This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal’s standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
The generalization of the PICVib approach [M. V. P. dos Santos et al., J. Comput. Chem. 2013, 34, 611] for calculating infrared intensities is shown to be successful and to preserve all interesting features of the procedure such as easiness of implementation and parallelization, flexibility, treatment of large systems and at high theoretical levels. It was tested and validated for very diverse molecular systems: XH3 (D2h), YH4 (D4h), conformers of RDX, S82 and E2 reaction product complexes, [W(dppe)2(NNC5H10)] complex, carbon nanotubes, and hydrogen-bonded complexes (H2O···HOH, MeHO···HOH, MeOH···OH2, MeOH···OHMe) including guanine- cytosine pair. The PICVib has an excellent overall performance for calculating infrared intensities of localized normal modes and even mixed vibrations, whereas care must be taken for vibrations involving intermolecular interactions. DFT functionals are still the best combination with high level ab initio methods such as CCSD and CCSD(T).

1. Introduction

Despite being one of the oldest methods of molecular spectroscopy, infrared (IR) is still one of the most important techniques for quantification and characterization of molecules and materials, mainly because of its appealing features and wide range of samples and analyses.1-7 For instance, vibrational frequencies of functional chemical groups are well characterized and transferable, thus resulting in prompt IR spectral assignments and interpretations. Noteworthy that in many instances only the values of the vibrational frequencies are not sufficient to unambiguously assign the spectrum, so the intensities of the IR transitions become indispensable. However, IR intensities are more difficult to estimate and require more elaborate models based on atomic charges and charge fluxes parameters.8 In fact, atomic charges and charge fluxes deduced from infrared intensities have been successful in describing intramolecular effects9-10 such as back donation and inductive effects as well as intermolecular ones.11-12

Currently, three models have been employed to estimate, interpret and even predict IR intensities, namely, the charge-charge flux-overlap (CCFO) model13 and its modification (CCFOM)14-16 based on the decomposition of the atomic polar tensor;17 the equilibrium charges charge-fluxes (ECCF) model18-19 based on electro-optical parameters that has also shown to be expressed in terms of the atomic polar tensor;20-21 and the atomic charge-charge-flux-dipole flux (CCFDF) model22-23 which includes atomic charges, charge fluxes and dipole fluxes contributions.16 These models are important in the interpretation of IR intensities in terms of atomic charges and their fluxes, however, they are limited to planar and linear molecules. Regarding quantum chemistry methods, the IR intensities can be calculated from the derivative of the dipole moment components with respect to the nuclear displacements.24-26

Usually, these calculations require better basis sets, including diffuse functions, and inclusion of electronic correlation when compared to the calculations of vibrational frequencies.25-27
step is thus the generalization of PICVib to perform IR intensity calculations with the same accuracy and computational demand. A proper choice of the calculation levels provides accurate IR intensities that can also be used to select vibrational transitions of interest, usually the most intense ones, similarly to the intensity-tracking scheme. For high-level ab initio methods, the computational demand increases steeply, so the PICVib approach becomes an important tool to provide accurate vibrational frequencies at the cost of single point energy calculations. Validation of the IR intensities calculated within the PICVib approach was performed for similar molecular systems used in the development of PICVib, namely, XH\(_2\) (D\(_{2h}\)) with X = B, Al, Ga, N, P, As, O, S and Se, YH\(_2\) (D\(_{2h}\)) with Y = C, Si and Ge, conformers of RDX, S\(_8\)C\(_2\) and E2 reaction product complexes, [W(dppe)\(_2\)(NCC\(_6\)H\(_{10}\))] complex, carbon nanotubes, and hydrogen-bonded complexes including guanine-cytosine pair. The performance of the PICVib approach was also excellent for IR intensities and its generalization to Raman intensities, anharmonicities, and solid state systems is currently under development.

2. Computational methods

2.1 The PICVib (Procedure for Investigating Categories of Vibrations) Approach

The PICVib approach explores the weak dependence of the normal modes on the level of theory. Briefly, from the equilibrium structure obtained at a desired high level of theory, the PICVib approach performs a simpler and feasible calculation of all normal modes and uses these normal modes to generate four displaced structures (±0.001 and ±0.01 Å) along a selected normal mode, \(Q_a\). The energies of each displaced structure, \(E(Q_a)\), are calculated at the desired level of theory and used to determine the force constant from a second-order polynomial fit to the energy,

\[
E(Q_a) = h_0^{(a)} + h_1^{(a)}Q_a + h_2^{(a)}Q_a^2
\]

(1)

where \(h_0^{(a)}\), \(h_1^{(a)}\) and \(h_2^{(a)}\) are the fitted polynomial coefficients for the selected normal mode and provide the corresponding force constant \(k_a\),

\[
k_a = 2h_2^{(a)}E_a \text{Å}^{-2} = 8.71948868 \times h_2^{(a)} \text{mdyn Å}^{-1}
\]

(2)

and the associated harmonic frequency, \(\nu_a\),

\[
\nu_a \text{cm}^{-1} = \frac{1}{\pi} \sqrt{\frac{h_2^{(a)}}{2m_a}} = 3.84698425 \times 10^{-1} \sqrt{\frac{h_2^{(a)}}{\text{mdyn Å}^{-1}}} m_a \text{amu}^{-1}
\]

(3)

with \(m_a\) being the reduced mass (in amu) of the selected normal mode \(a\) and \(h_2^{(a)}\) in mdyn Å\(^{-1}\) to yield \(\nu_a\) in cm\(^{-1}\).

