JCTC Journal of Chemical Theory and Computation

Fully Integrated Approach to Compute Vibrationally Resolved Optical Spectra: From Small Molecules to Macrosystems

Vincenzo Barone,*[,†,‡] Julien Bloino,[†] Malgorzata Biczysko,[†] and Fabrizio Santoro[‡]

Dipartimento di Chimica "Paolo Corradini" and CR-INSTM Village, Universita di Napoli Federico II, Complesso Univ. Monte S. Angelo, via Cintia, 80126 Napoli, Italy, and Istituto per i Processi Chimico-Fisici, Area della Ricerca-CNR, via G. Moruzzi, 56124 Pisa, Italy

Received November 5, 2008

Abstract: A general and effective time-independent approach to compute vibrationally resolved electronic spectra from first principles has been integrated into the Gaussian computational chemistry package. This computational tool offers a simple and easy-to-use way to compute theoretical spectra starting from geometry optimization and frequency calculations for each electronic state. It is shown that in such a way it is straightforward to combine calculation of Franck−Condon integrals with any electronic computational model. The given examples illustrate the calculation of absorption and emission spectra, all in the UV−vis region, of various systems from small molecules to large ones, in gas as well as in condensed phases. The computational models applied range from fully quantum mechanical descriptions to discrete/continuum quantum mechanical/molecular mechanical/polarizable continuum models.

1. Introduction

Nowadays the characterization of complex biological systems or nanomaterials of direct technological and social interest relies more and more on computational approaches, e.g., for the evaluation and rationalization of structural, energetic, electronic, and dynamic features.[1−6] On the experimental side accurate information can be gained, in principle, by a number of spectroscopic techniques, magnetic as well as optical. Nevertheless, up to very recently direct comparisons between experimental and computed spectroscopic data have been rather scarce. Integrated approaches, capable of accurately simulating spectra but at the same time easily accessible to nonspecialists, are highly desirable. Such tools would allow for the exploitation of the recent and ongoing developments that are taking place in the field of computational spectroscopy[7−15] resulting in easy and, ideally, automatic vis-a-vis comparisons between experimental and theoretical results. In the present work, we introduce an approach to model quantitatively vibronic spectra, in line with such a demand. The integration among the different computational steps is particularly straightforward, since all calculations are performed within the same computational package: this enables the fully automatic computation and visualization of vibrationally resolved UV−visible spectra, a feature that is routinely available for other spectroscopic ranges (e.g., IR/Raman).

In the framework of the Franck−Condon (FC) principle,[16−18] time-independent ab initio approaches to simulate vibronic spectra are based on the computation of overlap integrals (known as FC integrals) between the vibrational wave functions of the electronic states involved in the transition. The computation of FC integrals requires a detailed knowledge of the multidimensional potential energy surfaces (PES) of both electronic states or, within the harmonic approximation, at least computation of equilibrium geometry structures and vibrational properties. Till recently, computations of vibronic spectra have been limited to small systems or approximated approaches, mainly a consequence of the difficulties in obtaining accurate descriptions of excited electronic states of polyatomic molecules. Recent develop-

* To whom correspondence should be addressed. E-mail: baronev@unina.it.

[†] Universita di Napoli Federico II.

[‡] Istituto per i Processi Chimico Fisici.

10.1021/ct8004744 CCC: $40.75
Published on Web 02/19/2009

ments in electronic structure theory for excited states within the time-dependent density functional theory (TD-DFT)[19,20] and resolution-of-the-identity approximation of coupled cluster theory (RI-CC2)[21] have paved the route toward the simulation of spectra for significantly larger systems. When treating such large systems, the inclusion of vibrational contributions becomes very challenging, since the number of vibrational states to be taken into account increases steeply with the dimension of the molecule and the spectral width. Nonetheless, most of the possible vibronic transitions do not contribute significantly to the spectrum. Therefore the availability of effective selection criteria to individuate a priori the most relevant vibronic transitions within the dense bath of possible final states, can make feasible the calculation of the spectrum line shape also for these systems. Several schemes have been proposed,[22−27] ranging from the simplest approach, based solely on the energy window of the spectrum,[22,23] up to rigorous prescreening techniques based on analytically derived sum rules.[27] To maximize the efficiency of calculations, it is necessary to adopt a fast and a priori selection scheme of general applicability for a variety of different systems that is able to correctly choose all the non-negligible transitions. To satisfy these conditions we derived a general and robust tool starting from a method recently introduced by some of us[25,26] in the frame of harmonic approximation, which has been proven to provide very accurate spectra of medium-to-large systems with a limited computational cost. This new computational tool has been integrated into the Gaussian package.[53] A particular care has been taken to avoid the introduction of any prefixed constraint neither for the number of allowed quanta in a single mode nor for the number of simultaneously excited vibrations and for the spectrum energy range. Moreover, an automated procedure controls the spectrum convergence to avoid unnecessary calculations, and it alerts the user when incorrect settings might cause missing of important transitions. Thanks to these characteristics the integrated approach here presented can be applied to simulate in a fully automatic manner any vibronic spectrum for systems where nonadiabatic couplings are negligible and harmonic approximation is reliable.

Our integrated approach starts with the computation of the equilibrium geometries, frequencies, and normal modes for both electronic states involved in the electronic transition. A dedicated procedure performs reorientation of molecular coordinates to minimize the effect of rotations between the displaced equilibrium structures. After superposition of molecular geometries, the requested vibrationally resolved absorption or emission spectrum is generated. While the presented computational tool is developed within the harmonic approximation, it can be extended to take into account anharmonic effects. As a first step in this direction a correction scheme to derive excited state's anharmonic frequencies from ground-state data has been implemented.

In general, the accuracy of a simulated spectrum depends on the quality of the description of both the initial and the final electronic states of the transition. This is obviously related to the proper choice of a well-suited computational model: a reliable description of equilibrium structures, harmonic frequencies, normal modes, and electronic transition energy is necessary. In this respect, the advantage of the integrated approach we present here is the possibility of combining various computational schemes to create user-defined and/or problem-tailored approaches as for example the refinement of the electronic transition energy at a higher level of theory than the one used for the computation of structures and frequencies. Moreover, if a stand-alone software is used to compute the spectrum, some errors and/or inaccuracies in computations may be introduced by data processing and limited printing precision in the outputs of the electronic structure codes. The integrated tool here presented solves these problems by directly working with machine precision data available inside the Gaussian computational package. On the other side, the possibility of using available visualization or data analyzing tools is strictly related to the ease of use and accessibility of the approach to the nonspecialists. The aforementioned advantages of an integrated approach are illustrated with a few examples showing simulations of absorption and emission spectra, all in the UV−vis region, in the gas and condensed phases.

The paper is organized as follows. Section 2 describes the general theoretical frame of computation of FC and Herzberg−Teller spectra along with the details of the current implementation. Computational models that have been applied to the determination of structures, vibrations, and energies to provide the needed information for spectra calculation are gathered in section 3. Section 4 reports the application of this integrated tool to the simulation of: the vibronically resolved $S_1 \leftarrow S_0$ absorption spectrum of anisole (section 4.1), the photodetachment spectrum of SF_6^- (section 4.2), the emission phosphorescence the spectrum of $T_1 \rightarrow S_0$ transition of chlorophyll c2 (section 4.3), the UV absorption spectrum of $n \rightarrow \pi^*$ transition of acrolein in gas phase and in aqueous solution (section 4.4), and the photoelectron spectra of the isolated adenine molecule and adenine adsorbed on the Si(100) surface (section 4.5).

2. Spectra Calculation

2.1. Theory. In this section we briefly summarize the general mathematical frame for the spectra computations. The absorption spectrum, defined as the rate of energy absorption by a single molecule per unit radiant energy density is given by the expression[8]

$$\sigma_{abs}(\omega) = \frac{4\pi^2\omega}{3}\sum_f |\langle\Psi_i|\mu|\Psi_f\rangle|^2\delta(E_f - E_i - \hbar\omega) \qquad (1)$$

On the same foot the emission spectrum in photon counting experiments, defined as the rate of photon emission due to a single molecule, is[8]

$$\sigma_{em}(\omega) = \frac{4\omega^3}{3\hbar c^3}\sum_f |\langle\Psi_i|\mu|\Psi_f\rangle|^2\delta(E_f - E_i + \hbar\omega) \qquad (2)$$

The stick spectra in eqs 1 and 2 are usually convoluted with a Lorentzian or a Gaussian to simulate homogeneous or inhomogeneous broadening, respectively.

The intensity of a line of absorption or emission depends therefore on the transition dipole moment integral $\langle\Psi_i|\mu|\Psi_f\rangle$,

where Ψ_i and Ψ_f are the molecular wave functions and μ is the electric dipole moment.

