The complex vibrational spectrum of proline explained through the adiabatically switched semiclassical initial value representation

Giacomo Botti, Chiara Aieta, and Riccardo Conte*

Dipartimento di Chimica, Università degli Studi di Milano, via Golgi 19, 20133 Milano, Italy

(Dated: April 5, 2022)

Proline, a 17-atom amino acid with a closed-ring side chain, has a complex potential energy surface characterized by several minima. Its IR experimental spectrum, reported in the literature, is of difficult and controversial assignment. In particular, the experimental signal at 3559 cm^{-1} associated with the OH stretch is interesting because it is inconsistent with the global minimum, *trans*-proline conformer. This suggests the possibility that multiple conformers may contribute to the IR spectrum. The same conclusion is obtained by investigating the splitting of the CO stretch at 1766 and 1789 cm^{-1} and other, more complex spectroscopic features involving CH stretches and COH/CNH bendings. In this work, we perform full-dimensional, on-the-fly adiabatically switched semiclassical initial value representation simulations employing the ab initio DFT-D3-B3LYP level of theory with aug-cc-pVDZ basis set. We reconstruct the experimental spectrum of proline in its main features by studying the vibrational features of *trans*-proline and *cis1*-proline, and provide a new assignment for the OH stretch of *trans*-proline.

* riccardo.conte1@unimi.it

I. INTRODUCTION

Proline is the only natural amino acid whose side chain forms a covalent bond with the amidic nitrogen, creating a pyrrolidinic ring and making proline formally an imino acid. The ring imposes strict constraints on the conformational space,[1] thus reducing the Ramachandran plot of both proline and the preceeding amino acid along the primary structure of the protein.[2] The conformational rigidity plays a fundamental role in protein stability and folding. This is the reason for the common presence of proline in specific positions of β sheets and γ-turns. Furthermore, proline gets organized in peculiar structures, like prolinic left elices, and takes part in the stress-response metabolism.[1, 3–7]

The particular structure of proline has also influenced the biochemical nomenclature, and a peptidic bond involving proline is commonly dubbed as "prolyl bond". It is well known that a generic peptidic bond is planar, and the two resulting *trans*- and *cis*- isomers are generally well separated in energy. The prolyl bond is different, because, due to steric hindrance of the pirrolidinic ring, the *trans*- and *cis*- isomers are pretty close in energy.[8] However, the presence of the "wrong" proline isomer in the tertiary structure may hinder protein folding, making the prolinic *cis/trans* isomerism a rate determining step in certain folding kinetics.[9, 10] Many enzimes that catalyze such isomerism employing a variety of mechanisms are present in nature. One involves partial stabilization of the nitrogen doublet through a hydrogen bond from the enzimatic pocket or a neighbouring residue.[9–11] This prompts N pyramidalization that reduces the isomerization barrier. In certain cases, the enzime stabilizes the peptide in a conformation in which the H atom comes from the peptide itself: this mechanism is a form of autocatalysis. [12–14]

Proline is also a key player in asymmetric catalysis due to its coordinating properties. For instance, in aldol condensation, a multi-step process, proline is faster in forming the enamine than any other secondary amine, thanks to the carboxylic H atom which stabilizes the transition state.[15] The mechanism of proton relay further facilitates enamine formation.[16] Then, in a second step of the process, a new transition state involving a single proline is formed, and the carboxylic H guides the electrophile through an H bond. This proton transfer is fundamental in aldol and Mannich reactions. [17–19].

The vibrational spectrum of proline was obtained by Adamowicz and coworkers in an Ar matrix at 14 K.[20] Their assignment was mostly based on two features: the OH *str* peak at

3559 cm^{-1} and the split C=O *str* peak at 1766 and 1789 cm^{-1}. Annealing analysis excluded any matrix effect in the splitting, suggesting the presence of two conformers in the spectrum. In agreement with the general nomenclature for amino acids, the main proline conformers were classified as type I, with the N−H···O=C hydrogen bond (see Fig.1, panel B), and type II, with the N···H−O hydrogen bond (see Fig.1, panel A). The OH *str* frequency is commonly used to distinguish between these two types of conformers: for type I the frequency is around 3560 cm^{-1}, whereas for type II it is typically around 3200 −3300 cm^{-1}. For this reason, the 3559 cm^{-1} signal was assigned to the OH *str* of a type I conformer, and the lack of any signal around 3200 −3300 cm^{-1} initially excluded the presence of any type II conformer. Therefore, the split in the C=O *str* seems to be due to two conformers, both with a non-hydrogen-bonded OH.

This qualitative picture, however, is at odds with theoretical energy calculations which predict a type II conformer as the most stable one. Those calculations were performed in the paper by Adamowicz et al. at MP4(SDTQ) and CCSD(T) level with different basis sets at the equilibrium geometry found using the MP2/aug-cc-pVDZ level of theory and basis set. The harmonic frequencies and zero-point energy (ZPE) were obtained at DFT-B3LYP/aug-cc-pVDZ level. The predicted scaled harmonic OH *str* frequency returned a value close to that of the CH band, suggesting that the OH *str* could be identified by subtracting the CH contribution from the spectrum. This was done by means of the deuterated proline spectrum obtained at the same temperature conditions. In spite of a very weak spectral difference signal, those authors concluded that the OH *str* of a type II conformer was actually present at 3025 cm^{-1}, and that both type I and type II conformers were part of the original spectrum.

In a very preliminar and partial simulation, two of us employed a type II proline while developing the on-the-fly adiabatically switched semiclassical initial value representation (AS SCIVR) method for spectroscopy.[21] However, there was no agreement between AS SCIVR, scaled harmonic, and previously assigned OH *str* frequencies. Due to this controversial result, we found it interesting to study the vibrational spectrum of proline in details by performing a thorough spectral analysis using AS SCIVR.

Semiclassical vibrational spectroscopy has largely been proven as a reliable method for vibrational and vibronic analysis and spectral assignment, especially in presence of high anharmonicity and quantum effects.[22–35] In recent years, several semiclassical theories have been developed and applied to systems of various complexity and dimensionality.[36–

50] A breakthrough in the spectroscopy of moderate and large dimensional systems has been achieved upon introduction of the divide-and-conquer semiclassical initial value representation (DC SCIVR).[51] This approach builds on the previous semiclassical time-averaged method by Kaledin and Miller[52, 53] and on Ceotto's semiclassical multiple coherent approach.[54, 55] DC SCIVR makes high dimensional systems treatable with high accuracy by defining appropriate subspaces in which to perform the semiclassical calculations in reduced dimensionality.[56–58] In this way a sensible spectroscopic signal is obtained and the issue of the noisy and non-assignable full-dimensional spectrum overcome. A novel semiclassical approach is represented by AS SCIVR.[21, 59] This method is based on a pre-treatment of semiclassical trajectories performed by means of the adiabatic switching technique.[60–65] In practice, the starting conditions of the trajectory are chosen from harmonic quantization, in agreement with other semiclassical methods. However, in AS SCIVR the trajectories are evolved under the increasing perturbative influence of the true vibrational Hamiltonian. Only when the latter is completely switched on the semiclassical run starts. If the perturbation is switched on slowly, then action variables are conserved by virtue of the adiabatic theorem and the trajectories are representative of semiclassical quantization and numerically more stable. This leads to the possibility to get an assignable full-dimensional spectrum for larger systems, a better estimate of vibrational frequencies, and more precise spectral features. AS SCIVR involves either a trajectory distribution, when an analytical potential energy surface (PES) is available, or a single ab initio "on-the-fly" trajectory when a PES is not available. The latter has been the case for the full-dimensional proline study here presented.

