the last digit in parentheses).

Rotational Spectroscopy

The rotational constants are the leading parameters in the field of rotational spectroscopy and about 95–98% of their values is due to the equilibrium contribution, which only depends on the equilibrium structure. Therefore, accurate determinations of the molecular structure, as discussed in the previous section, play a fundamental role in predicting the rotational fingerprints. The prediction of rotational spectra also requires the determination of the vibrational corrections to rotational constants as well as the centrifugal-distortion constants; both of them imply the computation of cubic force fields (e.g., see refs 13 and 23 for further details) and the corresponding VPT2^{8,10,24,25} analysis. The effectiveness and accuracy of the VPT2 approach are well assessed when coupled with electronic computations at the

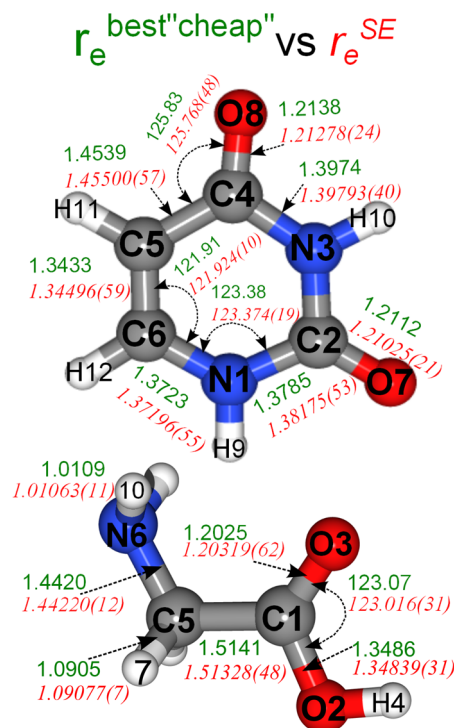


Figure 3. Molecular structure of uracil¹⁹ and of the most stable glycine conformer (Ip):¹¹ selected geometrical parameters obtained by means of the best-cheap composite approach (green) compared with the corresponding r_e^{SE} ones (red and italic, with uncertainty on the last digit in parentheses).

CCSD(T) level in conjunction with basis sets of at least triple- ζ quality.^{9,13} Nevertheless, it is well demonstrated that hybrid force fields where DFT anharmonic corrections are added to CCSD(T) harmonic frequencies provide remarkably accurate results at a much reduced computational cost.^{11,19,22,26}

Prediction of the transition intensity and the type of spectra observable requires also dipole moments, which like other first-order properties, can be computed with very high accuracy.¹³

Rotational Spectra of Oxirane Isotopologues. Focusing on astronomical observations, millimeter- and submillimeter-wave spectroscopy (corresponding to the frequency regions where rotational transitions lie) permits us to “watch” molecules in space by detecting the light reaching the Earth from the Universe’s earliest galaxies. In this view, the ALMA with its increased sensitivity and spatial resolution is opening up the possibility to detect organic and prebiotic species from bright, nearby sources with narrow emission lines or from the coldest and earliest stages in the formation of solar-type systems.²⁷ Furthermore, its high sensitivity will allow for the observation of isotopic species of well-assessed astrophysical species, which can provide useful information on nucleosynthesis.³

The accuracy of QM calculations in predicting rotational spectra of interest to astrophysical investigations is exemplified by Figure 4, which compares the predicted rotational spectra of oxirane and its monodeuterated isotopic species (oxirane- d_1) in band 3 of ALMA (84–116 GHz) to the available experimental data. It is noted that the calculated transitions for the main isotopic species are usually overestimated by less than 0.1%, and similar accuracy is expected for oxirane- d_1 . Therefore, theoretical predictions are sufficiently accurate for supporting either laboratory or astronomical assignments.

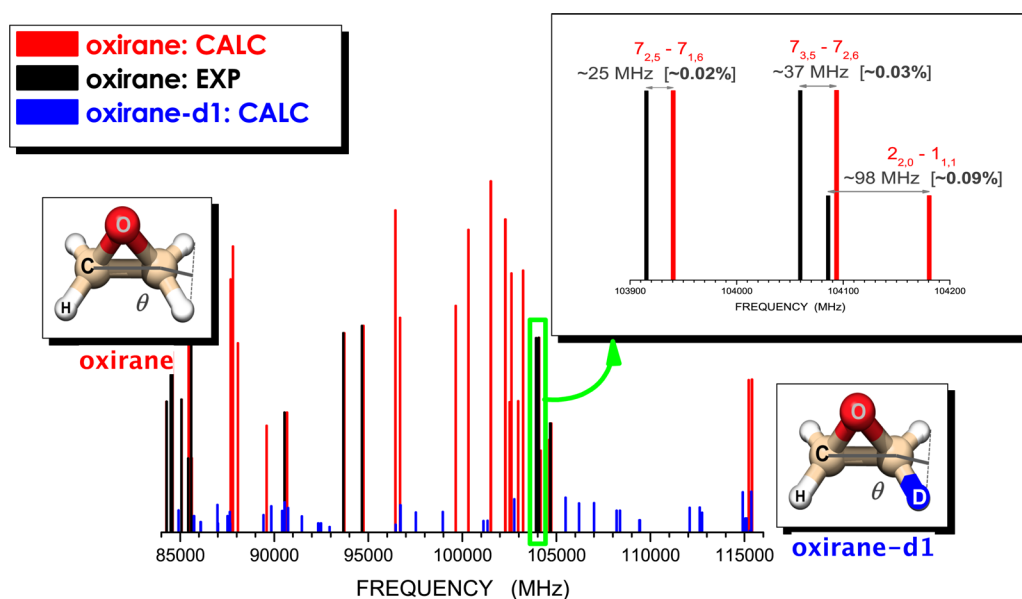


Figure 4. Computed rotational spectra (for $J \leq 10$) of oxirane (in red) and oxirane- d_1 (in blue); for the latter, intensities were reduced by 1 order of magnitude. The comparison with available experimental frequencies for oxirane (in black) is provided.