2.2 The PICVib Approach for Infrared Intensities

Within the mechanic and electric harmonic approximations, first order time-dependent perturbation theory, dipole approximation, and thermal averaging, the infrared absorption coefficient, \(A_{\nu_a}\), of a fundamental band associated with the normal mode \(a\) can be expressed as,

\[
A_{\nu_a} = \frac{N_A\pi}{3c^2} \left( \frac{\partial \mu}{\partial Q_a} \right)^2 = \frac{N_A\pi}{3c^2} \left( \frac{\partial \mu_x}{\partial Q_a} \right)^2 + \left( \frac{\partial \mu_y}{\partial Q_a} \right)^2 + \left( \frac{\partial \mu_z}{\partial Q_a} \right)^2 \]

(4)

where \(N_A\) is the Avogadro number, \(c\) is speed of light, \(\mu_x\), \(\mu_y\), and \(\mu_z\) are the Cartesian components of the dipole moment. The infrared intensity of a particular transition can be obtained from the corresponding absorption coefficient and the experimental setup, namely, the intensity of the incident light, the molar concentration, and the optical path through the sample cell. However, in the electronic structure literature and many other sources, the absorption coefficient is commonly referred as infrared intensity, so we shall also refer the quantity calculated in eq. (4) as the intensity, \(I_{\nu_a}\), that is, \(I_{\nu_a} = A_{\nu_a}\).

Within the PICVib formalism, the partial dipole derivatives for the IR intensity can be determined by the following linear expansion,

\[
\mu_i(Q_a) = a_{i0}^{(a)} + a_{i1}^{(a)}Q_a, \text{ with } i = x, y, z
\]

(5)

where the coefficients \(a_{i0}^{(a)}\) and \(a_{i1}^{(a)}\) are obtained by fitting a set of five-point values. Notice that these points are the same used to calculate the energies and provide the force constant, eqs. (1) and (2). Thus, once the energy and dipole moment components are obtained at the same (single point) calculation, there is no additional cost to perform infrared intensity calculation. From the fitted coefficients, the partial dipole derivatives can be approximated as,

\[
\frac{\partial \mu_i}{\partial Q_a} \approx \frac{a_{i1}^{(a)}}{\sqrt{m_a}} \text{ D Å}^{-1/2}
\]

(6)

and the IR intensity \(I_{\nu_a}\) (in km mol\(^{-1}\)) can be rewritten as,

\[
I_{\nu_a} \text{ (km mol}^{-1}\) = \frac{422547}{m_a \text{amu}} \sum_{\nu_{ij}} \left| a_{ij}^{(a)} / (\text{D Å}^{-1}) \right|^2
\]

(7)

for \(a_{ij}^{(a)}\) in D Å\(^{-1}\) and the reduced mass \(m_a\) in amu. Analogously to the PICVib approach of vibrational frequencies, the IR intensity calculation requires only the reduced mass and the linear fitted coefficients \(a_{ij}^{(a)}\) for a selected normal mode \(a\). For the IR intensities, this procedure requires all three partial derivatives of the dipole components \(a_{x1}^{(a)}\), \(a_{y1}^{(a)}\), \(a_{z1}^{(a)}\), which means three linear fittings, eq. (5), instead of one quadratic fitting, eq. (2), for the force constant calculation. One the other hand, linear fittings are expected to be more accurate than quadratic ones. However, because the IR intensity is estimated by the fitted \(a_{ij}^{(a)}\) terms squared, eq. (7),
the errors may be enhanced, resulting in worse estimates for the intensities than for the vibrational frequencies.

The PICVib approach\textsuperscript{10} use the notation method1:method2, where "method1" denotes the high-level quantum chemical method used to obtain the equilibrium molecular structure and to perform the single-point (energy and dipole moment) calculations for the displaced geometries, while "method2" is the least demanding method employed in the vibrational frequency calculation at the structure provided by "method1".

2.3 Molecular Systems and Computational Procedures

The systems used in the validation of the PICVib approach to vibrational frequencies\textsuperscript{10} were diverse and relevant, thus similar structures were employed in the present evaluation, which are depicted in Figure 1. This set includes systems with few atoms, namely, planar \textit{XH}_3 (D\textsubscript{3h} symmetry) with \textit{X} = B, Al, Ga, N, P, As, O, S, and Se and planar \textit{YH}_4 (D\textsubscript{4h} symmetry) with \textit{Y} = C, Si, and Ge closed-shell molecules (Figures 1a and 1b). The \textit{XH}_3 systems can be neutral or cationic according to the central atom, whereas the \textit{YH}_4 systems are neutral. The selected normal modes used for infrared intensities calculations involve the out-of-plane motion of the central atom. The equilibrium structures were obtained with the B3LYP,\textsuperscript{51} MP2,\textsuperscript{52} CCSD and CCSD(T) methods\textsuperscript{53} employing 6-311+G(2d,p) basis sets.\textsuperscript{74} The Cartesian coordinates of the normal modes were determined at the following theoretical levels: AM1/PM6 (PM6\textsuperscript{55} was employed when the AM1 method\textsuperscript{56} had no parameters for the central atom), B3LYP, MP2, and CCSD also with 6-311+G(2d,p) basis sets. The PICVib calculated IR intensities were compared to those obtained by analytical MP2/6-31G(d) results.