The paper is organized as follows: In Sec. II the theoretical aspects of AS SCIVR are briefly reviewed. In Sec. III the main regions of proline IR spectrum are analyzed and discussed. Finally Sec. IV is devoted to a brief summary and conclusions.

II. THEORETICAL METHODS

We adopted the on-the-fly adiabatically switched semiclassical initial value representation to determine the vibrational features of proline. AS SCIVR is the most recently developed among semiclassical methods for calculation of vibrational density of states (power spectra) and determination of quantum vibrational frequencies.[59] In its "on-the-fly" version AS

SCIVR is able to reproduce the spectral features from a single trajectory.[21] This feature is not new to SCIVR methods being already present in Ceotto's multiple coherent approach. We employed Grimme's DFT-D3[66] with B3LYP exchange-correlation functional and aug-cc-pVDZ basis set for both geometry optimizations and *ab initio* on-the-fly dynamics. This choice has been driven by the necessity to keep computational times affordable and the recognised improvement in H-bonded description when empirical diffusion corrections are included in DFT calculations.

The peculiarity of AS SCIVR is that it is based on a two-step procedure. In the first part, initial phase-space trajectory conditions ($\mathbf{p}_0$, $\mathbf{q}_0$) are chosen from the usual harmonic quantization

$$\begin{aligned} q_{0,i} &= q_{eq,i} \\ p_{0,i} &= p_{eq,i} = \sum_i \hbar\omega_i(\nu_i + 1/2), \end{aligned} \tag{1}$$

where ν_i is the vibrational quantum number of the i-th normal mode of vibration, ω_i is the associate harmonic frequency, and $\mathbf{q}_{eq}$ is the equilibrium geometry. The momenta associated to the equilibrium geometry according to the above quantization are for brevity also called equilibrium momenta ($\mathbf{p}_{eq}$). So, if ground state quantization is demanded, all ν_i are set to 0. Eq.(1) is used for single-trajectory "on-the-fly" simulations. When working with analytical potential energy surfaces[67] and a large number of trajectories, an appropriate angular phase is added to the definition of $q_{0,i}$ and $p_{0,i}$ to generate a set of different trajectories characterized by the same harmonic energy.[59]

Then, the dynamics is propagated through the adiabatic switching Hamiltonian (H_{as}) by switching on slowly the true anharmonic vibrational Hamiltonian of the system (H_{anh}), i.e.

$$H_{\text{as}} = [1 - \lambda(t)]\, H_{\text{harm}} + \lambda(t) H_{\text{anh}}, \tag{2}$$

where λ is the following switching function chosen in agreement with the literature[64]

$$\lambda(t) = \frac{t}{T_{\text{as}}} - \frac{1}{2\pi} \sin\left(\frac{2\pi t}{T_{\text{as}}}\right), \tag{3}$$

and T_{as} is the total length of the adiabatic switching trajectory. If the switching is slow, then action variables are conserved due to the classical adiabatic theorem.[68] In agreement with previous AS-SCIVR tests and calculations, we chose $T_{\text{as}} = 25\,000$ au (about 0.6 ps)

with a time step of 10 au. In the second part of the AS-SCIVR procedure, the adiabatic switching run is followed by an SC dynamics of equal length and time step started from the final phase space coordinates obtained from the previous AS procedure $(\mathbf{p}_{\rm as}, \mathbf{q}_{\rm as})$. By indicating the momenta and positions along this SC dynamics as

$$\mathbf{p}'_t(\mathbf{p}_{\rm as},\ \mathbf{q}_{\rm as}), \mathbf{q}'_t(\mathbf{p}_{\rm as}, \mathbf{q}_{\rm as}) \tag{4}$$

we can write the AS-SCIVR working formula[21] for the vibrational spectral density $I_{\rm as}(E)$

$$I_{\rm as}(E; \mathbf{p}_{\rm as}, \mathbf{q}_{\rm as}) = \left(\frac{1}{2\pi\hbar}\right)^{N_v} \frac{1}{2\pi\hbar T} \left| \int_0^T dt e^{\frac{i}{\hbar}[S_t(\mathbf{p}_{\rm as},\mathbf{q}_{\rm as})+Et+\phi_t(\mathbf{p}_{\rm as},\mathbf{q}_{\rm as})]} \langle \Psi(\mathbf{p}_{eq}, \mathbf{q}_{eq}) | g_t(\mathbf{p}'_t, \mathbf{q}'_t) \rangle \right|^2, \tag{5}$$

where N_v=45 is the number of vibrational degrees of freedom in proline, and T the simulation length of the semiclassical dynamics. The power spectrum depends on the AS conditions $(\mathbf{p}_{\rm as}, \mathbf{q}_{\rm as})$, via the evolved positions and momenta $(\mathbf{p}'_t, \mathbf{q}'_t)$ appearing in the instantaneous classical action (S_t), the phase of the prefactor (ϕ_t), and the coherent state $|g_t\rangle$.

As anticipated, the potential energy V is calculated ab initio in an "on-the-fly" fashion. The instantaneous action is defined as usual

$$S_t(\mathbf{p}'_t, \mathbf{q}'_t) = \int_0^t \left(\frac{\mathbf{p}'^2_{t'}}{2} - V(\mathbf{q}'_{t'}) \right) dt'. \tag{6}$$

The phase of the prefactor is[52]

$$\phi_t = \text{phase} \left[\sqrt{\left| \frac{1}{2} \left(\frac{\partial \mathbf{q}'_t}{\partial \mathbf{q}_{\rm as}} + \Gamma^{-1} \frac{\partial \mathbf{p}'_t}{\partial \mathbf{p}_{\rm as}} \Gamma - i\hbar \frac{\partial \mathbf{q}'_t}{\partial \mathbf{p}_{\rm as}} \Gamma + \frac{i\Gamma^{-1}}{\hbar} \frac{\partial \mathbf{p}'_t}{\partial \mathbf{q}_{\rm as}} \right) \right|} \right], \tag{7}$$

which requires evaluation of the monodromy matrix elements along the dynamics and, consequently, of the instantaneous Hessian matrix. Γ is the width matrix usually chosen to be a diagonal matrix with elements equal to the harmonic frequencies of vibration.
Coherent states are defined in the configuration space as[52]

$$\langle \tilde{\mathbf{q}} | g(\mathbf{p}', \mathbf{q}') \rangle = \left(\frac{\det(\Gamma)}{\pi^{N_v}} \right)^{1/4} \exp \left\{ -(\mathbf{q}' - \tilde{\mathbf{q}})^T \frac{\Gamma}{2} (\mathbf{q}' - \tilde{\mathbf{q}}) + \frac{i}{\hbar} \mathbf{p}'^T (\tilde{\mathbf{q}} - \mathbf{q}') \right\} \tag{8}$$