In the Born−Oppenheimer approximation, the wave function of each state can be expressed as a product of a nuclear ψ_n and an electronic ψ_e wave functions

$$\langle \Psi_i|\mu|\Psi_f\rangle = \langle \psi_n \psi_e|\mu|\psi'_e\psi'_n\rangle \quad (3)$$

The electric dipole moment can be separated into an electronic part μ_e and a nuclear part μ_n. By replacement of the electric dipole moment by these two components, the transition dipole moment integral can be divided into two terms

$$\langle \Psi_i|\mu|\Psi_f\rangle = \langle \psi_n \psi_e|\mu_e|\psi'_e\psi'_n\rangle + \langle \psi_n \psi_e|\mu_n|\psi'_e\psi'_n\rangle \quad (4)$$

Because of the orthogonality of the electronic wave functions of different electronic states, the second term of the rhs in eq 4 vanishes. As a consequence, the transition dipole moment integral depends on the nuclear wave functions and on the electronic transition moment $\mu_{if} = \langle \psi_e|\mu_e|\psi'_e\rangle$. Equation 4 can then be written

$$\langle \Psi_i|\mu|\Psi_f\rangle = \langle \psi_n|\mu_{if}|\psi'_n\rangle \quad (5)$$

However, since no general analytical expression exists for the electronic transition dipole moment, this integral must be approximated. A first approach, proposed by Franck and Condon[16−18] is based on the assumption that molecular geometry does not change significantly during the transition, and therefore the electronic transition dipole moment can be treated as a constant. While this approximation is known to lead to very good results in many cases, it becomes not satisfactory for the dipole-forbidden or weakly allowed transitions. In these cases, one needs to improve the model by expanding the transition dipole moment in a Taylor series of the normal coordinates (either the set Q' of the final state or the set Q of the initial state). For computational reasons that will become clear later we choose the expansion on Q'

$$\mu_{if}(Q') \cong \mu_{if}(Q'_0) + \sum_{k=1}^{N} \frac{\partial \mu_{if}}{\partial Q'_k} Q'_k + \frac{1}{2}\sum_{k=1}^{N}\sum_{l=1}^{N} \left(\frac{\partial^2 \mu_{if}}{\partial Q'_k\, \partial\, Q'_l}\right)_0 Q'_k Q'_l + \ldots \quad (6)$$

where Q'_0 refers to the equilibrium geometry of the final electronic state, while N is the number of normal modes.

In the current implementation we will consider the Taylor expansion up to the second derivatives. The zero-order, assuming that the transition dipole moment is unchanged during the transition, is the FC approximation. The Herzberg−Teller (HT) approximation corresponds to the development at the first order. In this case, we take into account a limited change in the structure during the transition.

Application of the Eckart conditions[28] allows minimization of the coupling between the rotational and vibrational motions of the nuclei in a molecule and as much as possible separate the nuclear wave function into rotational and vibrational contributions. Then, switching to the Dirac notation and assuming that the harmonic approximation can be used to represent the vibrational wave function of the initial and final states as eigenstates of the N-dimensional harmonic oscillator, i.e., by defining $|\psi_n\rangle = |v\rangle$ and $|\psi'_n\rangle = |v'\rangle$, it is possible to write eq 5 as

$$\langle \Psi_i|\mu|\Psi_f\rangle = \mu_{if}(Q'_0)\langle v|v'\rangle + \sum_{k=1}^{N}\left(\frac{\partial \mu_{if}}{\partial Q'_k}\right)_0 \langle v|Q'_k|v'\rangle + \frac{1}{2}\sum_{k=1}^{N}\sum_{l=1}^{N}\left(\frac{\partial^2 \mu_{if}}{\partial Q'_k\, \partial\, Q'_l}\right)_0 \langle v|Q'_k Q'_l|v'\rangle \quad (7)$$

where μ_{if} has been replaced by its Taylor expansion given in eq 6.

Computation of the overlap integrals between initial and final vibrational states requires the use of a common coordinates set. Duschinsky proposed a solution to this problem by considering a linear transformation between the normal modes of the initial state and the final state[29]

$$Q' = JQ + K' \quad (8)$$

The Duschinsky matrix J describes the projection of the normal coordinates basis vectors of the final state on those of the initial state and represents the rotation of the normal modes upon the transition. The shift vector K' represents the displacements of the normal modes between the initial state and the final state structures.

Since J in principle is not diagonal, the calculation of the vibrational overlap integrals is not straightforward, and several methods were devised to compute the different terms of the rhs of eq 7. Schematically, we can divide them into two categories, analytical approaches[30,31] and recursive approaches.[32−34] In our case both methodologies have been employed, exploiting their respective advantages. Analytical methods can quickly and accurately compute the transition dipole moment integrals through ad hoc formulas, but the latter need to be generated beforehand and individually coded in the program to be truly efficient. As a consequence, the possible transitions, in the sense of a particular combination of quantum numbers for the initial state and final state, need to be predicted. This task is really cumbersome and in practice unfeasible when considering medium- or large-sized molecules. Moreover, it is difficult to design general-purpose programs based solely on this approach.

On the other side, recursive approaches provide methods to compute the overlap between given initial and final states through formulas that express it in terms of sums of integrals involving states with lower vibrational quantum numbers. Once the overlap between ground vibrational states has been calculated directly, that of any other transition can be obtained by recursively applying these formulas. While the derivation of the latter formulas has been a prerequisite for the calculation of the harmonic spectrum for a generic system, additional conceptual and technical difficulties remain to be faced with. First, in principle there is an infinite number of possible transitions, and therefore some limitations must be applied as will be described in section 2.2. Second, while the need to avoid redundant calculations would suggest to store each computed overlap integral, this strategy would easily fail leading to a saturation of the computer memory, and therefore efficient algorithms for their storage and/or partial recomputation must be figured out.

For many studied systems, the FC and HT approximations are sufficient to correctly describe both absorption and emission spectra, but for symmetry-forbidden transitions a correct reproduction of intensity of weak bands may require the inclusion of higher-order terms.[35] Thus, as a first attempt to deal with more general cases we will introduce also the expansion to the second order of eq 6.

Recursion formulas for the overlap integrals between vibrational states have been derived using several methods.[34] Briefly, using second quantization, the terms $\langle v|Q'|_k v'\rangle$ and $\langle v|Q'_k Q'|_l v'\rangle$ can be expanded in simpler terms, and then the second and third terms of rhs of eq 7 can be written as a sum of overlap integrals

$$\begin{aligned}\langle \Psi_i|\mu|\Psi_f\rangle = {} & \mu_{\mathrm{if}}(Q'_0)\langle v|v'\rangle + \\ & \sum_{k=1}^{N}\left(\frac{\partial \mu_{\mathrm{if}}}{\partial Q'_k}\right)_0 \sqrt{\frac{\hbar}{2\omega'_k}}\left[\sqrt{v'_k}\,\langle v|v'-1'_k\rangle + \sqrt{v'_k+1}\,\langle v|v'+1'_k\rangle\right] + \\ & \sum_{k=1}^{N}\left(\frac{\partial^2 \mu_{\mathrm{if}}}{\partial Q'^2_k}\right)_0 \frac{\hbar}{4\omega'_k}\left[\sqrt{v'_k(v'_k-1)}\,\langle v|v'-2'_k\rangle + \right. \\ & \left. (2v'_k+1)\langle v|v'\rangle + \sqrt{(v'_k+1)(v'_k+2)}\,\langle v|v'+2'_k\rangle\right] + \\ & \sum_{k=1}^{N-1}\sum_{l=k+1}^{N}\left(\frac{\partial^2 \mu_{\mathrm{if}}}{\partial Q'_k \partial Q'_l}\right)_0 \frac{\hbar}{2\sqrt{\omega'_k\omega'_l}}\left[\sqrt{v'_k v'_l}\,\langle v|v'-1'_k-1'_l\rangle + \right. \\ & \sqrt{v'_k(v'_l+1)}\,\langle v|v'-1'_k+1'_l\rangle + \sqrt{(v'_k+1)v'_l}\,\langle v|v'+1'_k-1'_l\rangle + \\ & \left. \sqrt{(v'_k+1)(v'_l+1)}\,\langle v|v'+1'_k+1'_l\rangle\right] \quad (9)\end{aligned}$$

where $|v' + x'_l\rangle$ is the vibrational state with all quantum numbers equal to $|v'\rangle$ apart from mode l which has x'_l quanta more. In a recursive approach, when computing $\langle v|v'\rangle$, $\langle v|v'+1'_l\rangle$, $\langle v|v'+2'_l\rangle$, $\langle v|v'+1'_l-1'_k\rangle$, and $\langle v|v'+1'_l+1'_k\rangle$ are not yet calculated and need to be expressed with respect to integrals involving states with lower quanta. In those electronic methods where the excited-state Hessian is obtained by numerical derivatives of the analytical gradient (with positive and negative increments), first derivatives as well as diagonal second derivatives of the transition dipole moment are automatically obtained. We have taken into account both of them in our spectrum calculation (i.e., the second and third terms in eq 9) in the spirit of using all the available information provided by the electronic computational model.

2.2. Computational Strategy. The recursive approach described in the previous section can be successfully applied to the computation of spectra of medium-to-large molecular systems. However, in many cases the number of overlap integrals that must be taken into account can become extremely large, with a consequent increase in the required computational times and memory usage.