Single-wall carbon nanotubes (SWNTs) with different sizes were also studied by the PICVib-intensity approach. More specifically, we investigated some relevant IR intensities of isolated (8,0)\times3 and (11,0)\times3 SWNTs (Figures 11 and 1m). The PICVib procedure employed the density functional based tight binding (DFTB)\textsuperscript{63} and B3LYP/STO-3G methods at the lower-level and the B3LYP/6-31G(d) higher-level method within the Gaussian 09 revision C.01 program.

We have applied the PICVib approach in the study of selected IR intensities of the [W(dppe)$_2$($^{14}$N$^{15}$NC$_3$H$_{10}$)] and [W(dppe)$_2$($^{14}$N$^{15}$NC$_3$H$_{10}$)] isotopomer complexes, in Figure 1i. The structure was obtained at the B3LYP/LANL2DZ level and the two-layer ONIOM(B3LYP/LANL2DZ:PM6) method\textsuperscript{64} provided the normal modes using [Mo(NNC$_3$H$_{10}$)(PC$_3$H$_2$)$_2$] as the model system for the low-layer.\textsuperscript{30,65} The PICVib approach employed this structure and calculated the dipole moments of the displaced structures at the B3LYP/LANL2DZ level. All calculations were performed with the Gaussian 09 revision C.01 program.

Vibrational frequencies and IR intensities of hydrogen-bonded (H-bonded) systems such as water-water (HOH···OH$_2$), water-methanol (MeOH···OH$_2$ and HOH···OHMe), and methanol-methanol (MeOH···OHMe), depicted in Figures 1c-1f, were studied at the CCSD(T) level and MP2 levels of theory with 6-311+G(d,p) basis sets. The vibrational frequencies and IR intensities of the O–H bonds and OH···O hydrogen-bonds were estimated by the PICVib procedure using several DFT functionals (B3LYP, PBE1PBE,\textsuperscript{66} M06-2X,\textsuperscript{67} and wB97X-D\textsuperscript{68}) to calculate the normal mode coordinates. The H-bonded base-pair G-C (guanine-cytosine), Figure 1g, was investigated using three semiempirical methods (AM1, PM6 and PDDG\textsuperscript{69}), DFTB and wB97X-D functional in the low-level calculation and MP2/6-31++G(d,p) as the high-level method. The CCSD(T) calculations were performed with the AICESII and CFOUR programs and the remaining ones with the Gaussian 09 revision C.01 program.

For all these applications, the displacements along the selected normal modes were 0.0, ±0.001 and ±0.01 Å. The Gaussian and NWChem programs were slightly modified to increase the number of decimal figures printed for the Cartesian coordinates, energies and dipole moments to provide adequate numerical accuracy for the results. Simple computer programs based on Linux/Unix shell script as well as C/C++ were written to perform the PICVib analysis. These programs read the normal mode coordinates and the displacement values and provide the Cartesian coordinates of the structures along the normal mode for energy and dipole moment calculations in a format consistent with the chosen quantum chemistry program. These programs are available upon request by email to picvib.project@gmail.com.
Fig. 1. Molecular systems employed in the validation of the PICVib approach for IR intensities: (a) planar XH$_3$ (X = B, Al, Ga, N, P, As, O, S, and Se), (b) planar YH$_4$ (Y = C, Si, and Ge), (c) water–water (HOH···OH$_2$), water–methanol (d) MeOH···OH$_2$ and (e) HOH···OHMe, (f) methanol–methanol (MeOH···OHMe), (g) G–C (guanine–cytosine) base-pair, (h) RDX (1,3,5-trinitro-1,3,5-triazacyclohexane), (i) [W(dppe)$_2$(NNC$_5$H$_{10}$)] complex, (j) MeCOOEt···BDM···Cl$^-$···H$_2$C=CH$_2$, (k) MeCOOH···BDM···Cl$^-$···H$_2$C=CH$_2$, (l) single-wall carbon nanotube (8,0)$\times$3, and (m) single-wall carbon nanotube (11,0)$\times$3.
3. Results and discussions

As in the case of vibrational frequency calculations, the PICVib procedure has a very good overall performance to describe the IR intensities of several kinds of normal modes. Initially, it was validated for the out-of-plane modes $A'_2$ and $A'_1$ of $XH_3 (D_{3h})$ and $YH_4 (D_{4h})$ planar systems depicted in Figures 1a and 1b, respectively. These molecules and normal modes (positive or negative force constants), with IR intensities ranging from 82 to 750 km mol$^{-1}$, were selected because they have unusually distorted electronic densities and thus providing stringent tests on the quality of the wavefunction and on the accuracy of the PICVib algorithm. In addition, because these are small systems their vibrational frequencies and IR intensities can be calculated analytically at higher theoretical levels such as CCSD(T). The PICVib approach provides IR intensities with an average accuracy of ca. 100% (Table S4 in the Supplementary Information) with respect to the analytical calculations (Table S2). In fact, the errors associated with the AM1/PM6 semiempirical methods are the same as those provided by DFT and MP2 methods. This is an interesting result because semiempirical methods have very small computational demand compared to DFT or ab initio methods.