The reference state $|\Psi(\mathbf{p}_{eq}, \mathbf{q}_{eq})\rangle$ is built from equilibrium positions and momenta $(\mathbf{p}_{eq}, \mathbf{q}_{eq})$, as in the multiple coherent approach (we now write explicitly 45 for the number of vibrational degrees of freedom of proline)[54]

$$|\Psi(\mathbf{p}_{eq}, \mathbf{q}_{eq})\rangle = \prod_{j=1}^{45} \epsilon_{1,j} |p_{eq,j}, q_{eq,j}\rangle + \epsilon_{2,j} |-p_{eq,j}, q_{eq,j}\rangle, \tag{9}$$

and the full-dimensional character of the calculation can be appreciated. $|p_{eq,j}, q_{eq,j}\rangle$ and $|-p_{eq,j}, q_{eq,j}\rangle$ are coherent states, and $\epsilon_{1,j}$ and $\epsilon_{2,j}$ are the coefficients used to enforce parity. By setting both $\epsilon_{1,j}$ and $\epsilon_{2,j}$ equal to 1 for all j the peaks associated to the ZPE and to states with an even number of quanta of excitation are reproduced. The peak relative to the vibrational state with one quantum of excitation in mode j can be obtained by changing $\epsilon_{2,j}$ from 1 to -1. All spectral peaks are centered at energy values which give an estimate of the quantum energy of the corresponding state. Quantum frequency of vibrations are related to the transitions between vibrational states and can be obtained as energy differences between the energies of the two states involved in the transition. Therefore, by managing the coherent states in the reference state, one is capable to extract all vibrational features even from a single ZPE trajectory. However, similarly to the multiple coherent approach, in some instances refinement of results may be necessary. This can be achieved by running a calculation based on a trajectory characterized by a quantum of excitation in the target vibrational state. We notice that Eq.(5) provides the vibrational density of states as a function of the absolute quantum vibrational energy E. By setting the zero of energy at the ZPE level it is possible to label the density of states as $I(\omega)$ and compare the calculated frequencies with the experimental ones.

For additional and supporting calculations regarding the absorption intensity of the spectral signals we employed two approaches. One is the usual (double) harmonic approximation provided by the NWChem software for electronic calculations.[69] The other one is based on the Fourier transform of the dipole-dipole autocorrelation function

$$S(\omega) \propto \int_0^T e^{i\omega t} \langle \mu(t)\mu(0) \rangle dt. \tag{10}$$

T is the total simulation time, corresponding in our simulations to that of the semiclassical part of the dynamics, and μ is the molecular dipole moment. Eq.(10) provides a classical estimate of the absorption intensity (commonly indicated with the time-independent symbol S and not to be confused with the time-dependent instantaneous action indicated by a similar symbol in Eq.(6)) and includes anharmonic effects.

III. RESULTS AND DISCUSSION

In this Section we present the results of our investigation performed by means of the ab initio "on-the-fly" AS-SCIVR method at DFT-D3-B3LYP/aug-cc-pVDZ level of theory. This choice was substantiated by energetics calculations performed for trans- and cis1- proline at MP2 and CCSD(T) level of theory demonstrating the good accuracy of DFT-D3. We report these results in Section 3 of the Supplementary Material. Furthermore, previous semiclassical work on glycine undertaken at the same level of theory and same basis set provided very accurate results in comparison to both MP2-based calculations and experimental data.[70] One trajectory with initial harmonic ZPE quantization has been employed for the global minimum (GM) *trans*-proline and the same procedure (using the appropriate harmonic ZPE quantization) has been adopted for another conformer labeled as *cis*1-proline. In both cases the total evolution time was set to 50 000 au (approximately 1.2 ps) divided into 25 000 au for the adiabatic switching run, and the remaining 25 000 au for the semiclassical dynamics. The time step was chosen equal to 10 au (about 0.25 fs). Following a well-consolidated procedure in semiclassical dynamics,[54, 71] to determine more accurately the OH *str* frequency of *cis*1-proline a simulation starting with one quantum of energy in mode 45 (corresponding in normal mode approximation to the OH *str* of *cis*1-proline) was run. AS-SCIVR simulations provide vibrational spectral densities, so they include fundamental transitions, overtones, combination bands, and Fermi resonances. The reference experimental values are taken from Ref. 20. In that work matrix-isolation IR spectroscopy combined with ab initio calculations was used to investigate conformations of gas-phase amino acids frozen in low temperature inert gas matrices.

A. Energy and optimization calculations

To start our investigation we performed a search of the main minima characterizing the potential energy surface of proline. As anticipated, we employed the DFT-D3-B3LYP level of theory with Dunning's aug-cc-pVDZ basis set. Following Blagoi,[72] we classified the conformers on the basis of the relative position of the NH and COOH groups with respect to the ring: *trans*- conformers have them pointing in opposite directions, while *cis*- conformers have them pointing in the same direction. An additional classification to distinguish within

trans- and *cis*- conformers is based on the type of hydrogen bond formed by the conformer. This standard classification labels three main conformations: one with the N−H···O=C hydrogen bond (Fig.1, panel B), one with the N···H−O hydrogen bond (Fig.1, panel A), and the last one with the N−H···O−H hydrogen bond (Fig.1, panel C).[73] These correspond, respectively, to Adamowicz's type I, II, and III conformers. Using this classification we identified *trans*-proline as a type II conformer, *cis*1-proline as a type III conformer, and *cis*2-proline as a type I conformer. When performing the semiclassical part of the dynamics (at anharmonic zero-point energy) for *cis*2-proline, a quick interconversion to a higher electronic energy conformer takes place. We labeled this additional conformer, characterized by the absence of the internal hydrogen bond, as *trans*2-proline (Fig.1, panel D). Owing to this quick interconversion of *cis*2-proline, we employed *cis*1-proline in our simulations for the *cis*-type conformer. The ensemble of the four conformers is presented in Fig.1, while their Cartesian geometries and harmonic frequencies are listed in the Supplementary Material file. Table I reports the absolute and relative electronic energy values of the 4 conformers. Adiabatic electronic energies including harmonic and anharmonic AS-SCIVR ZPE estimates provide a similar picture and can be found in the Supplementary Material file.

TABLE I. Absolute and relative electronic energies of the four main conformers of proline at DFT-D3-B3LYP/aug-cc-pVDZ level of theory. GM stands for global minimum.