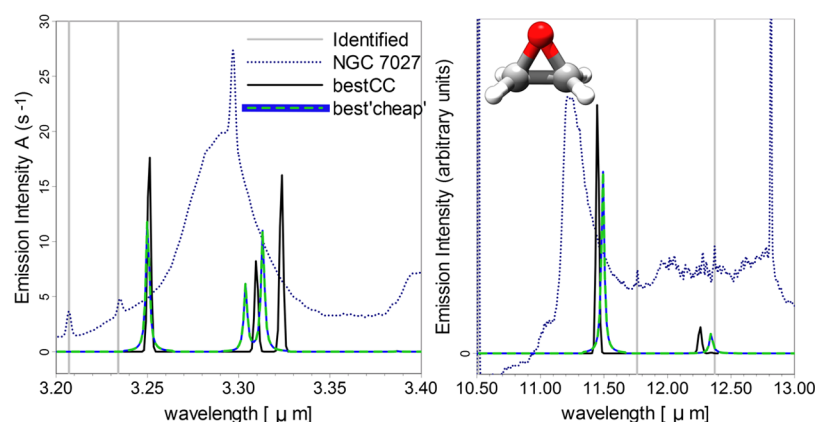


Figure 5. Comparison between simulated (black, best-CC/CCSD(T);⁴² blue/green, best-cheap/B3LYP) emission spectra of oxirane and observed spectra of the planetary nebula NGC 7027⁴³ (dark blue, dotted). Identified line fluxes⁴⁴ are marked by the gray lines. Theoretical stick spectra were broadened by Lorentzian functions with a half width at half maximum (HWHM) of 1 cm^{-1} .

Vibrational Spectroscopy

The information required for predicting or analyzing spectra in the field of vibrational spectroscopy are the *vibrational frequencies* and the corresponding *intensities*. To fulfill the accuracy (for frequencies) and interpretability (for intensities) requirements, it is mandatory to go beyond the so-called double-harmonic approximation, accounting for both mechanical and electrical anharmonic effects.^{28–30} In fact, not only the harmonic frequencies show too large deviations from the actual line positions, but also overtones and combination bands can be analyzed only once anharmonicity is accounted for, because their intensity vanishes at the harmonic level.³⁰

Among the approaches proposed for the computation of anharmonic frequencies,^{28–32} VPT2¹⁰ offers a remarkable compromise between reliability and effectiveness at least for semirigid molecules. VPT2 provides also intensities (infrared, vibrational circular dichroism, etc.),^{24,30,31,33–35} only requiring cubic and semidiagonal quartic force fields together with second and semidiagonal third derivatives of transition dipole moments in a normal mode representation.

As concerns the computational prerequisites, accurate harmonic frequencies require CCSD(T) computations possibly accounting for basis-set and core-correlation effects, whereas hybrid DFT functionals are suitable for deriving accurate anharmonic corrections.^{36,37} The resulting hybrid approaches are well tested for vibrational frequencies^{11,20,30,38,39} and show promising results for IR intensities.^{11,20,30} In the following, we denote these hybrid approaches as QM/QM', with QM and QM' referring to the harmonic and anharmonic parts of the force field, respectively.

Vibrational Spectrum of Oxirane: Assignment of Unidentified Infrared Bands. One of the most puzzling questions related to the identification of molecules and complex systems in space by means of IR spectroscopy concerns the so-called unidentified infrared bands (UIRs), for which several sources have been proposed, including small carbonaceous molecules (SCMs), like oxirane,⁴⁰ PAHs, or mixed aromatic and aliphatic molecules.⁴¹ However, considerable uncertainties in some assignments still persist because of the lack of accurate predictions of the spectroscopic parameters involved, also

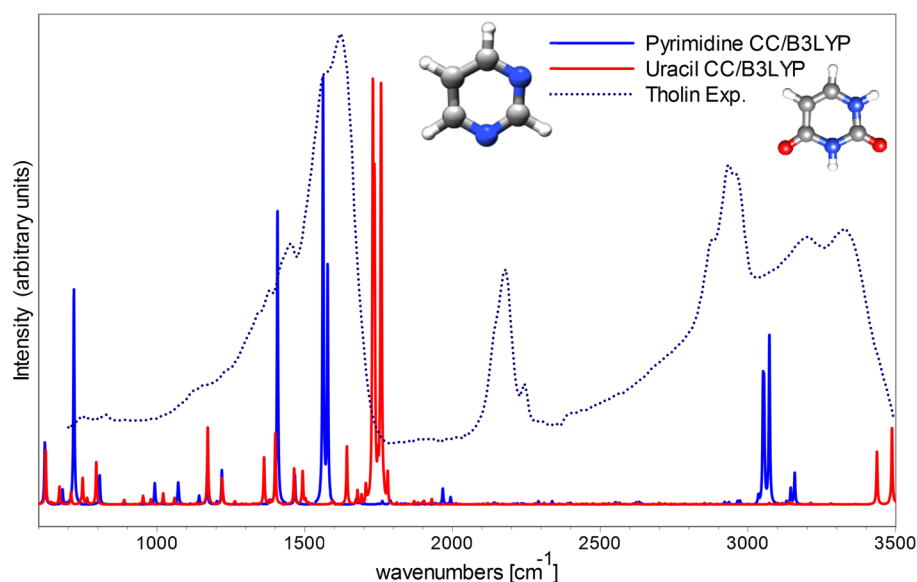


Figure 6. Comparison of the simulated IR spectrum of pyrimidine⁴⁵ and uracil³⁹ with the experimental spectrum of tholin (from Figure 1 of ref 51). Best-cheap/B3LYP stick spectra were broadened by Lorentzian functions with HWHM of 2 cm^{-1} .