In fact, the largest error (0.1%) associated to the IR intensity calculation is for planar $CH_4$ species at CCSD(T)/6-311+G(2d,p) level. The $GeH_4$ planar system was excluded from this error analysis because it has unstable wavefunction at several theoretical levels. The PICVib procedure is very accurate for predicting IR intensities of this set of compounds and an excellent probe to identify wavefunction instabilities.

RDX (1,3,5-trinitro-1,3,5-triazaacyclohexane), illustrated in Figure 1h, is an important explosive whose vibrational spectrum is typically employed for characterization. It has been investigated quite extensively and its main conformers have been studied at the MP2 level.66 Because it has less localized normal modes than the $XH_3$ and $YH_4$ planar systems, it is a more stringent test on the performance of the PICVib procedure. The normal modes involving the N–O stretches were selected because they have a wide range of IR intensities, ca. 0-314 km mol$^{-1}$, as presented in Table 1. These normal modes (see Figure S1) can involve up to three N–O asymmetrical stretches of the NO$_2$ groups at distant regions of the molecule. Because the dipole moment is very sensitive to the bond stretch between heteroatoms, predictions of these IR intensities are not a trivial task.

Table S5 presents the vibrational frequencies of the three normal modes for the conformers of RDX and shows that the normal mode $v_{(N-O)}(1)$ of AAE conformer presents the largest error. This is due to the RMSD (Root Mean Square Deviation) between the Cartesian coordinates of the MP2 and wB97XD normal modes being the largest one amongst this whole set of vibrations.

<table>
<thead>
<tr>
<th>Conformer</th>
<th>PICVib MP2:wB97XD</th>
<th>Analytical MP2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$v_{(N-O)}$ (1)</td>
<td>$v_{(N-O)}$ (2)</td>
</tr>
<tr>
<td>AAE</td>
<td>55 (+8.3)</td>
<td>203 (+11.5)</td>
</tr>
<tr>
<td>Twist</td>
<td>74 (-14.9)</td>
<td>201 (+5.8)</td>
</tr>
<tr>
<td>EEA</td>
<td>59 (-28.9)</td>
<td>220 (+16.4)</td>
</tr>
<tr>
<td>EEE</td>
<td>0 (+5.9)</td>
<td>307 (+5.9)</td>
</tr>
</tbody>
</table>

Table 1. The IR intensities (km mol$^{-1}$) of three normal modes for each RDX conformer. The value in parenthesis is the relative difference (in %) between the PICVib and the analytical IR intensities.

However, in the case of IR intensities (Table 1) the largest errors are found for the EEA conformer, probably because the IR intensities depend on vector property (dipole moment components), and not only the RMSD values are important, but also the directions of these RMSDs. Thus, unbalanced discrepancies in the normal modes can generate displaced structures that lead to slightly different values of the dipole moment components, which are enhanced by the quadratic dependence of the IR intensities on the dipole derivatives. Despite these considerations, the PICVib procedure provided the correct trends for all these IR intensities even for a floppy system as RDX.

Because the PICVib procedure investigates selected modes individually, it has the feature of producing simplified versions of vibrational spectra, where only the normal modes of interest, for instance, the most intense transitions, are calculated and displayed. In Figure 2, the complete IR spectrum (with all 99 vibrational transitions) of the intermolecular complex product (MeCOOEt···Cl···BDM) from the S$_{1}$2 reaction (Figures 1j and S2), calculated at the MP2/6-31G(d) level, is compared to only six normal modes calculated with the PICVib approach (see Table S6 in the Supplementary Information) at MP2:PBE1PBE/6-31G(d) level.

It can be seen that these six vibrational modes provide a clear picture of the IR spectrum with the most prominent features: C–O–H bending ($v_{16}$ and $v_{32}$) and O–H stretching of BDM ($v_{99}$ and $v_{90}$), C–O stretching ($v_{98}$) and C=O stretching ($v_{93}$) of ethyl acetate, with the $v_{16}$ mode being related to the interaction between Cl$^-$ and H–O of BDM.

Similarly, Figure 3 presents the complete IR spectrum (with all 99 vibrational transitions) of the intermolecular complex product (MeCOOH···H$_2$C=CH$_2$···Cl$^-$···BDM) from the E2 reaction (Figures 1k and S3) and compare it with the simplified ten normal modes calculated by the PICVib approach (Table S7). There are noticeable differences with respect to the product from the S$_{1}$2 pathway, such as an extra O–H stretching band ($v_{95}$) related to acetic acid and peaks assigned to C–H bending ($v_{44}$ and $v_{48}$) and stretching ($v_{87}$) of ethene.
The differences between these spectra could be used to characterize this reaction selectivity in gas-phase and thus provide information about the role of the BDM in these competitive reaction pathways. Regarding the accuracy of PICVib for computing IR intensities, its performance is very good with an unsigned average error smaller than 5.4% for intensities ranging from 10 to 855 km mol\(^{-1}\). As expected, the largest relative errors are related to the small intensity values (10-56 km mol\(^{-1}\)) for which absolute errors of 2-9 km mol\(^{-1}\) can lead to relative errors up to 19%.