Conformer	Energy (E_h)	Energy relative to GM (cm^{-1})
trans-proline (GM)	-401.23899	0
*cis*2-proline	-401.23589	679
*cis*1-proline	-401.23364	1172
*trans*2-proline	-401.23125	1697

B. The high frequency range: OH and NH stretches

In the high frequency region, the spectrum of proline is characterized by the OH and NH stretches. The sorting of harmonic frequencies for these vibrations is mixed and depends on the type of conformer. *Trans*-proline (Fig.1, panel A) has the OH *str* weakened by the internal hydrogen bond, so its highest harmonic frequency is related to the NH *str*. On the

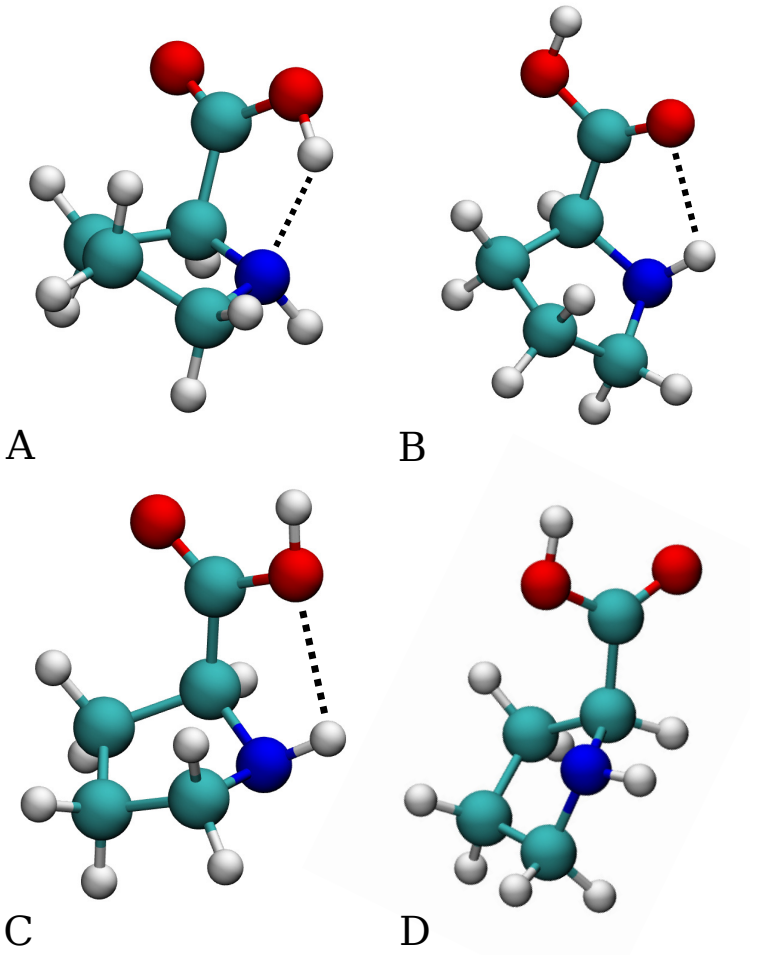

FIG. 1. A: *trans*-proline, B: *cis*2-proline, C: *cis*1-proline and D: *trans*2-proline. The dashed line indicates the intramolecular hydrogen bond.

other hand, the two *cis*- prolines (Fig.1, panel B and C) have the OH *str* at higher frequency. Harmonic estimates are reported in Table II from which it is clear that, considering the expected red shift due to anharmonicity, the OH *str* experimental signal at 3559 cm^{-1} can not be assigned to *trans*-proline. Therefore, a second conformer must contribute to the spectrum and, as anticipated, given the fast interconversion of *cis*2-proline to *trans*2-proline we focused on *cis*1-proline.

Fig.2 reports the AS-SCIVR density of states simulations for the OH stretch. These provide signals at 3329 cm^{-1} and 3522 cm^{-1} for the *trans*- and *cis*1- conformer respectively, confirming that the experimental peak at 3559 cm^{-1} is representative of the OH stretch of the *cis*1- conformer. Surprisingly, there is no evidence in the experimental spectrum of the OH signal relative to the *trans*-conformer. This is very unexpected since looking at

TABLE II. Double harmonic approximation (frequency ω, and intensity S) for the OH and NH stretches of the four main conformers of proline. Calculations performed at DFT-D3-B3LYP/aug-cc-pVDZ level of theory. Frequencies are in cm^{-1} and intensities in $km\,mol^{-1}$.

Conformer	ω_{OH}	ω_{NH}	S_{OH}	S_{NH}
trans-proline	3355	3560	346.09	6.64
cis1-proline	3740	3548	69.06	23.02
cis2-proline	3742	3520	65.94	1.28
trans2-proline	3744	3529	66.99	6.04

Table II the most intense OH signal should be the *trans*-proline one. As anticipated in the Introduction, a previous assignment in the literature indicated 3025 cm^{-1} for the frequency of the OH stretch of the *trans*- conformer.[20] Our estimate is quite different at 3329 cm^{-1}.

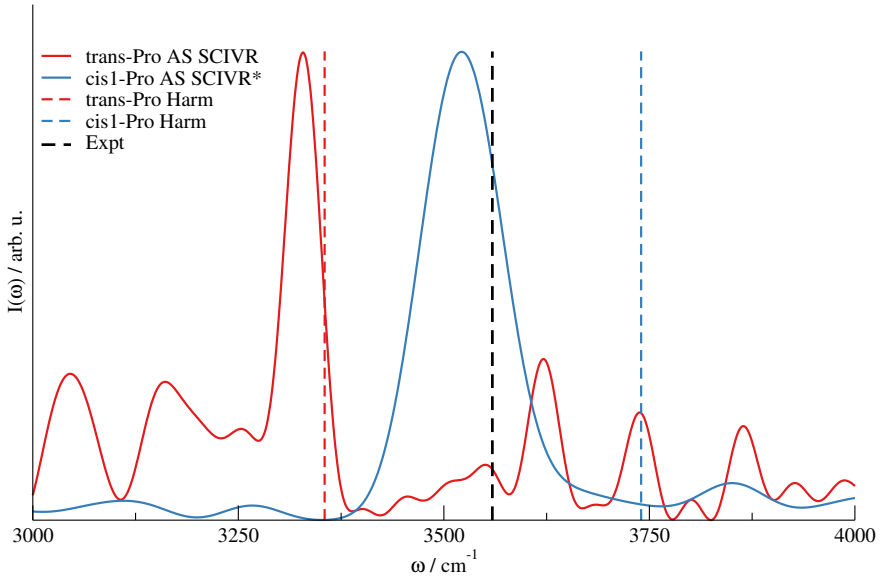

FIG. 2. AS-SCIVR density of states simulations for the OH stretches of *trans*-proline (red) and *cis*1-proline (blue). The corresponding harmonic frequencies are indicated by same-color vertical dashed lines. The experimental value for the *cis*1-proline OH stretch (from Ref. 20) is represented by the dashed black line. No experimental value for the *trans*-proline OH stretch is shown due to controversial assignment (see text). * excited trajectory employed (see text).