including Einstein spontaneous emission coefficients, A . To fill this gap, quantitative simulations of UIR spectra for different molecular species can be performed. To give an example, in a recent study on oxirane,⁴² the entire IR spectrum ($800\text{--}6000\text{ cm}^{-1}$) was simulated and compared with the corresponding observations, taking as reference the ISO-SW spectrum of the planetary nebula NGC 7027.⁴³ The portion corresponding to the $3.2\text{--}3.4\text{ }\mu\text{m}$ and $10.5\text{--}13.0\text{ }\mu\text{m}$ wavelength ranges are depicted in Figure 5, along with all identified line fluxes reported by Salas et al.⁴⁴ The chosen computational models (best-CC/CCSD(T) and best-cheap/B3LYP) reproduce well the observed UIR feature at $3.3\text{ }\mu\text{m}$, for both band position and absolute intensity. Moreover, the simulated band at $11.45/11.49\text{ }\mu\text{m}$, which is related to the ring deformation and can be therefore considered characteristic of oxirane, matches well (within 10 cm^{-1}) the $11.2\text{--}11.3\text{ }\mu\text{m}$ UIR features. In ref 42, it was demonstrated that the best theoretical estimates have an accuracy of about $5\text{--}10\text{ cm}^{-1}$ and $10\text{--}20\text{ cm}^{-1}$ for fundamental and nonfundamental transitions, respectively, with theoretical intensities deviating by 10% at most, from the high-resolution results. This accuracy allows for reliable simulations of IR spectra, which in turn can support astronomical observations.^{11,19,20,45}

■ CASE STUDIES

Several questions are still open concerning the chemical evolution of the universe. Particularly timely are those related to organic and prebiotic chemistry, for example: How did the solar system evolve to the present stage? How did life originate on earth? The first step toward the answer to these questions is the identification in space of organic and prebiotic molecules.

Semirigid Life Bricks: Pyrimidine and Uracil

DNA components have been confirmed to exist in extra-terrestrial meteorites,⁴⁶ and various studies have shown that nucleobases and other prebiotic molecules can be formed in irradiated ices (see, for example, ref 47) or in the complex organic aerosols produced in Titan's atmosphere (see, for example, refs 48 and 49), but to date, all searches for these molecules in the ISM turned out to be unsuccessful.

Titan's atmosphere is characterized by a rich and complex organic chemistry (see Ali et al.⁵⁰ and references therein) and is considered to be a model of primitive earth due to significant amounts of carbon ($\sim 2\%$ CH_4) and nitrogen (98% N_2), and traces of oxygen ($\sim 50\text{ ppm}$ of CO). The investigation of Titan's atmosphere might thus provide a unique opportunity to explain terrestrial prebiotic chemistry.⁴⁹ For this reason, several studies involving the simulation of Titan's atmosphere have been carried out.^{48,49} The complex organic mixtures that are formed in these simulation experiments (usually by irradiation of N_2 , CH_4 , H_2 , and CO) are known as "tholins".⁴⁹ Prebiotic molecules, in particular nitrogen-containing cyclic compounds such as purines and pyrimidines, are produced from non-biological matter in thermal heating and electric discharge experiments.⁴⁹ Detection of prebiotic species would thus confirm that Titan is a model of the prebiotic atmosphere of early earth. Indeed, Cassini ion neutral mass spectrometer measurements pointed out the presence of large molecules ($>100\text{ }\mu\text{m}$) in Titan's atmosphere: the measurements by instruments onboard the Cassini spacecraft clearly demonstrated that simple molecules in the gas phase do evolve into organic macromolecules characterized by the structural complexity typical of prebiological molecules.⁵⁰ The next step to be taken is thus the astronomical observation of prebiotic molecules and their identification in either the infrared or millimeter-wave region.

Concerning rotational fingerprints, the high spectral resolution and sensitivity of ALMA open up the possibility to detect prebiotic species. In this respect, the high accuracy of theoretical predictions reached for uracil¹⁹ and thioracil²⁶ confirms that we have at our disposal a computational strategy able to provide rotational parameters with the required accuracy for guiding laboratory or astronomical investigations.

Moving to vibrational spectroscopy, mid- and far-IR (FIR) spectroscopy has recently been applied to investigate the tholin spectra,⁵¹ which in turn have been compared with the spectra of Titan's aerosols derived from recent Cassini composite infrared spectrometer (CIRS) observations. While many similarities have been pointed out, several spectral features could not be assigned unambiguously.⁵¹ Here, we consider the case of uracil