Single-wall carbon nanotubes (SWNTs) are another challenging system for ascertaining the performance of PICVib mainly because of their computational demand (nanoscaled systems) and widespread potential applications. Despite some inherent difficulties,\(^{70}\) two normal modes with \(E_{1u}\) and \(A_{2u}\) symmetries around 1590 and 850 cm\(^{-1}\) were detected with considerable IR intensities.\(^{52}\) Theoretical studies have shown a strong dependence between diameter/chirality of the SWNTs and the frequencies of the IR-active modes.\(^{52}\) Thus, we employed the PICVib procedure to investigate two vibrational modes of isolated (8,0)\(\times\)3 and (11,0)\(\times\)3 SWNTs (Figures 1l and 1m). These normal modes for the (8,0)\(\times\)3 nanotube are depicted in Figure 4.

The calculated vibrational frequencies and IR intensities of the modes \(\nu_{C-H}(E_{1u})\) and \(\nu_{C-C}(A_{2u})\) are presented in Table 2, for the PICVib procedure at two levels: B3LYP/6-31G(d): DFTB and B3LYP/6-31G(d): B3LYP/STO-3G. The results from the analytical calculations at the B3LYP/6-31G(d) level are presented for comparison. Notice that the calculated IR intensities for the \(\nu_{C-H}\) normal mode of the (8,0)\(\times\)3 nanotube are ca. 3 times higher than those for the (11,0)\(\times\)3 one. Whereas the \(\nu_{C-C}\) normal mode is approximately 1.5 times more intense for the larger SWNT. In addition, for the smaller (8,0)\(\times\)3 nanotube, the intensity of the \(\nu_{C-H}\) vibration is larger than the \(\nu_{C-C}\) one, whereas this trend is reversed in the larger (11,0)\(\times\)3 SWNT.
Table 2. Vibrational frequencies, \( \nu \) (cm\(^{-1} \)), and IR intensities, \( I \) (km mol\(^{-1} \)), of two normal modes of the pure (8,0)×3 and (11,0)×3 SWNTs. The PICVib approach at the B3LYP/6-31G(d):DFTB and B3LYP/6-31G(d):B3LYP/STO-3G levels was employed and compared to the analytical B3LYP/6-31G(d) results. The number of functions in the 6-31G(d) basis sets is 1472 and 2024 for (8,0)×3 and (11,0)×3 SWNTs, respectively. Ratio means the ratio between the IR intensities of (8,0)×3 and (11,0)×3 SWNTs.

<table>
<thead>
<tr>
<th>Mode</th>
<th>( (8,0) \times 3 )</th>
<th>( (11,0) \times 3 )</th>
<th>( (8,0) \times 3 )</th>
<th>( (11,0) \times 3 )</th>
<th>( (8,0) \times 3 )</th>
<th>( (11,0) \times 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \nu )</td>
<td>( \nu )</td>
<td>( \nu )</td>
<td>( \nu )</td>
<td>( \nu )</td>
<td>( \nu )</td>
<td>( \nu )</td>
</tr>
<tr>
<td>C_H</td>
<td>786</td>
<td>1219</td>
<td>862</td>
<td>458</td>
<td>847</td>
<td>1497</td>
</tr>
<tr>
<td>Ratio</td>
<td>2.66</td>
<td>3.11</td>
<td>3.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_C</td>
<td>1503</td>
<td>963</td>
<td>1519</td>
<td>574</td>
<td>1504</td>
<td>777</td>
</tr>
<tr>
<td>Ratio</td>
<td>1.68</td>
<td>1.52</td>
<td>1.28</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These results are consistent with calculated and observed trends of the properties of nanotubes.\(^{70-73}\) The relationship between the diameter of the nanotube and the vibrational frequency, particularly, for the radial breathing mode is well established.\(^{73}\) Notice, however, that the IR intensities also can be very sensitive to the diameter of the carbon nanotube, so they may be an useful probe for characterizing SWNTs. Noteworthy that preserving the B3LYP functional for both levels (high and low methods) leads to more accurate results than the usual partitioning scheme that uses different methods. In fact, the use of minimum basis sets (STO-3G) allows performing this calculation with a small increment on the computational cost.

An armchair (5,5)×3 SWNT and its functionalized structures (with HF and \( F_2 \)) were the first large systems studied by the mode-tracking algorithm for the Raman active breathing mode\(^{49}\). The initial guess was provided by PM3 calculations of all normal modes and the number of iterations of the Davidson procedure was related to the goodness that the PM3 results (geometry and vibrational frequency) matched the DFT ones. This is quite similar to the performance of the PICVib approach that also depends upon the closeness of the low-level and high-level results. Because the selected mode was not active in the IR spectrum, its intensity was not calculated and a comparison (between these methods) was not possible; however, the ongoing generalization of the PICVib approach to Raman intensities (or activities) shall allow for a direct comparison between these procedures.