To clarify this point and the issue of the weak OH signal of *trans*-proline, we performed two calculations of the absorption intensity. One was obtained from the already presented

double harmonic estimate provided by the NWChem software. The other one was the result of a simulation based on the Fourier transform of the autocorrelation dipole-dipole function calculated on the semiclassical part of the same ZPE trajectory employed for the AS-SCIVR simulation of *trans*-proline. Therefore, we focused on the CO *str* and OH *str* of *trans*-proline. The CO *str* was chosen as a reference because it is characterized by a very intense signal. Fig.3 shows a comparison between the experimental IR spectrum and the simulations. From that figure it is clear that at the harmonic level the OH *str* is almost as intense as the CO one. However, once anharmonicities are taken into account, the intensity of the OH *str* decreases dramatically and the reason for the "missing" signal is eventually clarified. Furthermore, the classical anharmonic calculation is in agreement with the AS-SCIVR estimate for the frequency of vibration of the OH *str*.

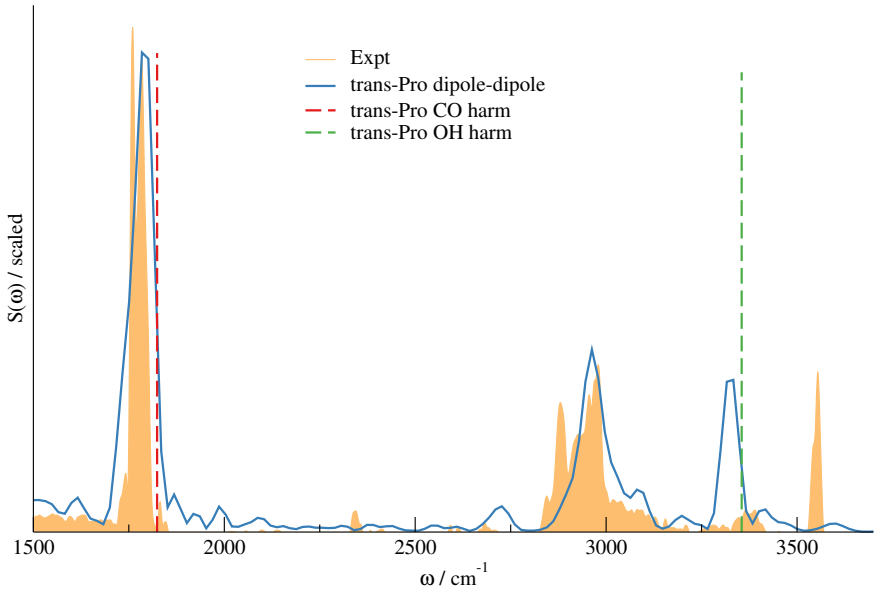

FIG. 3. *Trans*-proline: Experimental IR spectrum from Ref. 20 (orange area), double harmonic approximation for the CO (red dashed line) and OH (green dashed line) stretches, and Fourier transform of the classical dipole-dipole autocorrelation function (blue). Harmonic and classical intensity estimates have been scaled to match the experimental CO *str*.

Other signals of low intensity well above $3000\,\mathrm{cm}^{-1}$ in the experimental proline spectrum are related to the NH stretches. The two investigated conformers are characterized by AS-SCIVR frequencies for the NH *str* close to each other at around $3460\,\mathrm{cm}^{-1}$. This overestimates by about 70-90 cm^{-1} the experimentally assigned frequencies. We could have

obtained a better agreement (but at much additional computational cost!) by running excited trajectories as in the case of the OH *str* of *cis*1-proline. However, we decided not to pursue this strategy because we were already able to distinguish the two NH *str* from the OH *str* and we could confirm the experimental assignment of these low intensity modes.

C. The band of CH stretches

Moving to lower frequencies in the experimental spectrum the band of CH stretches is found. It includes seven normal modes of vibration for each conformer. The experimental band is rather extended in the frequency domain (about 150 cm^{-1}). Furthermore, it features an interesting double peak at 2885 and 2984 cm^{-1} (see Fig.3).

To better understand the origin of this spectroscopic feature we studied the CH modes of both *trans*- and *cis*1-proline. Calculated harmonic and AS-SCIVR frequency values are reported in Table III.

TABLE III. Harmonic and AS-SCIVR frequencies for the seven CH stretch modes (modes 37-43) of *trans*-proline (column 2 and 3) and *cis*1-proline (column 4 and 5). Calculations were performed at DFT-D3-B3LYP/aug-cc-pVDZ level of theory. All frequencies are in cm^{-1}

CH *str*	*trans*-proline		*cis*1-proline	
	ω_{harm}	ω_{ASSCIVR}	ω_{harm}	ω_{ASSCIVR}
mode 43	3135	3071	3138	3078
mode 42	3117	3060	3109	2953
mode 41	3086	2970	3090	2887
mode 40	3074	2949	3085	2982
mode 39	3057	2938	3065	2920
mode 38	3050	3017	3034	2946
mode 37	3003	2911	2941	2883

Analysis of the AS-SCIVR frequencies confirms the complexity of this band. The *trans*-proline frequencies are generally at higher values than the ones of the *cis*1 conformer. So, while the CH *str* signals of the *trans*- conformer seem to be responsible for the higher-frequency part of the band, the lower-frequency part is better described by the *cis*1- con-

former. This is even more evident from Fig.4, where mode 41 is plotted for both conformers. We notice that a couple of modes (modes 42 and 43) are estimated at frequencies somewhat higher than the experimentally assigned range, which extends between 2846 and 2984 cm^{-1}. Also in this case, agreement could have been improved by running additional excited trajectories, but this was not necessary since the description of the band given by the two basic AS-SCIVR simulations was already satisfactory and a remarkable computational overhead has been saved.