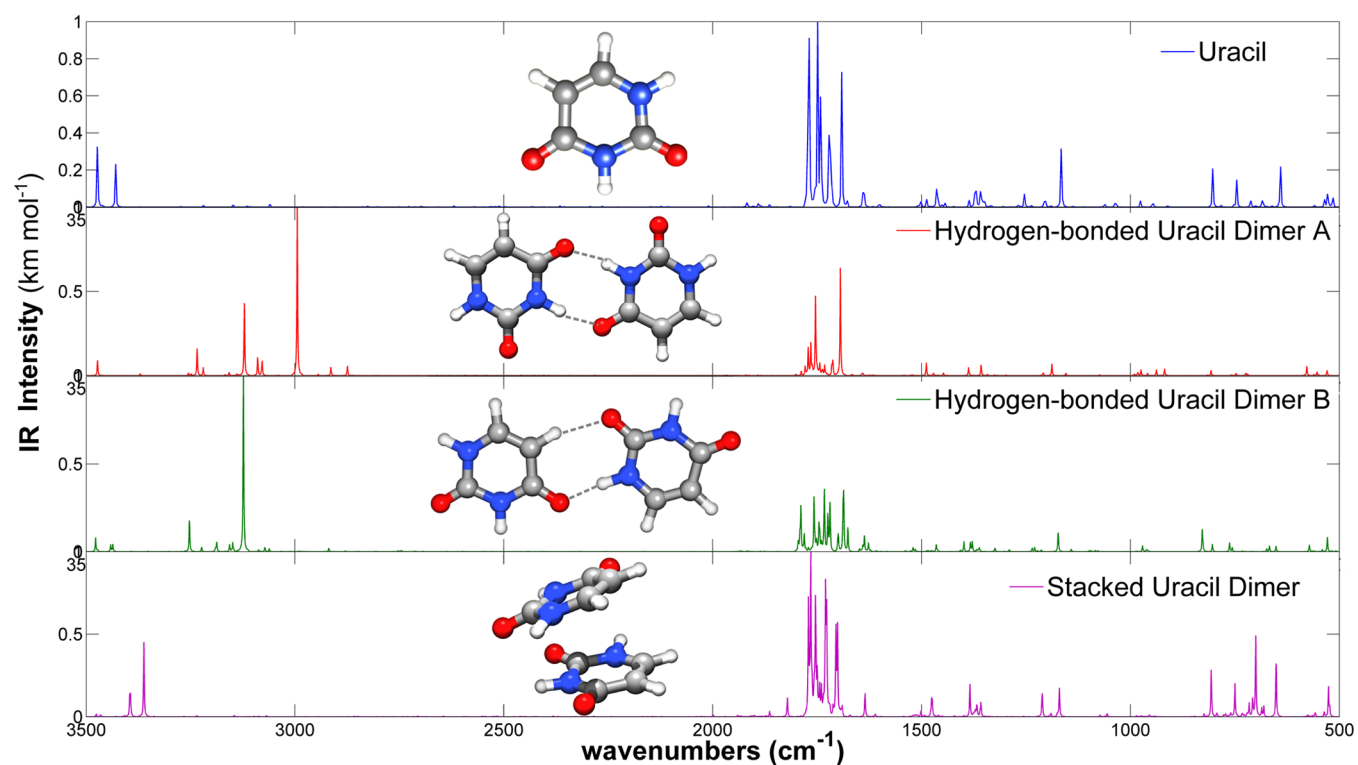


Figure 7. Comparison of the simulated IR spectra of uracil and its dimers. Best-cheap/B3LYP stick spectra were broadened by Lorentzian functions with HWHM of 2 cm^{-1} .

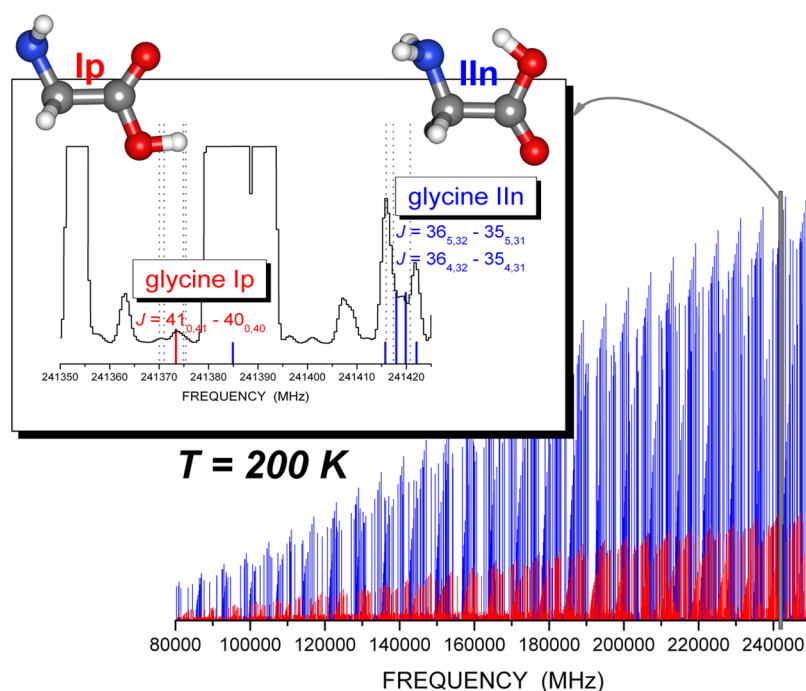


Figure 8. Simulation of the rotational spectra of glycine Ip and IIn at $T = 200\text{ K}$. Spectroscopic parameters are taken from CDMS;² population ratio is based on the thermodynamic properties computed in ref 11. In the inset, the 241.35–241.425 GHz portion is highlighted together with the simulated emission spectrum for the hot-core source NGC 6334 IRS1 (from Figure 5 of ref 27).

and pyrimidine as prebiotic systems possibly present in Titan's tholins. Figure 6 shows how well the computed IR spectrum of pyrimidine reproduces the intense spectral features of tholins, that is, the bands at 1562 and 1408 cm^{-1} , both due to the ring deformations involving C–N stretching vibrations, and the bands at 3000 – 3200 cm^{-1} , due to C–H stretchings. Moreover,

some less intense bands (CH bending at 1467 and 1366 cm^{-1} ; ring deformations, C–C and C–N, at 1143 and 1074 cm^{-1}) match well the unidentified features at 1450 , 1375 , 1140 , and 1070 cm^{-1} that have indeed been assigned to C–N aromatics and rings.⁵¹ On the other hand, the tholin spectrum does not show any intense transitions in the 1730 – 1760 cm^{-1} range,

which is characteristic for uracil due to the intense C=O stretchings and close-lying overtones involved in Fermi-resonance interactions.

Once prebiotic molecules are formed, more complex structures can be created, starting from molecular dimers toward larger complexes with higher degrees of self-organization. Detection of supramolecular structures and analysis of the different intermolecular interactions can be based on spectroscopic studies. Concerning uracil dimers, accurate simulations of IR spectra can be obtained by hybrid models (based on additivity assumptions) that combine accurate results for the monomer with intermolecular couplings computed by DFT models including dispersion corrections. Figure 7 clearly shows how different the spectral patterns of the isolated uracil molecule and its dimers either hydrogen bonded or stacked are.⁵²