Efficient computational strategies must be able to individuate a priori the relevant transitions among the infinite number of possible final states. A first simple scheme[22,23] selects the transitions on the basis of their energy, considering only those falling in a predetermined energy range (which is in principle arbitrary). Such an approach is not effective enough for large systems with broad spectra, and more robust methods need to be devised. An estimation procedure should work a priori, not require too many computations, and correctly choose the relevant transitions. Several strategies have been devised to decrease the number of required calculations, while retaining the accuracy of the resulting spectra.[24,25,36,37] Here, we use an a priori method called FCclasses,[25,37] which provides very accurate vibrationally resolved spectra of medium and large molecular systems with limited computational resources. According to this method, transitions are partitioned into classes C_n, depending on the number n of simultaneously excited normal modes of the final electronic state of the transition. The overlap integrals for single vibrations (class C_1) and combination of two normal modes (class C_2) are computed up to a chosen limit (it can be also very large, since computation is cheap; therefore any loss of accuracy in this step can be avoided). The probabilities of all these transitions are stored and then used in the computation of FC integrals for higher-order classes to obtain a priori estimates of the maximum quantum number that needs to be considered for each normal mode. To that end, for each class the allowed transitions are chosen iteratively on the ground of a minimum threshold for C_1 and C_2 probabilities, so that the number of overlap integrals to be computed stays approximatively below a user-defined limit ($N_I^{\max}$), that rules the calculation accuracy (see ref 25 for further details). As we will show in the following, the partition of transitions in "classes" C_n allows a breakthrough in the spectrum calculation, since this latter quickly converges with the increase of the class order n. From a technical point of view, such partition is particularly appealing for a high-efficiency implementation since it leads to a straightforward parallelization of the calculations. Furthermore, separation in classes allows for an easy control on the memory requirements for storage of the overlap integrals used in recursive procedures.

2.3. Convergence of the Spectra Computations. When using an a priori method to selectively compute transitions, a general issue to address is the evaluation of the convergence of the calculations and hence of the reliability of their outcome. In the calculation of vibrationally resolved spectra, this can be easily done on the ground of analytical sum rules, by comparing the actual computed spectrum intensity I^n_{tot} to the exact analytical limit $I^{\mathrm{a}}_{\mathrm{tot}}$. In the current approach spectrum convergence is always improved by increasing the limit for the number of integrals computed for each *class* of transitions: $N_I^{\max}$ (MAXINT). A higher number of allowed transitions obviously yields a better spectrum convergence but also directly increases the required computational time and memory usage.

Skipping prefactors and the dependence on the frequency, the intensity of a given transition is equal to $I(v, v') = |\langle v|\mu_{\mathrm{if}}(Q')|v'\rangle|^2$ and summing over all the possible final states one gets

$$I^{\mathrm{a}}_{\mathrm{tot}} = \sum_{v'} |\langle v|\mu_{\mathrm{if}}(Q')|v'\rangle|^2 = \sum_{\rho=x,y,z}\langle v|\mu^2_{\mathrm{if}\rho}(Q')|v\rangle \quad (10)$$

where the superscript a indicates that the sum has been carried out analytically by exploiting the closure relation. In the limit of a complete inclusion of all the possible final states the numerical sum of the state-to-state intensities $I^n_{\mathrm{tot}} = \sum_{v'} I(v, v')$ must approach $I^{\mathrm{a}}_{\mathrm{tot}}$, and the ratio $C = I^n_{\mathrm{tot}}/I^{\mathrm{a}}_{\mathrm{tot}}$ can be used as a control of the convergence, which is

complete when $C = 1$. It is possible that, for large systems, a great number of transitions has to be considered to reach convergence of the spectrum intensity, and calculations although feasible can become computationally demanding. On the other side, usually the main scope is to correctly reproduce the spectrum line shape and assign the most important vibronic transitions. It has been shown[25,26,37] that the spectrum line shape usually converges much faster than C. The line shape convergence can be easily checked by comparison of results calculated with two different thresholds for $N_I^{\max}$.

For FC calculations eq 10 trivially gives $|\mu_{if}|^2$ and the convergence only depends on the sum of FC factors $|\langle v|v'\rangle|^2$, which must tend to 1. For HT calculations it has been shown in ref 37 that the integral in eq 10 is more easily evaluated by expanding μ_{if} on the normal coordinates of the initial state Q. By consideration of the cartesian component ($\rho = x$, y or z)

$$\mu_{if\rho}(Q) \simeq \mu_{if\rho}(Q_0) + \sum_{k=1}^{N} \frac{\partial \mu_{if\rho}}{\partial Q_k} Q_k = \mu_{if\rho}(Q_0) + M_{1\rho}^T Q \quad (11)$$

where $M_{1\rho}$ is the vector of the dipole derivatives $\partial\mu_{if\rho}/\partial Q$. Writing an analogous expression for the expansion as a function of Q'

$$\mu_{if\rho}(Q') = \mu_{if\rho}(Q_0') + M_{1\rho}'^{\,T} Q' \quad (12)$$

and substituting eq 8 into eq 11 one gets

$$\mu_{if\rho}(Q_0) = \mu_{if\rho}(Q_0') + M_{1\rho}'^{\,T} K' \quad (13)$$

$$M_{1\rho}^T = M_{1\rho}'^{\,T} J \quad (14)$$

These relations provide the data for the computation of I_{tot}^a

$$I_{tot}^a = |\mu_{if}(Q_0)|^2 + \sum_{k=1}^{N} \frac{\hbar}{2\omega_k}(2v_k + 1) \sum_{\rho=x,y,z} M_{1k\rho}^2 \quad (15)$$

An analogous procedure can be followed to compute I_{tot}^a for an expansion of μ_{if} up to the second order. One has

$$\mu_{if\rho}(Q) = \mu_{if\rho}(Q_0) + M_{1\rho}^T Q + Q^T M_{2\rho} Q = \mu_{if\rho}(Q_0') + M_{1\rho}'^{\,T} Q' + Q'^T M_{2\rho}' Q' \quad (16)$$

where

$$\mu_{if\rho}(Q_0) = \mu_{if\rho}(Q_0') + M_{1\rho}'^{\,T} K' + K'^T M_{2\rho}' K' \quad (17)$$

$$M_{1\rho}^T = M'^T_{1\rho} J + K'^T M_{2\rho}' J \quad (18)$$

$$M_{2\rho} = J^T M_{2\rho}' J \quad (19)$$

Now, by consideration of eq 10, expanding μ_{if} up to the second order and taking into account that terms in Q_k^n do not contribute by symmetry if n is odd, one gets

$$I_{tot}^a = |\mu_{if}(Q_0)|^2 + \langle v|\{\sum_k \sum_{\rho=x,y,z} [(2\mu_{if\rho}(Q_0) M_{2kk\rho} + M_{1k\rho}^2) Q_k^2 + M_{2kk\rho}^2 Q_k^4] + 2\sum_{\substack{k,l \\ k>l}} \sum_{\rho=x,y,z} (M_{2kk\rho} M_{2ll\rho} + 2M_{2kl\rho}^2) Q_k^2 Q_l^2\}|v\rangle \quad (20)$$

It can be easily shown that

$$\langle v|Q_k^2|v\rangle = \frac{\hbar}{2\omega_k}(2v_k + 1) \quad (21)$$

$$\langle v|Q_k^{\,2} Q_l^{\,2}|v\rangle = \frac{\hbar^2}{4\omega_k\omega_l}(2v_k + 1)(2v_l + 1) \quad (22)$$

$$\langle v|Q_k^4|v\rangle = \frac{\hbar^2}{4\omega_k^2}(6v_k^2 + 6v_k + 3) \quad (23)$$

By substitution of eqs 21− 23 in eq 20, one finally obtains the analytical sum for a second order expansion of μ_{if}

$$I_{tot}^a = |\mu_{if}(Q_0)|^2 + \sum_k \sum_{\rho=x,y,z} \frac{\hbar}{2\omega_k} \times [2\mu_{if\rho}(Q_0) M_{2kk\rho} + M_{1k\rho}^2](2v_k + 1) + \sum_k \sum_{\rho=x,y,z} \frac{\hbar^2}{4\omega_k^2} M_{2kk\rho}^2 (6v_k^2 + 6v_k + 3) + 2\sum_{\substack{k,l \\ k>l}} \sum_{\rho=x,y,z} \frac{\hbar^2}{4\omega_k\omega_l} (M_{2kk\rho} M_{2ll\rho} + 2M_{2kl\rho}^2)(2v_k + 1)(2v_l + 1) \quad (24)$$

In the current implementation we restrict ourselves to the diagonal terms of the second derivative of μ_{if}, thus the last term of eq 24 is not taken into account.

2.4. Anharmonic Corrections for the Ground and Excited Electronic States. As already stated, the methods outlined in previous sections are based on the harmonic approximation. However, improving the accuracy of simulated spectra requires going beyond harmonicity and taking into account anharmonic effects, couplings between modes, and vibrational or vibronic resonances. As a first step in this direction, some of us introduced[38] anharmonic corrections to the vibrational frequencies of both ground and excited electronic states. Here we briefly summarize a simple scheme to derive excited-state mode-specific scaling factors starting from the ground-state ones. These latter should be provided by the user, e.g., from perturbative anharmonic frequency calculations,[39] or from easily accessible ground-state experimental data. For each particular normal mode Q_k, the frequency scaling vector α is computed first, using the formula $\alpha(k) = \nu(k)/\omega(k)$ where ν is the anharmonic frequency and ω is the harmonic frequency. To proceed further, we shall assume that, if there is a one-to-one relation between the normal modes Q_k and Q_k' of the initial and final state, the scaling factors α_k and α_k' are equal. However, the normal modes are in general not coincident ($J \neq I$), and α cannot be transferred directly to scale the frequencies of the final state. In other words, the scaling vector must be adapted to the excited-state frequencies. In the case of small-amplitude vibrations, this can be obtained by expressing the normal modes of the excited state as linear combinations of the normal modes of the initial state by means of the Duschinsky transformation. The J_{kl} coefficients can now be applied to derive the relation between the initial (k) and final (l) state mode-specific anharmonicity scaling factors

$$\alpha_l' = \sum_k^N J_{lk}^2 \alpha_k \quad (25)$$

2.5. Implementation. The inclusion of vibronic computations in a general purpose electronic code (here Gaussian) requires an effective control of the memory usage. However, the storage of the overlap integrals for the recursion formulas is often the main bottleneck in FC calculations, and by use of a conventional method, keeping most integrals in an array is quickly unviable when dealing with large systems. Taking advantage of the capabilities of current computers, it is possible to consider a semidirect method, storing only a group of overlap integrals at a time. The computational costs induced by the necessary recalculations are generally on par with the times required to find a specific element in a large array. The main problem of a semidirect approach is to devise a consistent scheme to split the calculations of transition dipole moment integrals in groups requiring only a moderate need for recalculation in order to not hamper the efficiency of the spectrum generation. In our implementation, we chose to part the overlap integrals following a two-step procedure. First, they are naturally divided in classes as described in section 2.2. Second, each possible combination of n modes defines a set belonging to class C_n, and only the integrals calculated for a given set are stored at a time, thus dropping the storage requirements.