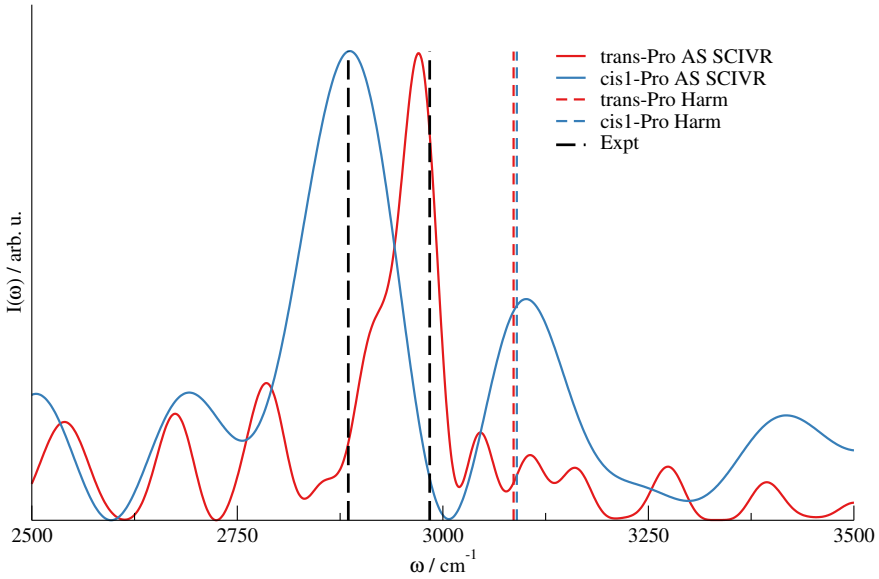

FIG. 4. AS-SCIVR density of states simulations for two representative CH modes of *trans*-proline (red) and *cis*1-proline (blue). The corresponding harmonic frequencies are indicated by same-color vertical dashed lines. Experimental values (from Ref. 20) are reported with black dashed lines.

D. The mid frequency range: The CO stretch and the COH/CNH bendings

The next relevant feature in the vibrational spectrum of proline is represented by the CO stretch. The experimental spectrum shows again a double peak signal at 1766 and 1789 cm^{-1} (see Fig.3). However, this time the gap (about 20 cm^{-1}) is much narrower than in the case of the OH *str* and CH *str*. In spite of this, we were able to resolve the double peak by means of our calculations for the CO *str* of *trans*- and *cis*1-proline. Results are plotted in Fig.5 and the AS-SCIVR estimates at 1766 and 1799 cm^{-1} are in excellent agreement with the experimental findings at 1766 and 1789 cm^{-1}, respectively.

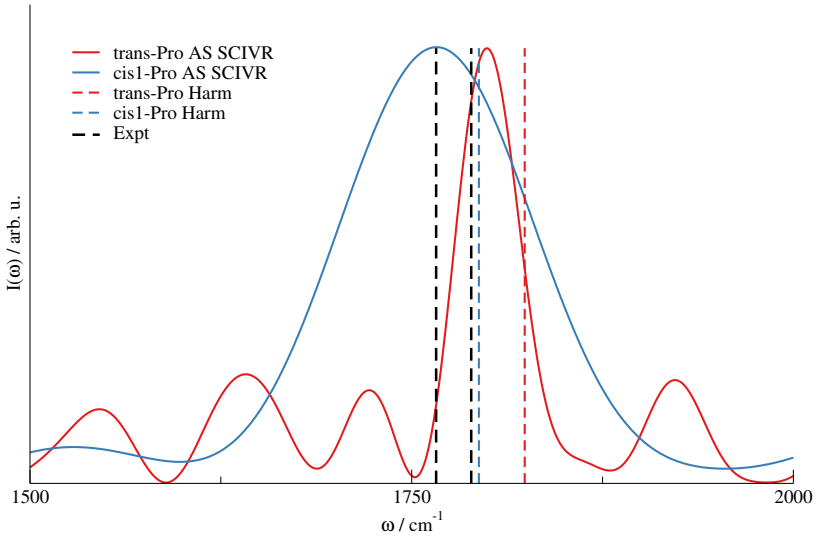

FIG. 5. AS-SCIVR density of states simulations for the CO stretches of *trans*-proline (red) and *cis*1-proline (blue). The corresponding harmonic frequencies are indicated by same-color vertical dashed lines. Experimental values (from Ref. 20) are reported with black dashed lines.

Finally, we focused on the region just below and around 1400 cm^{-1}, where a double peak structure with frequencies at 1384 and 1405 cm^{-1} is found in the experimental IR spectrum. Our AS-SCIVR calculations are reported in Fig.6 and we identified two very mixed motions that we could associate in normal mode approximation to the COH and CNH bendings. Our assignment is that the two CNH bendings, one of *trans*-proline and one of *cis*1-proline, are responsible for these two experimental peaks. We were also able to assign other two low-intensity experimental signals at 1284 and 1412 cm^{-1} to the two COH bendings with the lower-frequency one belonging to *cis*1-proline.

IV. SUMMARY AND CONCLUSIONS

In this paper we examined the main features of the complex vibrational spectrum of proline by means of the adiabatically switched semiclassical initial value representation technique. Table IV reports the principal results obtained for *trans*-proline and *cis*1-proline. Complete data can be found in the Supplementary Material file.

We started our investigation having in mind to solve the open issue concerning the assignment of the fundamental transition associated to the OH *str* of the *trans*-proline conformer. That signal was expected to be very bright, as confirmed by double harmonic calcula-

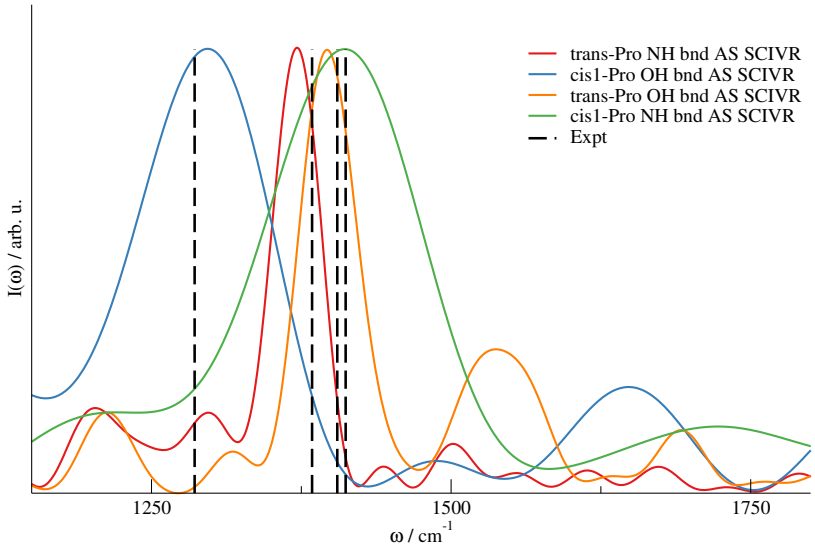

FIG. 6. AS-SCIVR density of states simulations for the COH and CNH bendings of *trans*-proline (orange and red) and *cis*1-proline (blue and green). Experimental values (from Ref. 20) are reported with black dashed lines.

tions, but was actually not evident in the experimental spectrum. The only OH *str* easily assignable was the one referred to a *cis*- type conformer at 3559 cm^{-1}. In the literature the OH *str* of *trans*-proline was assigned at 3025 cm^{-1} as the result of a difference spectrum upon selective deuteration.[20] We were able instead to assign the signal at 3329 cm^{-1} and a dynamical calculation of the absorption intensity, based on the Fourier transform of the dipole-dipole autocorrelation function, corroborated that conclusion revealing the crucial role of anharmonicity.