Flexible Life Bricks: Glycine and Its Peptide Analogue

Nucleobases represent semirigid molecules, and even their oligomers show well-defined structures. The situation is different for the protein building blocks, which are characterized by high flexibility. Thus, the PES of amino acids and peptides are usually characterized by several low-energy minima, and accurate predictions of their relative stabilities are needed for the proper averaging of spectroscopic parameters. Among the molecules identified in space,⁴ glycine (NH₂CH₂COOH), the simplest α -amino acid, has been found in several carbonaceous meteorites and is one of the two most abundant amino acids in Stardust-returned foil samples.⁵³ While its presence in the ISM is yet to be confirmed, it has been recently postulated that glycine could reach detectable levels in cold solar-type system precursors⁵⁴ or hot-core sources.²⁷ Local thermodynamical equilibrium (LTE) calculations pointed out for the L1544 prestellar core (taken as an example of solar-type system precursor) a rather strong rotational spectrum of the most stable conformer of glycine in ALMA band 2 (68–90 GHz), with several lines between 67 and 80 GHz expected to have peak intensities greater than 10 mK.⁵⁴ In ref 27, the emission spectra for the hot-core source NGC 6334 IRS1 were simulated, with the portion of spectrum in the 241.35–241.425 GHz range showing a clear emission at 241.373 GHz due to glycine.²⁷ In both works,^{27,54} only the I_p conformer (see Figure 8) was considered. While the typical temperatures of prestellar core justify this choice (i.e., at temperature ≤ 10 K only the most stable conformer is populated), this is not the case for the characteristic temperature of the hot-core source NGC 6334 IRS1, that is, 200 K. As shown in Figure 8, at that temperature the rotational spectrum of glycine I_n is predicted to be the most intense. This is due to its very large dipole moment; in fact, the rotational spectrum of the less stable conformer I_n was the first observed by microwave spectroscopy.⁵⁵ As evident in the inset of Figure 8, in the portion of spectrum considered by Garrod, two transitions of I_n, more intense than that of I_p, are present, and actually, they seem to match well unassigned features (the reader is referred to Figure 5 of ref 27). In ref 27, the I_n conformer of glycine was not considered based on the assumption that it lies ~ 700 cm⁻¹ above I_p, thus leading to a population lower than 1%. On the contrary, the accurate QM calculations of ref 11 pointed out that the energy difference is only 191 cm⁻¹ (312 cm⁻¹ once zero-point vibrational correction is included), thus leading to a population of $\sim 10\%$ for I_n at $T = 200$ K.

Moving to the issue of detecting glycine by means of its IR signatures, our most accurate computational results, showing a mean absolute error (MAE) with respect to available experimental data smaller than 8 cm⁻¹, were obtained at the VPT2 level employing a hybrid best-cheap/B3LYP anharmonic force field for the six low-energy glycine conformers.¹¹ In particular, the most intense C=O stretching vibration, computed with an accuracy of ± 5 cm⁻¹, allows one to distinguish the different conformers in the fingerprint spectral range: a finding of clear relevance to the detection of glycine and other prebiotic molecules in complex astrophysical environments.

Adsorption of molecules on different substrates present in the ISM, such as the dust particles composed of mixtures of amorphous and crystalline organic and inorganic solid supports, is important because of self-organization and self-interaction of (bio)molecules at interfaces.⁵⁶ From a spectroscopic point of view, the first consequence is the appearance of new characteristic features. In this context, computational studies permit us to simulate accurate spectra even for complex systems, where differently oriented and adsorbed molecules show specific spectral patterns. To give an example, computational spectroscopy studies of a hybrid organic–inorganic system, namely, glycine adsorbed on silicon, led to the correct interpretation of the experimental results.⁵⁷ For the purpose of astrochemical identification, this type of investigation can be further improved by correcting the harmonic wavenumbers of the isolated molecule based on an accurate composite approach,¹¹ while considering the absorption and anharmonic effects at the DFT level. The computational burden of anharmonic calculations, which might be prohibitively expensive for sizable systems, can be significantly scaled down by reduced-dimensionality computations (see ref 30 for a detailed account), thus combining the accuracy of the anharmonic treatment with feasibility. In the case of glycine adsorbed on a silicon cluster,⁵⁷ the overall conclusion is that we have at our disposal accurate predictions of the IR spectra of the isolated and adsorbed molecule that can be directly compared. The fingerprint region of the most intense transitions due to the C=O stretching vibrations shown in Figure 9 points out the typical red and blue shifts of the $\nu(\text{C}=\text{O})_{\text{free}}$ and $\nu(\text{C}=\text{O})$ bands.

It is well recognized^{58–60} that not only hot cores but also dark cloud and prestellar cores reveal a high chemical complexity. An important step toward an answer concerning how life originated would be provided by the identification of molecules containing the peptide bond. Also in this context, computational studies can play a key role. As shown by ref 22, QM composite schemes were able to accurately predict the rotational parameters required to guide the assignment of the rotational spectra of the C₃ and C₇ conformers of *N*-acetyl-glycinamide, the simplest peptide analogue. The accuracy was such that for the rotational line positions an overall relative discrepancy (with respect to experiment) of 0.03% was observed. This is demonstrated by Figure 10, which reports a small portion of the rotational spectrum recorded. While the experimental investigation allowed for the accurate determination of the rotational constants, the limited region considered (6–18 GHz) did not permit the evaluation of the centrifugal-distortion constants required to extend the prediction of rotational transitions up to the frequency ranges covered by the ALMA bands. Once again, computational spectroscopy is able to provide the missing information, also concerning the relative

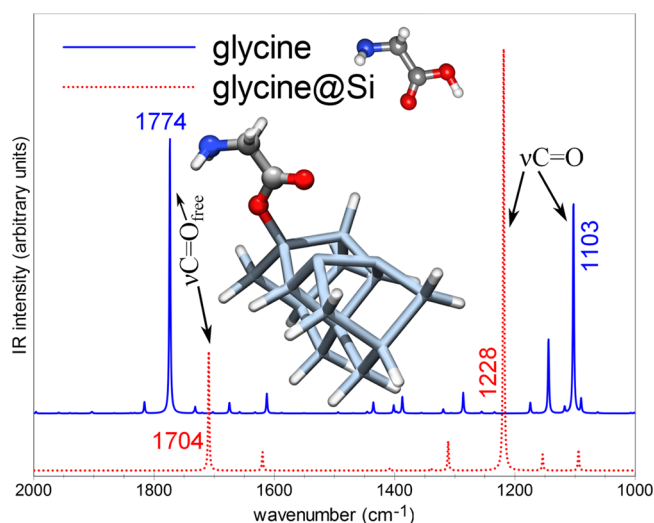


Figure 9. Infrared spectra of the glycine Ip conformer and glycine adsorbed on silicon cluster in the 2000–1000 cm^{-1} wavenumber range. Theoretical estimates are from hybrid best-cheap/B3LYP computations, with anharmonic and molecule–silicon interaction effects considered at the DFT level. The computed stick spectra were broadened by Lorentzian functions with HWHM of 1 cm^{-1} .