This partition technique also allows the implementation of several strategies to improve the efficiency of our calculations. First, it is not necessary to work on the full-dimensionality (N, the number of normal modes) Sharp and Rosenstock matrices (A, B, C, D, E),[30] and these latter, together with the diagonal matrices of reduced frequencies, can be replaced inside a given set by smaller matrices containing only the elements required by the calculations. The size of these new matrices, n for class C_n, is generally much lower than N speeding up the search for specific elements in these arrays. Additionally, the subdivision into well-defined sets allows the tailoring of the storage of the overlap integrals and their indexing. In fact, since the overlap integrals to be computed for a given set are known in advance, it is possible to choose a linear storage with a particular indexing function adapted to it, boasting the efficiency in retrieving a specific FC integral with respect to what would be allowed by a generic storage technique. Finally, different sets are independent, and therefore their calculation can be tackled in parallel, improving greatly the velocity of the spectrum simulation. Thanks to the double partition, two levels of parallelization can be pursued, either treating all the classes contemporaneously but their sets in serial or treating even these latter simultaneously.

In our computational strategy we calculate by default the intensities of the C_1 and C_2 transitions through analytical formulas. Nonetheless, a recursive calculation can be adopted also for these classes, if requested by the user. The present computational tool is designed for large systems where it may be necessary to take into account a large number of simultaneously excited modes to adequately simulate the spectra. Hence, a general method has been devised to let the order n of the highest class be set by the user.

The general integrated procedure to compute an absorption or emission spectrum is straightforward. As a first step the geometry structures of both electronic states and the frequencies of one of the two electronic states, by default the final one, have to be determined. These computations have to be performed beforehand and are stored in the internal checkpoint files. In this way all data needed in subsequent steps are extracted directly from internal files, and the accuracy of results is not arbitrarily truncated. The suite of modules dedicated to spectrum computations is integrated through a ad hoc subprogram, which proceeds to the treatment of the data available from previous calculations. Thus, the procedure to compute the vibrationally resolved spectrum extracts the stored data, and in a next step computes the frequencies for the second electronic state (by default the initial one) and the vibronic spectrum. When dealing with two structures computed independently, spurious effects might be introduced in the spectrum, if displacements due to translation and rotation are not carefully dealt with. While translation can be exactly removed by superposing the center of masses of the two equilibrium structures, rotational effects can be minimized by maximizing their mutual superposition. Our subprogram checks the superposition of the equilibrium structures and, if necessary, proceeds to their rotation and translation. All necessary data, such as frequencies and normal modes, are regenerated in this case.

The simplest computation of Franck−Condon spectrum requires the following data:

(1) Cartesian coordinates of the atoms.

(2) Masses of the atoms.

(3) Energy of the ground and excited states.

(4) Frequencies for the two electronic states involved in the transition.

(5) Normal modes for the two electronic states, expressed by the atom displacements.

All of which are available from frequency calculations performed for each electronic state. Additional data can be needed depending on the requested calculation. For HT calculations also the electronic transition dipole moment and its derivatives are required. For TD-DFT method, these data are directly available from the frequency calculations, since they are performed by numerical differentiation of the analytical TD-DFT energy gradients. If the electronic excited states are calculated with methods adopting analytical second derivatives by default, the numerical calculation of frequencies must be requested together with the HT calculation. Numerical derivatives of the transition dipole moment provide at no additional cost the diagonal terms of the second order of its Taylor expansion (see eq 6). Our method can take into account also these data, to utilize all the available information and maximize the accuracy of the computed spectra. Other data are also required to apply the anharmonic corrections. It is possible to provide anharmonic frequencies for both electronic states or derive anharmonic corrections for excited-state based on the ground-state data, as explained in section 2.4. In both cases proper sets of frequencies need to be included directly in the input for spectrum computation. In this way it is possible to apply anharmonic corrections by means of calculations performed beforehand (e.g., using the perturbative approach)[39] or making use of available (e.g., experimental) data. The spectrum convoluted with a Gaussian function [with half-width at half-maximum (SPECHWHM)

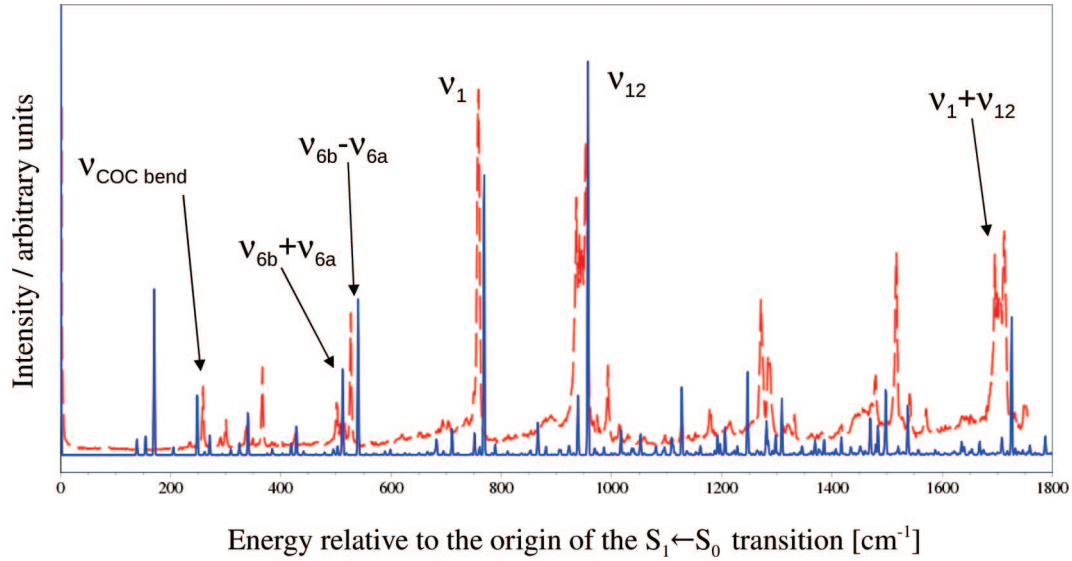

Figure 1. Computed TA (blue lines) and experimental REMPI[55] (red dashed lines) spectra of the $S_1 \leftarrow S_0$ transition of anisole along with the assingment of the most intense bands; see ref 38 for the details.

Table 1. Convergence of Spectra Computations for Adenine and Adenine@Si(100)[a]

	adenine		adenine@Si(100)	
class (n)	${}_NC_n$	progression	${}_NC_n$	progression
3	9.14×10^3	84.54%	4.27×10^7	87.31%
4	8.23×10^4	93.57%	6.75×10^9	94.82%
5	5.76×10^5	97.48%	8.54×10^{11}	97.37%
6	3.26×10^6	98.32%	8.98×10^{13}	97.88%
7	1.54×10^7	98.39%	8.08×10^{15}	97.93%

[a] For each class C_n the number of combinations of the n excited oscillators ${}_NC_n$ and spectrum progression are listed. The C_1 and C_2 transitions have been computed by analytical formulae allowing a maximum quantum number $v_i = 30$, and $v_1 = v_2 = 20$ (MaxC1 = 30, MaxC2 = 20) respectively. For the classes C_n, $n \geq 3$ the transitions to be computed have been selected setting the parameter N_I^{max} to 10^8 (the default value).

of 135 cm^{-1} by default] is generated as an outcome of vibronic calculations (Lorentzian broadening can be introduced as well). The assignment of vibronic transitions is available through the corresponding overlap integrals that are printed up to a chosen threshold (PRTINT), the transitions with intensity larger than 1% of $I(0, 0')$ are printed out as default. The convergence of the spectrum calculations and the contribution of the classes are also reported together with the requested matrices (e.g., J).