Other interesting features of the experimental spectrum that we were also able to describe by performing simulations for both *trans*- and *cis*1-proline include a set of split bands and peaks. For instance, in the case of the CO *str* a split signal is a clear clue of spectral contribution by at least two conformers, since a single CO bond is present in proline. Likewise, the signal associated to the COH/CNH bendings displays a double-peak and even this was accurately reproduced by means of our simulations.

We performed *ab initio* optimizations and electronic energy calculations to better characterize the potential energy surface. In agreement with the previous literature we found a global minimum conformer of *trans* geometry, a *cis*2 conformer at higher energy, and a *cis*1 conformer at even higher energy. However, when evolving the AS-SCIVR dynamics with

TABLE IV. Calculated AS-SCIVR frequencies of vibration for *trans*-proline and *cis*1-proline at DFT-D3-B3LYP/aug-cc-pVDZ level of theory. Experimental values are given in the last column. All values are in cm^{-1}.

Motion	*trans*-proline	*cis*1-proline	Exp[a]
OH *str*	3329	3522	3559 (*cis*)
NH *str*	3461	3467	3393,3369
CH *str*	3071	3078	
CH *str*	3060	2953	
CH *str*	2970	2887	
CH *str*	2949	2982	[2846,2984]
CH *str*	2938	2920	
CH *str*	3017	2946	
CH *str*	2911	2883	
CO *str*	1799	1766	1789,1766
NH *bnd*	1371	1411	1384,1405
OH *bnd*	1397	1297	1412,1286

[a] from Ref. 20. [2846,2984] indicates the frequency range of the experimental CH band.

zero-point energy, we found that the *cis*2 conformer quickly interconverted into a *trans*2 conformer higher in energy than *cis*1-proline. For this reason our calculations on the *cis* species were focused on the more stable *cis*1-proline. To strengthen this choice we probed the *cis*2-/*trans*2-proline system by means of an additional AS-SCIVR simulation not presented here. The resulting simulation provided more elaborated peaks, as expected on the basis of the more complex dynamics, but frequency values for the main modes of vibration were in very close agreement with those found for *cis*1-proline.

In general, our on-the-fly AS-SCIVR simulations were able to reproduce the experimental results with good accuracy. We stress that our results are fully (45-) dimensional and were mainly based on just two short-time semiclassical trajectories, as established by on-the-fly semiclassical theories. Evolution of these trajectories, one for each of the two conformers investigated, include Hessian calculations. We employed a third trajectory starting the

adiabatic switching procedure with a quantum of excitation in the OH *str* mode of *cis*1-proline only to refine the estimate of that specific frequency. Use of additional AS-SCIVR trajectories with one quantum of excitation in the NH and CH stretches would have probably led to a better agreement with the experiment, but we reckoned it was not necessary since assignment of those signals was not controversial. The semi-quantitative picture we got for NH and CH stretches is fully satisfactory and running several additional trajectories would have come at a high computational cost. For lower energy signals (CO *str* and COH/CNH bendings) our frequency estimates are in excellent agreement with the experiment. Finally, we also notice that our simulations provide narrower and more precise peaks for *trans*-proline and wider ones for *cis*1-proline. This is mainly an effect of the more energetically isolated well for the global minimum *trans*-proline.

Given the peculiar behavior of proline in protein folding and its primary role in asymmetric catalysis, it will be interesting to further investigate the interplay between *trans*- and *cis*- isomers and their interconversion mechanisms. We will focus on that in our future work.

SUPPLEMENTARY MATERIAL

The supplementary material contains the equilibrium geometries of the four conformers, the values of their frequencies and intensities in double harmonic approximation, energetics calculations for *trans*-proline and *cis1*-proline, complete AS-SCIVR data for *trans*-proline and *cis1*-proline, and estimates of the ZPEs and relative electronic energies including the ZPEs for the four conformers.

ACKNOWLEDGMENTS

Authors thank warmly Prof. Michele Ceotto for carefully reading the manuscript and useful comments. G.B. thanks Istituto Nazionale Previdenza Sociale (INPS) for funding his Ph.D. scholarship.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflict of interest to disclose.

DATA AVAILABILITY

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

---

[1] M. W. Macarthur and J. M. Thornton, J. Mol. Biol **218**, 397 (1991).

[2] N. L. Summerst and M. Karplus, J Mol. Biol **216**, 991 (1990).

[3] I. D. Reva, S. G. Stepanian, A. M. Plokhotnichenko, E. D. Radchenko, G. G. Sheina, and Y. P. Blagoi, J. Mol. Struct. **3**, 1 (1994).

[4] E. Czinki and A. G. Császár, Chem. Eur. J. **9**, 1008 (2003).

[5] J. Richardson and D. Richardson, Science **240**, 1648 (1988).

[6] A. A. Adzhubei and M. J. E. Sternberg, J. Mol. Biol. **229**, 472 (1993).

[7] J. Schmidt and S. R. Kass, J. Phys. Chem. A **117**, 4863 (2013).

[8] J. W. Silzel, T. A. Murphree, R. K. Paranji, M. M. Guttman, and R. R. Julian, J. Am. Soc. Mass Spectrom. **31**, 1974 (2020).

[9] G. Fischer, Chem. Soc. Rev. **29**, 119 (2000).

[10] U. Reimer, E. Mokdad, M. Schutkowski, and G. Fischer, Biochemistry **36**, 13802 (1997).

[11] J. Fanghänel and G. Fischer, Frontiers in Bioscience **9**, 3453 (2004).

[12] S. Fischer, R. L. J. Dunbrack, and M. Karplus, J. Am. Chem. Soc. **116**, 11931 (1994).

[13] K. B. Wiberg and K. E. Laidig, J. Am. Chem. Soc. **109**, 5935 (1987).

[14] S. Fischer, S. Michnick, and M. Karplus, Biochem. **32**, 13830 (1993).

[15] S. Bahmanyar and K. N. Houk, "The origin of stereoselectivity in proline-catalyzed intramolecular aldol reactions," (2001).

[16] C. Allemann, R. Gordillo, F. R. Clemente, P. H. Y. Cheong, and K. N. Houk, Acc. Chem. Res. **37**, 558 (2004).

[17] L. Hoang, S. Bahmanyar, K. N. Houk, and B. List, J. Am. Chem. Soc. **125**, 16 (2003).

[18] M. Nielsen, D. Worgull, T. Zweifel, B. Gschwend, S. Bertelsen, and K. A. Jørgensen, Chem. Commun. **47**, 632 (2011).

[19] S. Bahmanyar, K. N. Houk, H. J. Martin, and B. List, J. Am. Chem. Soc. **125**, 2475 (2003).

[20] S. G. Stepanian, I. D. Reva, E. D. Radchenko, and L. Adamowicz, J. Phys. Chem. A **105**, 10664 (2001).