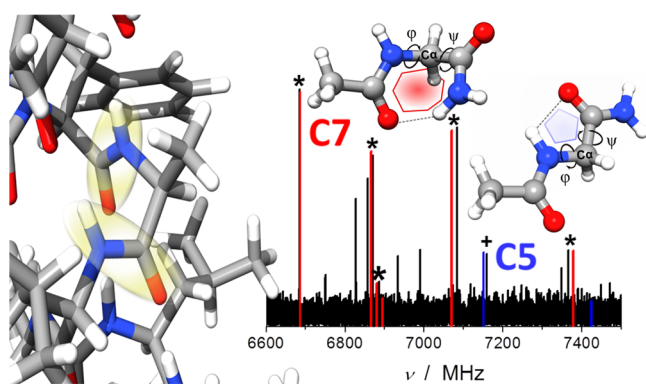


Figure 10. Schematic representation of the peptide bonding scheme along with the experimental (black) and computed rotational spectra of the C_7 (red) and C_5 (blue) glycine dipeptide conformers in the 6.6–7.5 GHz frequency range (see ref 22). For C_7 , $\varphi = \pm 82.1^\circ$ and $\psi = \mp 64.2^\circ$; for C_5 , $\varphi = 180.0^\circ$ and $\psi = 180.0^\circ$.

stability of the two conformers, at any temperature relevant to specific astrochemical environments.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: vincenzo.barone@sns.it.

*E-mail: malgorzata.biczysko@pi.iccom.cnr.it.

*E-mail: cristina.puzzarini@unibo.it.

Funding

This work was supported by Italian MIUR (PRIN 2012, PON01-01078/8) and by the University of Bologna (RFO funds).

Notes

The authors declare no competing financial interest.

Biographies

Vincenzo Barone is Full Professor of Theoretical and Computational Chemistry at the Scuola Normale Superiore in Pisa (from 1994 to

2007 at the Federico II University of Naples), where he also coordinates the Dreams Center. He is a fellow of the Accademia dei Lincei and IAQMS and has been President of the Italian Chemical Society (SCI). His main research interests include DFT, solvation, and computational spectroscopy.

Malgorzata Biczysko obtained her Ph.D. in Theoretical Chemistry in 2000 from the University of Wrocław. Since 2014, she has been a researcher at Italian Research Council (ICCOM-CNR) in Pisa. She works on the development and application of computational protocols to simulate vibrational and vibronic spectra.

Cristina Puzzarini is Associate Professor at the University of Bologna (Dipartimento di Chimica “Giacomo Ciamician”). After obtaining a Ph.D. in Theoretical Chemistry (1997), in 2000 she moved back to the Millimeter-/Submillimeter-wave Spectroscopy lab, where in 1993, she got her M.Sc. degree in Chemistry. Her research lines cover both computational and experimental aspects in the field of High-Resolution Molecular Spectroscopy.

REFERENCES

- (1) Cheung, A. C.; Rank, D. M.; Townes, C. H.; Thornton, D. D.; Welch, W. J. Detection of NH_3 Molecules in the Interstellar Medium by Their Microwave Emission. *Phys. Rev. Lett.* **1968**, *21*, 1701–1705.
- (2) Müller, H. S. P.; 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*, 215–227.
- (3) Tielens, A. G. G. M. The Molecular Universe. *Rev. Mod. Phys.* **2012**, *85*, 1021–1081.
- (4) Herbst, E.; van Dishoeck, E. F. Complex Organic Interstellar Molecules. *Annu. Rev. Astron. Astrophys.* **2009**, *47*, 427–480.
- (5) Herbst, E.; Yates, J. T. Introduction: Astrochemistry. *Chem. Rev.* **2013**, *113*, 8707–8709.
- (6) Brown, W. A. Astrochemistry. *Phys. Chem. Chem. Phys.* **2014**, *16*, 3343–3343.
- (7) Pulay, P.; Meyer, W.; Boggs, J. E. Cubic Force Constants and Equilibrium Geometry of Methane from Hartree-Fock and Correlated Wavefunctions. *J. Chem. Phys.* **1978**, *68*, 5077–5085.
- (8) Pawłowski, F.; Jørgensen, P.; Olsen, J.; Hegelund, F.; Helgaker, T.; Gauss, J.; Bak, K. L.; Stanton, J. F. Molecular Equilibrium Structures from Experimental Rotational Constants and Calculated Vibration-Rotation Interaction Constants. *J. Chem. Phys.* **2002**, *116*, 6482–6496.
- (9) Demaison, J.; Boggs, J. E.; Császár, A. G., Eds. *Equilibrium Molecular Structures: From Spectroscopy to Quantum Chemistry*; CRC Press, Taylor & Francis Group: Boca Raton, FL, 2011.
- (10) Mills, I. M. In *Molecular Spectroscopy: Modern Research*; Rao, K. N., Mathews, C. W., Eds.; Academic: New York, 1972.
- (11) Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. The Performance of Composite Schemes and Hybrid CC/DFT Model in Predicting Structure, Thermodynamic and Spectroscopic Parameters: The Challenge of the Conformational Equilibrium in Glycine. *Phys. Chem. Chem. Phys.* **2013**, *15*, 10094–10111.
- (12) Piccardo, M.; Penocchio, E.; Puzzarini, C.; Biczysko, M.; Barone, V. Semi-Experimental Equilibrium Structure Determinations by Employing B3LYP/SNSD Anharmonic Force Fields: Validation and Application to Semirigid Organic Molecules. *J. Phys. Chem. A* **2015**, *119*, 2058–2082.
- (13) Puzzarini, C.; Stanton, J. F.; Gauss, J. Quantum-Chemical Calculation of Spectroscopic Parameters for Rotational Spectroscopy. *Int. Rev. Phys. Chem.* **2010**, *29*, 273–367.
- (14) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A Fifth-Order Perturbation Comparison of Electron Correlation Theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
- (15) Heckert, M.; Kállay, M.; Tew, D. P.; Klopper, W.; Gauss, J. Basis-Set Extrapolation Techniques for the Accurate Calculation of

Molecular Equilibrium Geometries Using Coupled-Cluster Theory. *J. Chem. Phys.* **2006**, *125*, No. 044108.