The parameters that rule spectra calculations can be divided into two groups, the first being related to the spectrum output as the convolution and its spectral range and grid, etc., which can be freely chosen in a manner appropriate for the problem under study. The second set of parameters is related to the accuracy of the computed spectrum. These are the maximum quantum number for each mode in C_1 and C_2 integrals (MAXC1 and MAXC2, respectively) and the maximum number of simultaneously excited modes MAXBANDS, which all determine the transitions to be computed, together with an additional parameter N_I^{max} (MAXINT). This latter limits an a priori and computationally cheap estimate of the number of integrals N_I to be computed for each class (see ref 25). This quantity is evaluated through the expression: $N_I = {}_NC_n \times \langle v'_{max} \rangle^n$. The binomial coefficient represents the number of combinations of the n excited oscillators and $\langle v'_{max} \rangle$ is the arithmetic mean of the N maximum quantum numbers v'_{max}. In the following, for brevity, we will generically indicate N_I^{max} as the maximum number of integrals computed for each class. An internal check procedure has been devised to monitor computation convergence in order to prevent unnecessary calculations but also to warn for the cases when transitions to higher quanta need to be considered. Various aspects of spectrum convergence with respect to these parameters are discussed further in sections 4.2, 4.3, and 4.5.

3. Computational Chemistry Models

Full geometry optimizations and harmonic frequency calculations have been performed for all systems under study and in both final and initial electronic states. In some cases anharmonic perturbative[39] calculations have been performed for the ground electronic state, and subsequently used to account for anharmonicity in both S_0 and S_1 electronic states, as described in section 2.4. The computational chemistry methods have been chosen accordingly to the system under study in order to find a satisfactory balance between feasibility of calculations and accuracy of results. For anisole molecule the ground and excited states computations have been carried out using DFT and TD-DFT,[19] respectively, with the well-known B3LYP functional[40] and the 6-311+G(d,p) basis set. For SF_6 and its negative ion, computations have been performed at the MP2[41] level with aug-cc-pVTZ basis set.[42,43] For chlorophyll c2 both singlet ground and triplet excited electronic states have been calculated at the DFT level with the PBE0[44] functional and 6-31G(d) basis set. Solvent effects on the UV spectrum of acrolein have been introduced by a model where the solute molecule is treated explicitly by means of DFT with B3LYP functional and N07D double-ζ basis set,[45] and the continuum medium is described by the CPCM.[46] In the case of adenine molecule adsorbed on Si(100) surface the ONIOM[47] quantum me-

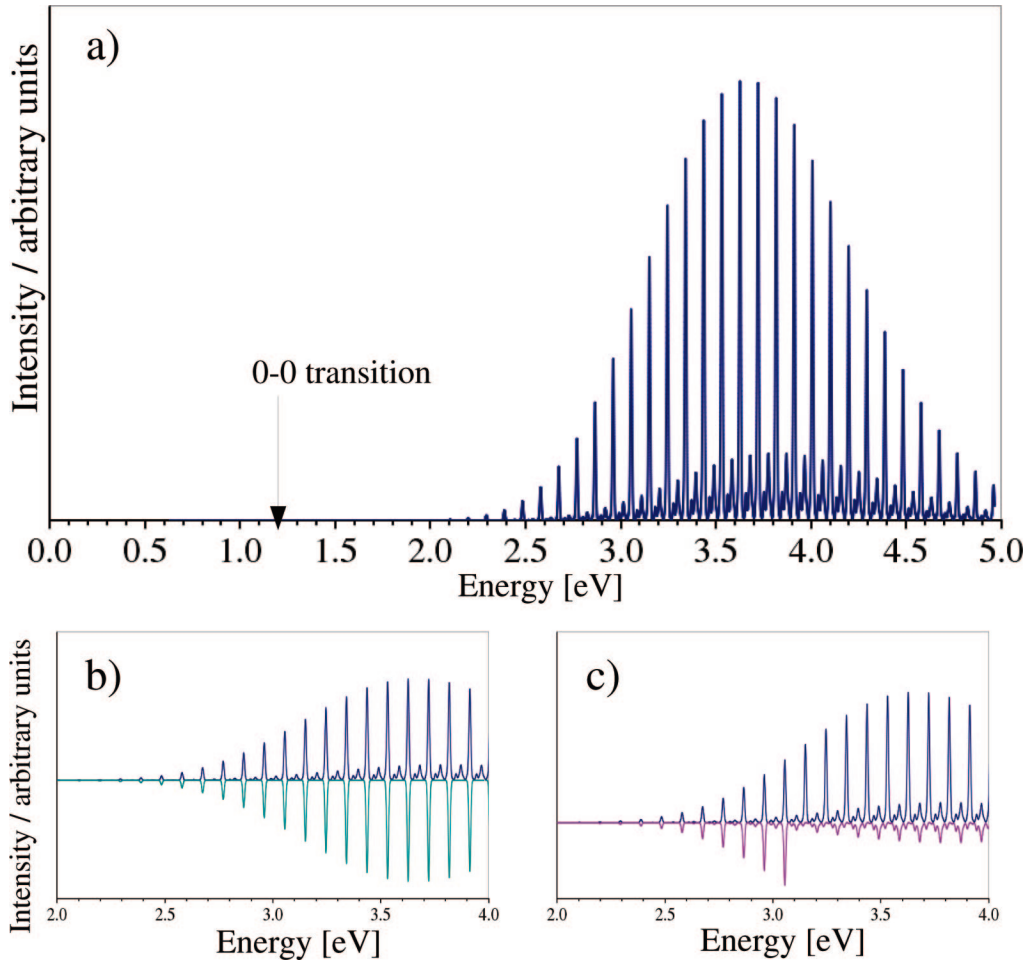

Figure 2. Theoretical electron photodetachment spectrum of SF_6^-. (a) Full spectrum in a range from 0 to 5 eV calculated within FC approximation with combinations between all modes considered; the energy of 0−0 transition is marked by an arrow. (b) Comparison between the full spectrum (upper panel) and a spectrum with coupling between modes excluded (lower panel). (c) Comparison between the full spectrum calculated with MAXC1 set to 100 (upper panel) vs the one with the maximum 20th overtone considered (lower panel).

chanical/molecular mechanical (QM/MM) scheme has been adopted with the Si(100) surface represented by a cluster of 119 silicon atoms. The QM part corresponding to the adenine molecule has been calculated at the B3LYP/6-31+G(d,p) level, while the cluster has been modeled by molecular mechanics using the UFF force field.[48] In the QM/MM calculation, the MM part has been polarized taking into account the interaction with the ground-state of the neutral molecule or its cation, respectively. Moreover in some cases the electronic energies of the initial and final electronic states have been refined at the coupled cluster level.[49−51] The CCSD and EOM-CCSD calculations have been performed with the MOLPRO[52] package. All other calculations have been performed with a locally modified version of the Gaussian suite of quantum chemistry programs.[53]

4. Applications

The integrated approach to compute vibrationally resolved optical spectra can be applied to a large variety of systems ranging from small molecules in the gas phase to macrosystems in condensed phases, whenever nonadiabatic couplings are negligible and harmonic approximation is reliable. The given examples of absorption spectrum of $S_1 \leftarrow S_0$ electronic transitions of anisole, photodetachment spectrum of SF_6^-, emission $T_1 \rightarrow S_0$ phosphorescence spectrum of chlorophyll c2, UV spectrum of acrolein in the gas phase and aqueous solution and a photoelectron spectra of adenine adsorbed on the Si(100) surface, are chosen to illustrate the flexibility of the present computational tool.

4.1. Vibrationally Resolved Optical Spectrum of Anisole. The recently published vibrationally resolved absorption spectrum of the $S_1 \leftarrow S_0$ electronic transitions of anisole[38] represents an example of the simulation accuracy achievable when good quality geometries and force fields for both electronic states are provided. The planarity of anisole in both electronic states has been determined by the high resolution laser-induced fluorescence (LIF) spectroscopy.[54] For anisole, methods based on the density functional theory and its time-dependent extension for electronic excited states [B3LYP/6-311+G(d,p) and TD-B3LYP/6-311+G(d,p)] have been applied to geometry optimizations and harmonic frequency calculations. The remarkable overall agreement between theoretical and experimental[54] rotational constants (average deviation of about 0.5%, for both electronic states) confirms good quality of the calculated geometry structures. The relative energy of the electronic states has been refined

by EOM-CCSD/CCSD//aug-cc-pVDZ computations. The vibrational frequencies in the first excited electronic state have been corrected according to the ground-state experimental frequencies (EA) or to the calculated perturbative anharmonic frequencies[39] (TA). The spectrum has been computed with the default choice of parameters yielding 99.6% of the total spectrum intensity. The simulated vibronic profile convoluted with a fwhm of 2 cm^{-1} is compared to the highly accurate experimental data from resonance-enhanced multiphoton ionization (REMPI) spectrum[55] in Figure 1. On the whole, a very good agreement has been achieved (the root-mean-square deviation between computed and experimental bands is 15 cm^{-1}), as detailed in ref 38. To reproduce correctly the band intensities and the rich vibrational structure of the REMPI spectra, it has been necessary to account for changes in structure, vibrational frequencies, and normal modes between the involved electronic states. It is worthwhile highlighting that the remarkable overall agreement, also as far as band positions are concerned, has only been possible by correcting the frequencies for anharmonicity. The discrepancy between the absolute position of experimental and simulated spectra remains the main shortcoming of the purely theoretical approach: to achieve a good match between spectra, the energy of the electronic transition should be computed with an accuracy of $\sim$10 cm^{-1}. Thus, even if DFT/TD-DFT computations are able to provide quite reasonable estimates of the relative energetics of the electronic states (within 0.2 eV) and despite the refinement based on coupled clusters calculations (0.05 eV), it was still necessary to compare spectra shifted to the 0−0 origin. Nevertheless, the remarkable agreement between theoretical and experimental spectra allowed for revision of some assignments of fundamental vibrations in the S_1 state of anisole. In particular, for many bands that had been assigned to S_1 fundamentals, consideration of the relative intensities has suggested instead a different interpretation as combinations or overtones.[38]