[21] G. Botti, M. Ceotto, and R. Conte, J. Chem. Phys. **155**, 234102 (2021).

[22] W. H. Miller and T. F. George, J. Chem. Phys. **56**, 5637 (1972).

[23] W. H. Miller, Proc. Natl. Acad. Sci. USA **102**, 6660 (2005).

[24] E. J. Heller, Acc. Chem. Res. **14**, 368 (1981).

[25] M. F. Herman and E. Kluk, Chem. Phys. **91**, 27 (1984).

[26] K. G. Kay, J. Chem. Phys. **101**, 2250 (1994).

[27] D. V. Shalashilin and M. S. Child, J. Chem. Phys. **115**, 5367 (2001).

[28] D. V. Shalashilin and M. S. Child, Chem. Phys. **304**, 103 (2004).

[29] F. Grossmann, J. Chem. Phys. **125** (2006).

[30] E. Pollak, "The Semiclassical Initial Value Series Representation of the Quantum Propagator," in *Quantum Dynamics of Complex Molecular Systems* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2007) pp. 259–271.

[31] M. Wehrle, M. Sulc, and J. Vanicek, J. Chem. Phys. **140**, 244114 (2014).

[32] M. Buchholz, F. Grossmann, and M. Ceotto, J. Chem. Phys. **148**, 114107 (2018).

[33] G. Bertaina, G. Di Liberto, and M. Ceotto, J. Chem. Phys. **151**, 114307 (2019).

[34] C. Aieta, M. Micciarelli, G. Bertaina, and M. Ceotto, Nat. Comm. **11**, 4384 (2020).

[35] A. Rognoni, R. Conte, and M. Ceotto, Chem. Sci. **12**, 2060 (2021).

[36] M. Buchholz, F. Grossmann, and M. Ceotto, J. Chem. Phys. **144**, 094102 (2016).

[37] M. Buchholz, F. Grossmann, and M. Ceotto, J. Chem. Phys. **147**, 164110 (2017).

[38] M. Buchholz, E. Fallacara, F. Gottwald, M. Ceotto, F. Grossmann, and S. D. Ivanov, Chem. Phys. **515**, 231 (2018).

[39] T. Begusic, J. Roulet, and J. Vanicek, J. Chem. Phys. **149**, 244115 (2018).

[40] M. Micciarelli, R. Conte, J. Suarez, and M. Ceotto, J. Chem. Phys. **149**, 064115 (2018).

[41] R. Conte, F. Gabas, G. Botti, Y. Zhuang, and M. Ceotto, J. Chem. Phys. **150**, 244118 (2019).

[42] T. Begusic, M. Cordova, and J. Vanicek, J. Chem. Phys. **150**, 154117 (2019).

[43] M. Micciarelli, F. Gabas, R. Conte, and M. Ceotto, J. Chem. Phys. **150**, 184113 (2019).

[44] T. Begušić and J. Vaníček, J. Chem. Phys. **153**, 184110 (2020).

[45] M. Cazzaniga, M. Micciarelli, F. Moriggi, A. Mahmoud, F. Gabas, and M. Ceotto, J. Chem. Phys. **152**, 104104 (2020).

[46] R. Conte, G. Botti, and M. Ceotto, Vib. Spectrosc. **106**, 103015 (2020).

[47] M. Gandolfi, A. Rognoni, C. Aieta, R. Conte, and M. Ceotto, J. Chem. Phys. **153**, 204104 (2020).

[48] C. Aieta, G. Bertaina, M. Micciarelli, and M. Ceotto, J. Chem. Phys. **153**, 214117 (2020).

[49] A. Rognoni, R. Conte, and M. Ceotto, J. Chem. Phys. **154**, 094106 (2021).

[50] M. Gandolfi and M. Ceotto, J. Chem. Theory Comput. **17**, 6733 (2021).

[51] M. Ceotto, G. Di Liberto, and R. Conte, Phys. Rev. Lett. **119**, 010401 (2017).

[52] A. L. Kaledin and W. H. Miller, J. Chem. Phys. **118**, 7174 (2003).

[53] A. L. Kaledin and W. H. Miller, J. Chem. Phys. **119**, 3078 (2003).

[54] M. Ceotto, S. Atahan, G. F. Tantardini, and A. Aspuru-Guzik, J. Chem. Phys. **130**, 234113 (2009).

[55] M. Ceotto, S. Valleau, G. F. Tantardini, and A. Aspuru-Guzik, J. Chem. Phys. **134**, 234103 (2011).

[56] F. Gabas, G. Di Liberto, R. Conte, and M. Ceotto, Chem. Sci. **9**, 7894 (2018).

[57] F. Gabas, R. Conte, and M. Ceotto, J. Chem. Theory Comput. **16**, 3476 (2020).

[58] F. Gabas, R. Conte, and M. Ceotto, J. Phys. Chem. Lett. **13**, 1350 (2022).

[59] R. Conte, L. Parma, C. Aieta, A. Rognoni, and M. Ceotto, J. Chem. Phys. **151**, 214107 (2019).

[60] E. A. Solovev, Zh. Eksp. Teor. Fiz. **75**, 1261 (1978).

[61] Q. Sun, J. M. Bowman, and B. Gazdy, J. Chem. Phys. **89**, 3124 (1988).

[62] S. Saini, J. Zakrzewski, and H. S. Taylor, Phys. Rev. A **38**, 3900 (1988).

[63] J. Huang, J. J. Valentini, and J. T. Muckerman, J. Chem. Phys. **102**, 5695 (1995).

[64] C. Qu and J. M. Bowman, J. Phys. Chem. A **120**, 4988 (2016).

[65] T. Nagy and G. Lendvay, J. Phys. Chem. Lett. **8**, 4621 (2017).

[66] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, The Journal of Chemical Physics **132**, 154104 (2010).

[67] R. Conte, C. Qu, P. L. Houston, and J. M. Bowman, J. Chem. Theory Comput. **16**, 3264 (2020).

[68] L. D. Landau and E. M. Lifshitz, *Mechanics* (Elsevier, 1982).

[69] M. Valiev, E. Bylaska, N. Govind, K. Kowalski, T. Straatsma, H. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. Windus, and W. de Jong, Comput. Phys. Commun. **181**, 1477 (2010).

[70] F. Gabas, R. Conte, and M. Ceotto, J. Chem. Theory Comput. **13**, 2378 (2017).

[71] M. Ceotto, S. Atahan, S. Shim, G. F. Tantardini, and A. Aspuru-Guzik, Phys. Chem. Chem. Phys. **11**, 3861 (2009).

[72] I. Reva, S. Stepanian, A. Plokhotnichenko, E. Radchenko, G. Sheina, and Y. Blagoi, J. Mol. Struct. **318**, 1 (1994).

[73] A. Lesarri, S. Mata, E. J. Cocinero, S. Blanco, J. C. Lûpez, and J. L. Alonso, Angew. Chem. Int. Ed. **41**, 4673 (2002).