(16) Heckert, M.; Kállay, M.; Gauss, J. Molecular equilibrium geometries based on coupled-cluster calculations including quadruple excitations. *Mol. Phys.* **2005**, *103*, 2109–2115.

(17) Stanton, J. F., Gauss, J., Harding, M. E., Szalay, P. G. CFOUR, A quantum chemical program package. 2011; with contributions from A. A. Auer, R. J. Bartlett, U. Benedikt, C. Berger, D. E. Bernholdt, Y. J. Bomble, O. Christiansen, M. Heckert, O. Heun, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, K. Klein, W. J. Lauderdale, D. Matthews, T. Metzroth, L. A. Mueck, D. P. O'Neill, D. R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, S. Stopkiewicz, A. Tajti, J. Vázquez, F. Wang, J. D. Watts and the integral packages MOLECULE (J. Almloef and P. R. Taylor), PROPS (P. R. Taylor), ABACUS (T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wuelen. For the current version, see <http://www.cfour.de> (accessed September 13, 2012).

(18) Puzzarini, C. Extrapolation to the Complete Basis Set Limit of Structural Parameters: Comparison of Different Approaches. *J. Phys. Chem. A* **2009**, *113*, 14530–14535.

(19) Puzzarini, C.; Barone, V. Extending the Molecular Size in Accurate Quantum-Chemical Calculations: The Equilibrium Structure and Spectroscopic Properties of Uracil. *Phys. Chem. Chem. Phys.* **2011**, *13*, 7189–7197.

(20) Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. Accurate Molecular Structures and Infrared Spectra of trans-2,3-Dideuteriooxirane, Methyloxirane, and trans-2,3-Dimethyloxirane. *J. Chem. Phys.* **2014**, *141*, No. 034107.

(21) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.

(22) Puzzarini, C.; Biczysko, M.; V, B.; Largo, L.; Peña, I.; Cabezas, C.; Alonso, J. L. Accurate Characterization of the Peptide Linkage in the Gas Phase: A Joint Quantum-Chemical and Rotational Spectroscopy Study of the Glycine Dipeptide Analogue. *J. Phys. Chem. Lett.* **2014**, *5*, 534–540.

(23) Puzzarini, C. Rotational Spectroscopy Meets Theory. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6595–6607.

(24) Gaw, F., Willetts, A., Handy, N., Green, W. SPECTRO - a program for derivation of spectroscopic constants from provided quartic force fields and cubic dipole fields; In *Advances in Molecular Vibrations and Collision Dynamics*. Bowman, J. M., Ed.; JAI Press: Greenwich, CT, 1991; Vol. 1B; pp 169–185.

(25) Barone, V. Anharmonic Vibrational Properties by a Fully Automated Second-Order Perturbative Approach. *J. Chem. Phys.* **2005**, *122*, No. 014108.

(26) Puzzarini, C.; Biczysko, M.; Barone, V.; Peña, I.; Cabezas, C.; Alonso, J. L. Accurate Molecular Structure and Spectroscopic Properties of Nucleobases: A Combined Computational-Microwave Investigation of 2-Thiouracil as a Case Study. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16965–16975.

(27) Garrod, R. T. A Three-phase Chemical Model of Hot Cores: The Formation of Glycine. *Astrophys. J.* **2013**, *765*, No. 60.

(28) Császár, A. G.; Fabri, C.; Szidarovszky, T.; Matyus, E.; Furtenbacher, T.; Czako, G. The Fourth Age of Quantum Chemistry: Molecules in Motion. *Phys. Chem. Chem. Phys.* **2012**, *14*, 1085–1106.

(29) Christiansen, O. Selected New Developments in Vibrational Structure Theory: Potential Construction and Vibrational Wave Function Calculations. *Phys. Chem. Chem. Phys.* **2012**, *14*, 6672–6687.

(30) 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*, 1759–1787.

(31) Barone, V.; Baiardi, A.; Biczysko, M.; Bloino, J.; Cappelli, C.; Lipparini, F. Implementation and Validation of a Multi-purpose Virtual Spectrometer for Large Systems in Complex Environments. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12404–12422.

(32) Roy, T. K.; Gerber, R. B. Vibrational Self-Consistent Field Calculations for Spectroscopy of Biological Molecules: New Algorithmic Developments and Applications. *Phys. Chem. Chem. Phys.* **2013**, *15*, 9468–9492.

(33) Vázquez, J.; Stanton, J. F. Simple(r) Algebraic Equation for Transition Moments of Fundamental Transitions in Vibrational Second-Order Perturbation Theory. *Mol. Phys.* **2006**, *104*, 377–388.

(34) Bloino, J.; Barone, V. A Second-Order Perturbation Theory Route to Vibrational Averages and Transition Properties of Molecules: General Formulation and Application to Infrared and Vibrational Circular Dichroism Spectroscopies. *J. Chem. Phys.* **2012**, *136*, No. 124108.

(35) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision D.01; Gaussian, Inc.: Wallingford, CT, 2009.

(36) Boese, A. D.; Martin, J. Vibrational Spectra of the Azabenzenes Revisited: Anharmonic Force Fields. *J. Phys. Chem. A* **2004**, *108*, 3085–3096.