Another system explored by the PICVib approach is the [W(dppe)_2(\text{NNC}_5\text{H}_{10})] (dppe = 1,2-bis(diphenylphosphino)ethane) - \( \text{B}^\text{W} \) complex (see Figure 1i) and its \(^{15}\text{N} \) isotopomer.\(^{74}\) This molecule was chosen because in our previous investigation,\(^{30}\) the normal mode analysis revealed an extensive mixing between the N–N and W–N vibrations of the metal–N–N core with the vibrations of the piperidine ring. Even with this mixing, the PICVib procedure estimated these vibrational frequencies in very good agreement with experimental and QCC-NCA values.\(^{74}\) Thus, in order to properly reproduce this relevant part of the IR spectrum, it is important to validate the PICVib approach for calculating the intensities. The comparisons between the full analytical IR spectra of the \(^{14}\text{N} \)-\( \text{B}^\text{W} \) and \(^{15}\text{N} \)-\( \text{B}^\text{W} \) isotopomer complexes with the PICVib approach for selected vibrations are presented in Figures 5 and 6, respectively. These selected normal modes involve mainly the N–N moiety in the NNC\(_5\text{H}_{10}\) ligand (see Figure S4).

The PICVib approach does provide reasonable results for the IR intensities (see Table S8) showing the correct isotopic trends in Figures 5 and 6.
Notice that the intensities of modes VIIIa and VIIIb are calculated with unusually larger errors because they are delocalized over the phenyl rings that are part of the ONIOM low-layer, which is treated by the PM3 semiempirical method. The PM3 method has not been properly validated for tungsten type-complexes and is likely to be a source of large errors. This can be considered a critical test case of the PICVib procedure, as the Schrock’s dinitrogen molybdenum complex was for the intensity tracking approach. Indeed, both procedures provided good results for the IR intensities of relevant normal modes at a small fraction of the computational demand of the full analytical Hessian diagonalization.

Infrared intensities can be a sensitive tool for probing H-bonded systems and the following intermolecular complexes were used to validate the PICVib approach: water-water (Figure 1c), water-methanol (Figures 1d e 1e), methanol-methanol (Figure 1f), and C-G (cytosine-guanine) base-pair (Figure 1g). The first three systems were selected because they are prototypical H-bonded solvents and their small sizes allows for high-level analytical calculations. Whereas, the latter is one of the most relevant biomolecular system. The selected normal modes involve the intramolecular O–H bond and intermolecular OH···O interaction because they are potential candidates for probing hydrogen-bonding by vibrational spectroscopy. As observed previously, calculations of certain normal modes of intermolecular complexes are strongly dependent upon the density functional employed. Therefore, we performed a more detailed analysis using four DFT functionals: B3LYP, PBE1PBE, M06-2X and wB97XD.

Taking the MP2 calculations as the reference values (Table S9), the B3LYP functional presents the worst performance for predicting the vibrational frequencies the normal mode related to OH···O interaction in the water dimer and the MeHO···OH, while the other functionals had an excellent accuracy for calculating the frequencies of the whole set of normal modes.
Table 3. Infrared intensities (km mol$^{-1}$) of the vibrational modes involving the OH···O and O−H bonds in hydrogen-bonded complexes of water-water, water-methanol (MeOH) and methanol-methanol. PICVib calculations at the MP2:B3LYP, MP2:PBE1PBE, MP2:M06-2X and MP2:wB97XD levels are compared to the MP2 values. All calculations employed 6-311+G(d,p) basis sets.

<table>
<thead>
<tr>
<th>H-bond complex</th>
<th>MP2</th>
<th>MP2:B3LYP</th>
<th>MP2:PBE1PBE</th>
<th>MP2:M06-2X</th>
<th>MP2:wB97XD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OH···O</td>
<td>O−H</td>
<td>OH···O</td>
<td>O−H</td>
<td>OH···O</td>
</tr>
<tr>
<td>H$_2$O···HOH</td>
<td>31</td>
<td>274</td>
<td>14</td>
<td>292</td>
<td>3</td>
</tr>
<tr>
<td>MeHO···HOH</td>
<td>4</td>
<td>361</td>
<td>0</td>
<td>380</td>
<td>4</td>
</tr>
<tr>
<td>MeOH···OH$_2$</td>
<td>6</td>
<td>379</td>
<td>11</td>
<td>375</td>
<td>3</td>
</tr>
<tr>
<td>MeOH···OHMe</td>
<td>1</td>
<td>461</td>
<td>1</td>
<td>460</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4. Infrared intensities (km mol$^{-1}$) of the vibrational modes involving the OH···O and O−H bonds in hydrogen-bonded complexes of water-water, water-methanol (MeOH) and methanol-methanol. PICVib calculations at the CCSD(T):B3LYP, CCSD(T):PBE1PBE, CCSD(T):M06-2X and CCSD(T):wB97XD levels. All calculations employed 6-311+G(d,p) basis sets.