4.2. Photodetachment Spectrum of SF_6^-. SF_6^- represents an interesting case of a relatively small and highly symmetric system where accurate ab initio methodologies can be applied, giving possibility for an easy computation of accurate theoretical spectra which can be of great value for the interpretation of the best available experimental spectroscopic data in the gas phase. The photodetachment spectrum of SF_6^- is characterized by a broad progression, with the band maximum shifted by more than 2 eV from the 0−0 transition corresponding to the adiabatic electron affinity (AEA). These features have been attributed to the significant changes in the geometry between ionic and neutral species.[56,57] Indeed MP2/aug-cc-pVTZ calculations show a significant elongation of the S−F bond upon electron attachment (from 1.5750 to 1.7146 Å), while the molecule octahedral symmetry remains unchanged. Moreover, the theoretical [CCSD(T)/aug-cc-pVTZ] AEA of 1.06 eV is in very good agreement with the experimental value of 1.0 eV.[58] The fully theoretical photodetachment spectrum, calculated in an energy range from 0 to 5.0 eV within the FC approximation on the basis of the aforementioned ab initio results (see Figure 2) clearly resembles its recently measured

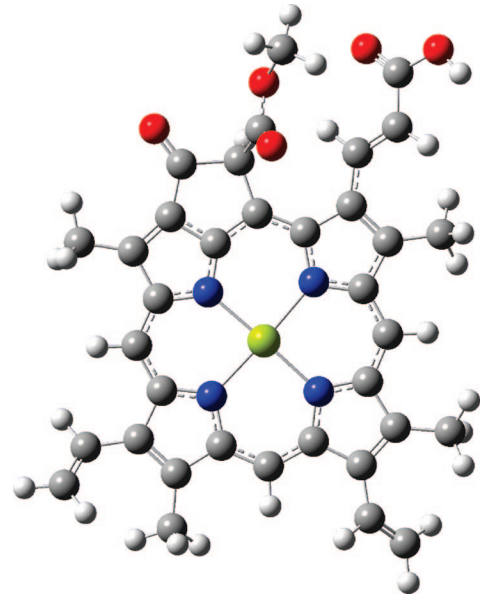

Figure 3. Structure of chlorophyll c2.

experimental counterpart.[56] For the SF_6^- photodetachment spectrum most of the vibrational progression derives from excitation of the single totally symmetric S−F stretching mode (ν_1). Indeed, the computed spectrum shows a regular pattern of bands, the most intense corresponding to ν_1. Panel b of Figure 2 shows a comparison between the spectrum obtained with only the FC integrals from class C_1 and the complete spectrum, where couplings between all modes are also taken into account. In the latter case it is found that the weaker bands gain intensity from excitation of this totally symmetric mode, being related to the combinations between the overtones of ν_1 and doubly excited degenerate mode ν_4 or ν_5. As a consequence of the large changes in the S−F bond length, the most intense transitions are related to high overtones, which must therefore be considered to reproduce accurately the spectrum features. Our approach allows such a flexibility, through the change of the MAXC1 parameter to a user-defined value instead of the default value of 20. Indeed in the latter case only about 30% of the spectrum intensity has been achieved in comparison to 94% when all necessary excitations have been taken into account. The comparison of the spectra calculated with MAXC1 set to 100 and to 20 is shown on panel c) of Figure 2.

4.3. Phosphorescence Spectrum of Chlorophyll c2. Despite ongoing experimental and theoretical research, the understanding of the molecular mechanism of light harvesting in photosystem II is not yet satisfactory. Quantum chemical computations of optical properties combined with spectroscopic experiments can undoubtedly contribute to shed further light on this phenomenon.[59] The triplet states of chlorophylls are of particular interest due to their dual photodamage and photoprotective role in photosystem II.[60] Chlorophyll c2 (see Figure 3) is a large molecule with 73 atoms and 213 normal modes. Its $T_1 \rightarrow S_0$ phosphorescence spectrum has been chosen to demonstrate various aspects related to spectrum convergence and applicability of the integrated approach. More detailed studies of spectra of photosynthetic pigments are planned in future works. The

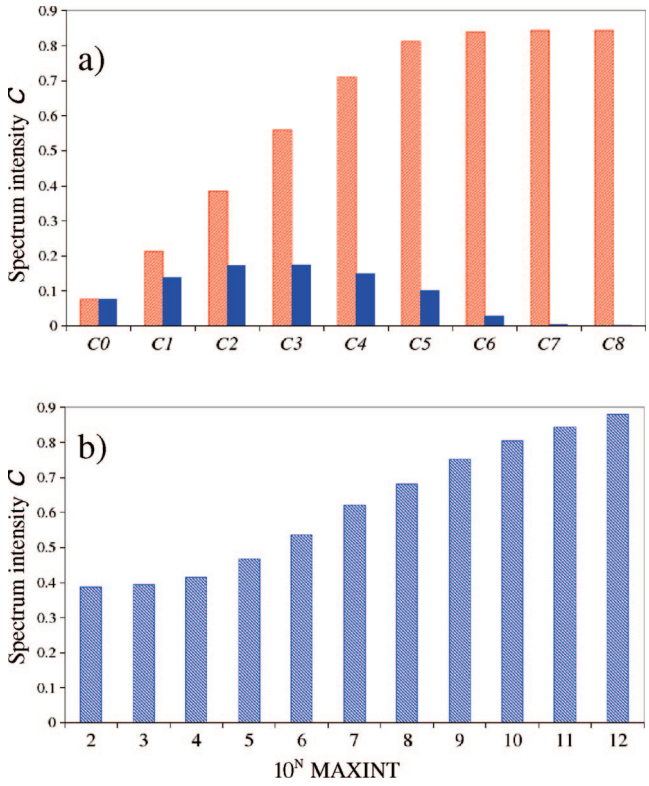

Figure 4. Convergence of the spectrum calculation for chlorophyll c2 with classes (upper panel) and with the maximum number of integrals MAXINT set for each class (lower panel). For spectrum convergence with classes the total convergence up to class C_n is shown as red blocks, while contribution of class C_n is shown in blue.

main factor which rules the accuracy of the calculations is the maximum number of integrals which are computed. Higher number of allowed transitions obviously improves the spectrum quality but at the same time strongly increases the required computational times and memory usage. For larger systems it becomes crucial to select only the most important transitions. In the present approach transitions are partitioned in classes, and the pool of C_1 and C_2 transitions is used to estimate a priori the maximum quantum number that has to be considered for each mode (see section 2.2) in order to keep the number of integrals computed for each class within the limit determined by MAXINT. This procedure has been proved to be very cost-effective.[25,26,37] Panel a of Figure 4 shows the spectrum convergence with classes calculated with MAXINT set to 10^{11}, while the spectrum convergence with increase of MAXINT (from 10^2 to 10^{12}) is depicted on panel b. As expected, a very small number of integrals is not sufficient and leads only to about half of the spectrum intensity. Moreover, even as many as 10^{12} integrals cannot provide the full convergence of spectrum intensity (C). On the other hand it is apparent that the contribution of classes higher than C_5 decreases steeply and that the difference in spectrum intensity calculated up to C_7 and up to C_6 is smaller than 1%, confirming spectrum convergence with respect to classes.

Nevertheless, in most cases the convergence of the spectrum line shape is much faster[25] than the full convergence of the total spectrum intensity. This fact is particularly

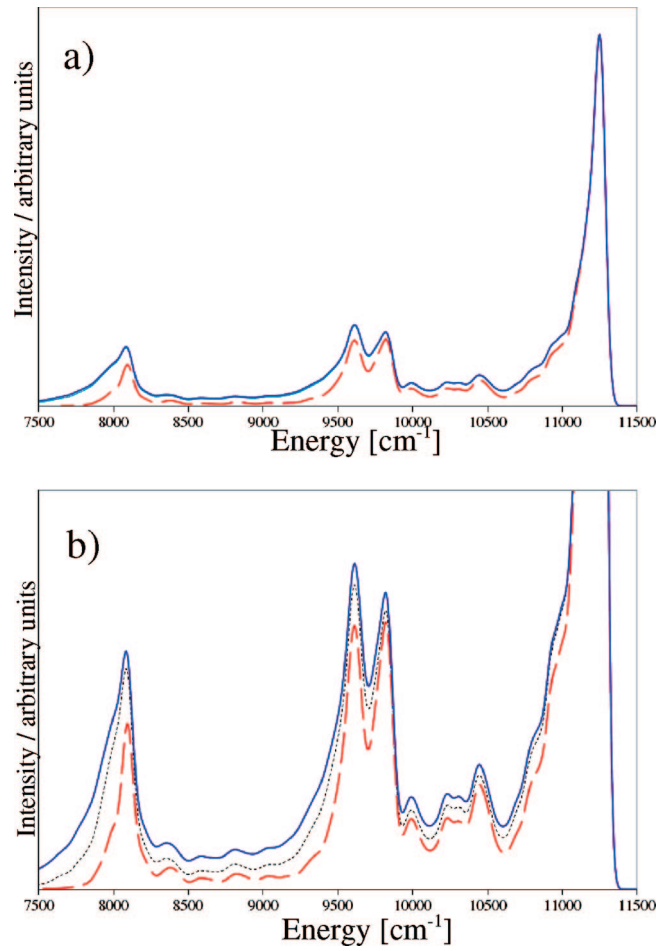

Figure 5. Convergence of the spectrum calculation for chlorophyll c2 with the threshold on the number of computed integrals. Comparison of spectrum shape calculated with MAXINT set to 10^2 (dashed red line) and 10^9 (solid blue line) is shown on upper panel, while the onset with spectra calculated with MAXINT set to 10^2, 10^6 (fine-dashed black line), and 10^9 is shown on lower panel.