(37) Carbonniere, P.; Lucca, T.; Pouchan, C.; Rega, N.; Barone, V. Vibrational computations beyond the harmonic approximation: Performances of the B3LYP functional for semirigid molecules. *J. Comput. Chem.* **2005**, *26*, 384–388.

(38) 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*, 1015–1036.

(39) 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*, 3702–3710.

(40) Bernstein, L. S.; Lynch, D. K. Small Carbonaceous Molecules, Ethylene Oxide (c-C₂H₄O) and Cyclopropenylidene (c-C₃H₂): Sources of the Unidentified Infrared Bands? *Astrophys. J.* **2009**, *704*, No. 226.

(41) Kwok, S.; Zhang, Y. Mixed Aromatic-Aliphatic Organic Nanoparticles As Carriers of Unidentified Infrared Emission Features. *Nature* **2011**, *479*, 80–83.

(42) 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*, No. 107.

(43) Sloan, G. C.; Kraemer, K. E.; Price, S. D.; Shipman, R. F. A Uniform Database of 2.4–45.2 μm Spectra from the Infrared Space Observatory Short Wavelength Spectrometer. *Astrophys. J.* **2003**, *147* (Suppl. S), 379–401.

(44) Salas, J. B.; Pottasch, S. R.; Beintema, D. A.; Wesselius, P. R. The ISO-SWS Spectrum of Planetary Nebula NGC 7027. *Astron. Astrophys.* **2001**, *367*, 949–958.

(45) Biczysko, M.; Bloino, J.; Brancato, G.; Cacelli, I.; Cappelli, C.; Ferretti, A.; Lami, A.; Monti, S.; Pedone, A.; Prampolini, G.; Puzzarini, C.; Santoro, F.; Trani, F.; Villani, G. Integrated Computational Approaches for Spectroscopic Studies of Molecular Systems in the Gas Phase and in Solution: Pyrimidine As a Test Case. *Theor. Chem. Acc.* **2012**, *131*, No. 1201.

(46) Callahan, M. P.; Smith, K. E.; Cleaves, H. J.; Ruzicka, J.; Stern, J. C.; Glavin, D. P.; House, C. H.; Dworkin, J. P. Carbonaceous Meteorites Contain a Wide Range of Extraterrestrial Nucleobases. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 13995–13998.

(47) Materese, C. K.; Nuevo, M.; Bera, P. P.; Lee, T. J.; Sandford, S. A. Thymine and Other Prebiotic Molecules Produced from the

Ultraviolet Photo-irradiation of Pyrimidine in Simple Astrophysical Ice Analogs. *Astrobiology* **2013**, *13*, 948–962.

(48) Hörst, S. M.; Yelle, R. V.; Buch, A.; Carrasco, N.; Cernogora, G.; Dutuit, O.; Quirico, E.; Sciamma-O'Brien, E.; Smith, M. A.; Somogyi, A.; Szopa, C.; Thissen, R.; Vuitton, V. Formation of Amino Acids and Nucleotide Bases in a Titan Atmosphere Simulation Experiment. *Astrobiology* **2012**, *12*, 809–817.

(49) Cable, M. L.; Hörst, S. M.; Hodyss, R.; Beauchamp, P. M.; Smith, M. A.; Willis, P. A. Titan Tholins: Simulating Titan Organic Chemistry in the Cassini–Huygens Era. *Chem. Rev.* **2012**, *112*, 1882–1909.

(50) Ali, A.; Sittler, E. C., Jr.; Chornay, D.; Rowe, B. R.; Puzzarini, C. Cyclopropenyl Cation - the Simplest Huckel's Aromatic Molecule - and Its Cyclic Methyl Derivatives in Titan's Upper Atmosphere. *Planet. Space Sci.* **2013**, *87*, 96–105.

(51) Gautier, T.; Carrasco, N.; Mahjoub, A.; Vinatier, S.; Giuliani, A.; Szopa, C.; Anderson, C. M.; Correia, J.-J.; Dumas, P.; Cernogora, G. Mid- and Far-Infrared Absorption Spectroscopy of Titan's Aerosols Analogues. *Icarus* **2012**, *221*, 320–327.

(52) Fornaro, T.; Biczysko, M.; Monti, S.; Barone, V. Dispersion Corrected DFT Approaches for Anharmonic Vibrational Frequency Calculations: Nucleobases and Their Dimers. *Phys. Chem. Chem. Phys.* **2014**, *16*, 10112–10128.

(53) Elsilá, J. E.; Glavin, D. P.; Dworkin, J. P. Cometary Glycine Detected in Samples Returned by Stardust. *Meteorit. Planet. Sci.* **2009**, *44*, 1323–1330.

(54) Jimenez-Serra, I.; Testi, L.; Caselli, P.; Viti, S. Detectability of Glycine in Solar-type System Precursors. *Astrophys. J. Lett.* **2014**, *787*, No. L33.

(55) Godfrey, P. D.; Brown, R. D. Shape of Glycine. *J. Am. Chem. Soc.* **1995**, *117*, 2019–2023.

(56) Desai, R. C.; Kapra, R. *Dynamics of Self-Organized and Self-Assembled Structures*; Cambridge University Press: Cambridge, UK.

(57) Carnimeo, I.; Biczysko, M.; Bloino, J.; Barone, V. Reliable Structural, Thermodynamic, And Spectroscopic Properties of Organic Molecules Adsorbed on Silicon Surfaces from Computational Modeling: The Case of Glycine@Si(100). *Phys. Chem. Chem. Phys.* **2011**, *13*, 16713–16727.

(58) van Dishoeck, E. F.; Blake, G. A. Chemical Evolution of Star-Forming Regions. *Annu. Rev. Astron. Astrophys.* **1998**, *36*, 317–368.

(59) Marcelino, N.; Cernicharo, J.; Agúndez, M.; Roueff, E.; Gerin, M.; Martín-Pintado, J.; Mauersberger, R.; Thum, C. Discovery of Interstellar Propylene (CH₂CHCH₃): Missing Links in Interstellar Gas-Phase Chemistry. *Astrophys. J. Lett.* **2007**, *665*, L127–L130.

(60) 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*, No. L12.