<table>
<thead>
<tr>
<th>H-bond complex</th>
<th>CCSD(T)</th>
<th>CCSD(T): B3LYP</th>
<th>CCSD(T): PBE1PBE</th>
<th>CCSD(T): M06-2X</th>
<th>CCSD(T): wB97XD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OH···O</td>
<td>O−H</td>
<td>OH···O</td>
<td>O−H</td>
<td>OH···O</td>
</tr>
<tr>
<td>H$_2$O···HOH</td>
<td>37</td>
<td>231</td>
<td>17</td>
<td>251</td>
<td>1</td>
</tr>
<tr>
<td>MeHO···HOH</td>
<td>6</td>
<td>315</td>
<td>7</td>
<td>335</td>
<td>8</td>
</tr>
<tr>
<td>MeOH···OH$_2$</td>
<td>8</td>
<td>332</td>
<td>6</td>
<td>332</td>
<td>2</td>
</tr>
<tr>
<td>MeOH···OHMe</td>
<td>1</td>
<td>416</td>
<td>3</td>
<td>416</td>
<td>2</td>
</tr>
</tbody>
</table>

The same trends and performances were obtained when compared to the CCSD(T) level of calculation (Table S10), where the PBE1PBE and wB97XD functionals showed a better performance than the M06-2X one.

Regarding the infrared intensities, Tables 3 and 4 present comparisons between the PICVib procedure and the analytical calculations at both MP2 and CCSD(T) levels. The PICVib combined the four DFT functionals with the ab initio methods, namely, M1:B3LYP, M1:PBE1PBE, M1:M06-2X and M1:wB97XD, with M1 = MP2 and CCSD(T).

Overall, for these DFT functionals, the PICVib procedure is very accurate for IR intensities of O−H normal modes (errors *ca.* 2.3%). The smallest deviations are those provided by the wB97XD functional (errors *ca.* 1.9%). Except for the H$_2$O···HOH complex, all functionals predict IR intensities of the O−H normal modes (errors *ca.* 1.2%) quite accurately. The intermolecular normal mode of water dimer is quite difficult to be calculated by the PICVib approach because of its strong dependence upon the DFT functional, with the wB97XD functional presenting the best performance.

Indeed, the behavior of the PICVib approach for calculating infrared intensities is distinct from the calculations of vibrational frequencies. This is probably due to the vector dependence of the IR intensities that can be sensitive to small variations of the intermolecular orientations, which barely affects the energy.

A similar analysis was performed with the CCSD(T) method as presented in Table 4. At this level of theory, the behavior of the DFT functionals was similar to that of MP2 method, with the PICVib procedure presenting slightly larger average errors (*ca.* 3.2%) for the IR intensities of O−H normal modes. The smallest error is associated to the M06-2X functional (errors *ca.* 2.4%), while the highest error is associated to the B3LYP functional (errors *ca.* 3.8%). The accuracy of the IR intensities of the intermolecular normal mode shows similar behavior to the MP2 results. However, unlike the previous analysis, the CCSD(T):B3LYP approach presents, on average, the best accuracy amongst the employed functionals.

The G-C (guanine-cytosine) DNA base pair was the large H-bonded system studied. For the vibrational frequency calculations, the PICVib approach employed the AM1 method for providing the selected vibrational modes. In the present work we expanded the range of methods employed to evaluate the infrared intensities: MP2:AM1, MP2:PM6, MP2:PDDG, MP2:DFTB, and MP2:wB97XD. The selected normal modes are the N−H bending, δ(N−H), and the N−H stretching, ν(N−H), vibrations involving the intermolecular hydrogen-bonding.
In summary, the performance of the PICVib approach for calculating infrared intensities is excellent for localized normal modes, is very good for mixed vibrations and a proper choice of the low-level method must be made for vibrations involving intermolecular interactions. The procedure for IR intensities is more sensitive to the quality of the low-level method and to the wavefunction instabilities effects than the vibrational frequencies ones.

Table 6 presents comparisons between the calculated intensities with the low-level method and with the PICVib approach for some selected systems and normal modes. Indeed, the results are significantly improved by using the PICVib approach that provides results very close to the analytical high-level ones. For the nanotubes, this improvement is even more relevant because it causes an inversion of the intensities, namely, the mode $ν_{C-H}$ is more intense than $ν_{C-C}$ when calculated with the low-level method, but becomes less intense at the PICVib level.

The computational demand, both in CPU time savings and the amount of RAM memory, of the PICVib procedure is a very small fraction of the complete analytical calculation, especially for the large systems such as nanotubes, G-C base pair, and $[W(dppe)$_2$(NHC)$_2$H$_{10}$]$ complex. The PICVib approach is thus a viable tool for exploring selected normal modes of large systems because of its accuracy to provide vibrational frequencies and IR intensities. Its easiness of implementation, parallelization, and flexibility allows using any quantum chemistry program or electronic structure method. It is also less expensive than the mode-tracking protocol because it employs only single-point calculations at the high-level, whereas the latter approach requires numerical differentiation of analytical gradients and the number of iterations and the subset of the Hessian matrix depends upon the initial guess for the selected normal mode. The accuracy of the PICVib approach for both vibrational frequencies and IR intensities is related to the closeness between the low-level and the high-level geometries and normal modes. However, there are internal consistency checks in the PICVib approach that allows for estimates of the errors and thus help guiding the selection of the low-level method. As mentioned earlier, the performance of the PICVib procedure has encouraged us to generalize this approach to calculate Raman intensities (or activities) and anharmonic corrections, as well as to extend it to crystalline systems. This generalization is under development and treatments of resonances using the PICVib approach are also being considered.