encouraging for large systems like chlorophyll c2. Figure 5 compares spectrum line shapes calculated with MAXINT set to 10^2, 10^6, and 10^9. It is clear that the main spectral features are well reproduced even if total spectrum intensity is far from convergence. The spectra calculated with MAXINT = 10^9 or larger are identical on this scale. Thus, inspection of the spectrum line shape indicates that the most important transitions have been taken into account, and that reliable spectra have been computed already with N_I^{max} set to 10^9. The analysis of classes contributions to the total spectrum (Figure 6) shows that most of the spectrum bands are composed from classes up to C_4, with C_1 and C_2 influencing most the spectrum line shape. Contributions of the classes related to the simultaneous excitation of five and more modes are much flatter and of little importance for the spectrum line shape, although they are not negligible for the spectrum intensity. The present case of chlorophyll turns out to be much more challenging than that adopted as a benchmark by Dierksen et al.[24] and Jankowiak et al.,[27] a very large polycyclic aromatic hydrocarbon (PAH) derivative with 462 normal modes. Such PAH has a rather narrow photoelectron spectrum, and our method is able to converge it up to values

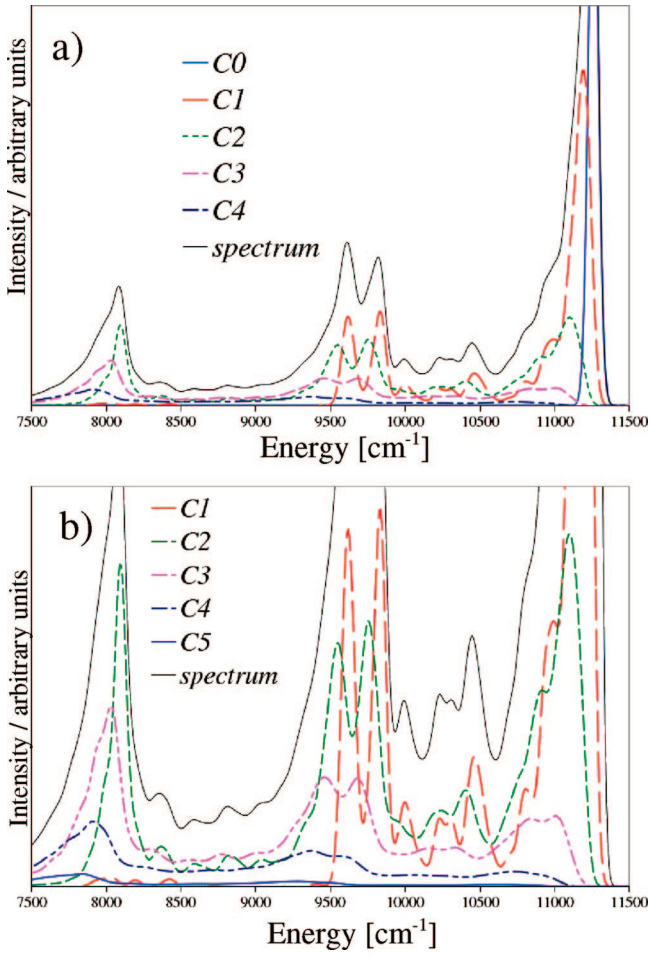

Figure 6. Convergence of the spectrum calculation for chlorophyll c2 with classes; contributions of specific classes are compared with total spectrum (see legend). Classes C_0-C_4 are shown on upper panel, while the onset with contribution of classes C_1-C_5 is shown on lower panel. Contribution of higher classes are not visible in this scale.

of C larger than 0.99. These tests show that our methodology can satisfactorily compute converged spectra also for large challenging systems. When the interest is focused on the high-energy wing of the spectrum (the one suffering of the largest relative error) as for instance for computation of nonradiative transition rates, a careful check of convergence in that energy region must be performed and purposely tailored methods may result more suitable.

4.4. UV Spectrum of Acrolein in the Gas Phase and Aqueous Solution. The UV absorption spectrum of acrolein has attracted significant attention since this molecule exhibits two conjugated chromophores C═C and C═O, a common feature for many natural systems. In particular, a blueshift of the $n \rightarrow \pi^*$ transition of the C═O group has been observed in going from gas phase to aqueous solution.[61] Our integrated approach allows a straightforward computation of the gas phase and aqueous solution absorption spectra of acrolein, giving direct insights into the experimentally observed effect. The structures and frequencies of acrolein have been determined by DFT/TD-DFT computations with the B3LYP functional and N07D polarized double-ζ basis set, both in gas phase and in aqueous solution. The effect of water solvent has been included by means of the polarizable continuum

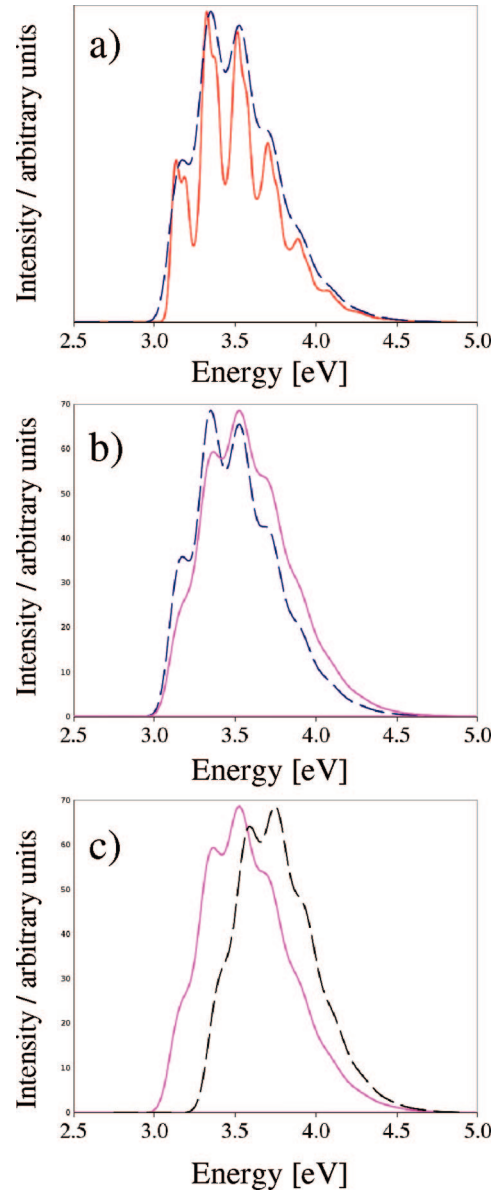

Figure 7. Theoretical absorption UV spectra of $n \rightarrow \pi^*$ electronic transition of acrolein. (a) Gas phase spectrum in a range from 2.5 to 5 eV calculated within FC approximation with fwhm of 500 (red line) and 1000 cm^{-1} (blue dashed line). (b) Comparison between gas-phase spectra calculated within either the FC (blue dashed line) or the FC-HT approximations (pink line). (c) Comparison between calculated spectra for acrolein in gas phase (pink line) and in water solution described by the CPCM model (black dashed line).

model, where the solvent is represented by a homogeneous dielectric polarized by the solute, placed within a cavity built as an envelope of spheres centered on solute atoms.[46] The solvent is described in the nonequilibrium limit where only its fast (electronic) degrees of freedom are equilibrated with the excited-state charge density while the slow (nuclear) degrees of freedom remain equilibrated with the ground state. This assumption is sufficient to describe absorption spectrum in solution, because of the different time scales of the electronic and nuclear response components of the solvent reaction field.[25] To simulate the spectrum line-shape it is necessary to convolute the stick-spectrum with a Gaussian