### Table 5. Infrared intensities (km mol$^{-1}$) of the vibrations involving the N-H bending and N-H stretching bonds for the guanine-cytosine (G-C) base pair complex. PICVib calculations at the MP2:AM1, MP2:PM6, MP2:PDDG, MP2:DFTB and MP2:wB97XD levels. The value in parenthesis is the percentage difference between the PICVib and analytical MP2 results. All calculations employed 6-31+G(d,p) basis sets.

<table>
<thead>
<tr>
<th>Mode</th>
<th>MP2:AM1</th>
<th>MP2:PM6</th>
<th>MP2:PDDG</th>
<th>MP2:DFTB</th>
<th>MP2:wB97XD</th>
<th>MP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$δ_{N-H}$</td>
<td>947 (+26%)</td>
<td>932 (+28%)</td>
<td>1432 (+11%)</td>
<td>987 (+23%)</td>
<td>1204 (+7%)</td>
<td>1288</td>
</tr>
<tr>
<td>$ν_{N-H}$</td>
<td>737 (+352%)</td>
<td>228 (+40%)</td>
<td>887 (+444%)</td>
<td>561 (+244%)</td>
<td>141 (+14%)</td>
<td>163</td>
</tr>
</tbody>
</table>

The PICVib calculated IR intensities are presented in Table 5 (and Table S11) and are validated by comparing to the analytical MP2 calculations.

For vibrational frequencies, the PICVib procedure using semiempirical methods describes both modes very accurately (error smaller than 2.5%). However, for IR intensities these methods presented a much poorer performance mainly for the intermolecular N–H stretching $ν_{N-H}$ normal mode. Even for the best semiempirical performance (MP2:PM6), we found errors of 40%. However, the wB97XD functional significantly improves the results with the PICVib approach despite of the delocalized nature of these normal modes and the vector behavior of the IR intensity contributions.

### Table 6. Comparisons between the low-level and PICVib calculated infrared intensities (km mol$^{-1}$).

<table>
<thead>
<tr>
<th>System/Mode</th>
<th>Low-level</th>
<th>PICVib</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAE/ $ν_{(N-O)}$ (1)</td>
<td>322</td>
<td>55</td>
</tr>
<tr>
<td>EEA/ $ν_{(N-O)}$ (2)</td>
<td>434</td>
<td>220</td>
</tr>
<tr>
<td>EEE/ $ν_{(N-O)}$ (3)</td>
<td>579</td>
<td>307</td>
</tr>
<tr>
<td>Nanotube (8,0)$×3$/ $ν_{C-H}$</td>
<td>744</td>
<td>1219</td>
</tr>
<tr>
<td>Nanotube (11,0)$×3$/ $ν_{C-C}$</td>
<td>826</td>
<td>574</td>
</tr>
<tr>
<td>G-C base pair/$ν_{(N-H)}$</td>
<td>1408$^a$</td>
<td>887</td>
</tr>
<tr>
<td>H$<em>2$O···HOH dimer/$ν</em>{O-H}$</td>
<td>327$^b$</td>
<td>243$^c$</td>
</tr>
<tr>
<td>S$<em>{n2}$ reaction/$ν</em>{59}$</td>
<td>305</td>
<td>344</td>
</tr>
<tr>
<td>E2 reaction/ $ν_{97}$</td>
<td>979</td>
<td>756</td>
</tr>
</tbody>
</table>

$^a$PDDG. $^b$M06-2X. $^c$CCSD(T):M06-2X.
Conclusions

The generalization of the PICVib approach for calculating infrared intensities was successful and involves only dipole moment calculations at the displaced structures used to compute the vibrational frequencies. Thus, there is no additional cost to obtain the IR intensities and the generalization preserves all the interesting features of the PICVib procedure, namely, easiness of implementation and parallelization, flexibility to use any computational quantum chemistry program or any combination of electronic structure methods. It can treat large systems at only a very small fraction of computational demand needed for complete analytical calculations at high-level quantum chemical methods. The only caveat is related to the decrease of relative accuracy of calculations of vibrations with very low intensities in the infrared region. However, these transitions in the IR spectrum are usually not useful, turning the PICVib procedure into a very appropriate and helpful tool, because it allows selecting the vibrations of interest based not only upon their frequencies, but also on their intensities.

Acknowledgements

The authors would like to thank the Brazilian System of High Performance Processing (SINAPAD), specially, CENAPAD-SP, LNCC-RJ, CENAPAD-PE, CESUP-RS, and CENAPAD-UFC. Authors M. V. P. S. and Y. G. P. would like to thank CAPES and FACEPE, respectively, for their scholarships. The Brazilian agencies CAPES, PRONEX-CNpq-FACEPE (APQ-0859-1.06/08), FINEP, and inct-INAMI are gratefully acknowledged for partial funding. The authors also thank Carlos Henrique B. Cruz (dQF-UFPE) for helping with the analytical frequency calculation of the C-G base pair.

Notes and references

* Departamento de Química Fundamental, Universidade Federal de Pernambuco, Cidade Universitária, Recife-PE, 50740-560, Brazil.
* Fax: 55 81 2126 8442; Tel: 55 81 2126 8459; E-mail: longo@ufpe.br

5 P. Larkin, Infrared and Raman Spectroscopy; Principles and Spectral Interpretation; Elsevier: New York, 2011.