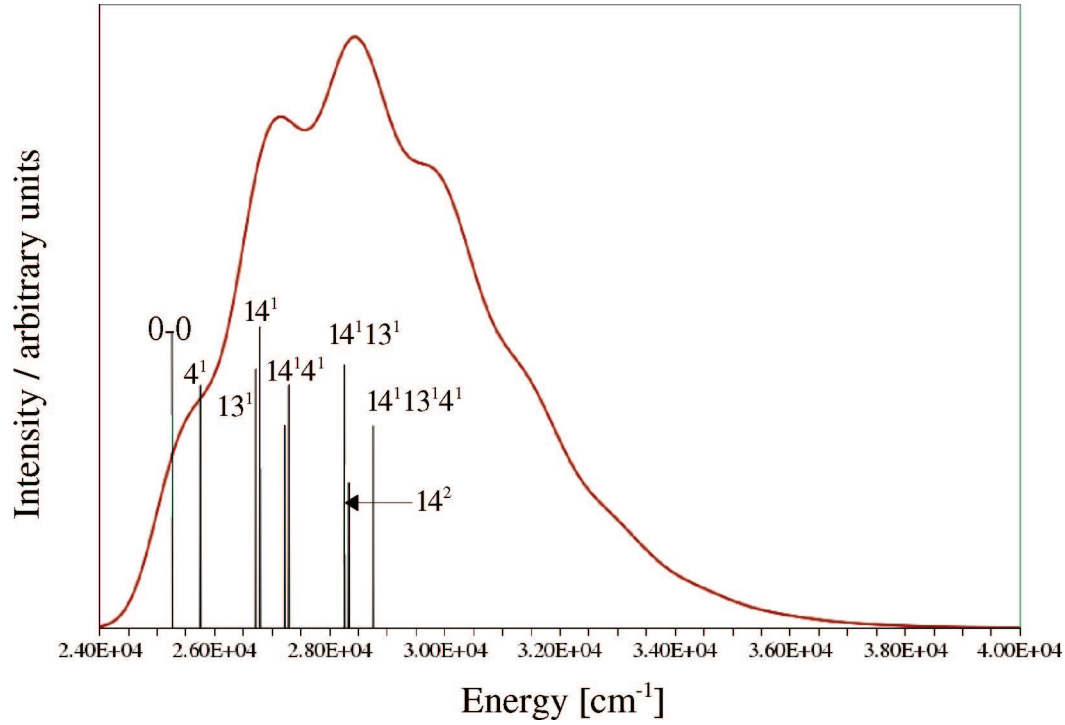

Figure 8. Assignment of the main bands of the theoretical absorption UV spectrum of $n \rightarrow \pi^*$ electronic transition of acrolein in gas phase. Solid line reports the spectrum in a range of 24000−40000 cm^{-1} calculated within FC-HT approximation with fwhm of 1000 cm^{-1}. The main stick bands are assigned as n^x where n is the excited normal mode and x its quantum number.

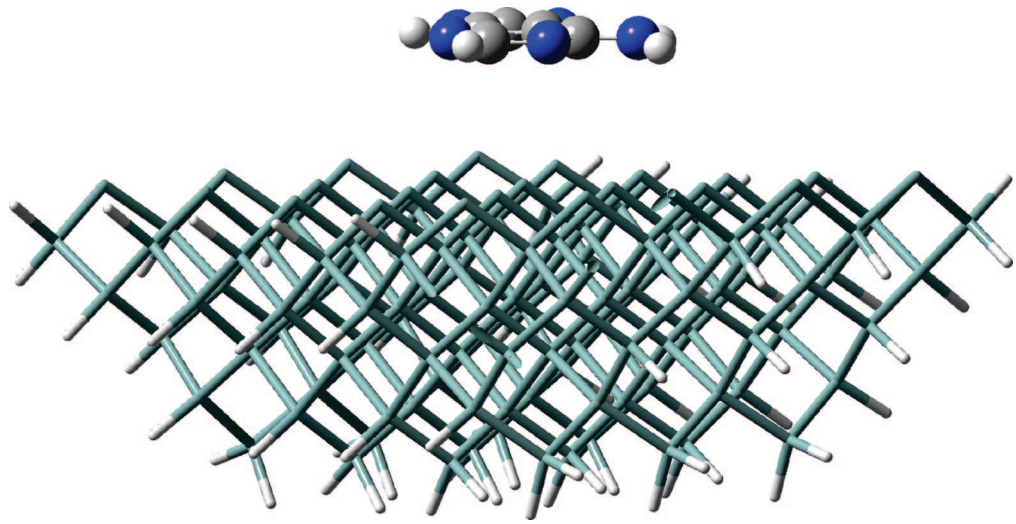

Figure 9. Adenine adsorbed on a cluster of 119 silicon atoms, modeling the Si(100) surface.

with an appropriate full-width at the half-maximum (fwhm): panel a of Figure 7 compares spectra calculated with the values of the fwhm set to 500 and 1000 cm^{-1}. For acrolein the latter choice better reproduces the broad structure of the experimental spectrum. In the present approach it is possible to improve the FC spectrum by considering changes of the transition dipole moment with the geometry. It is worth mentioning that in the present case inclusion of the HT term does not require any additional quantum mechanical computation, since the TD-DFT frequencies are calculated numerically giving direct access to the necessary derivatives of the transition dipole moment with respect to the normal coordinates of the excited electronic state. Inclusion of the HT term is particularly important for dipole forbidden or weakly allowed transitions where the FC approximation is unreliable. This is the case of the weakly dipole allowed $n \rightarrow \pi^*$ transition of acrolein ($\mu = 0.0463$ au), where the HT contribution indeed influences significantly the spectrum lineshape, as shown by the comparison of the FC and FC-HT spectra on panel b of Figure 7. Both the FC and the FC-HT spectra are fully converged (100.0%) to their respective limits (see section 2.3). The FC-HT spectra calculated in the gas phase and in the aqueous solution are compared in panel c of Figure 7. It is evident that not only the solvent shift is well reproduced by the theory but that also changes in the band shapes agree with the recent results obtained with a

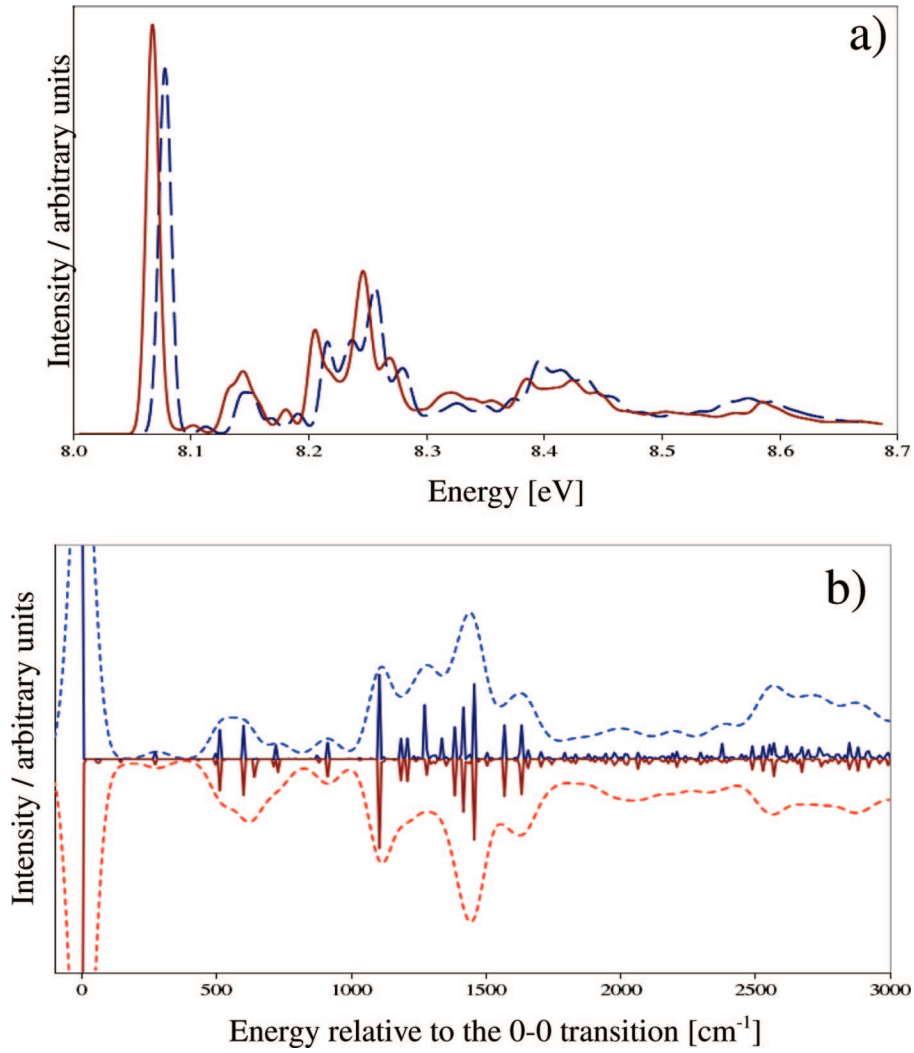

Figure 10. Comparison between the theoretical FC photoionization spectra in gas phase of isolated adenine (blue dashed line) and adenine adsorbed on a Si(100) surface (solid red line): (a) Spectra in an absolute energy range from 8.0 to 8.7 eV calculated within FC approximation with fwhm = 100 cm^{-1}. (b) Spectra shifted to the relative origins of the 0−0 electronic transitions, isolated molecule (upper panel), and adenine@Si(100) (lower panel); the stick bands show the most important transitions.

more accurate but computationally demanding time-domain approach.[61] Moreover, the present method works directly in the frequency domain, so that all the individual vibronic contributions to the total spectrum are computed and can be easily assigned, as shown in Figure 8. We foresee that the accessibility to an easy and straightforward method for the computations of vibrationally resolved spectra within the integrated approach here described, may lead to breakthroughs in the studies of UV−vis spectra in condensed phases.

4.5. Photoelectron Spectrum of Adenine Adsorbed on Si(100). Reliable computational studies of optical properties for large nanosystems in condensed phases can support the design of new materials relevant for optics, photonics, and sensoristics. Our approach is in line with such a demand as illustrated by simulation of the photoelectron spectrum of adenine adsorbed on Si(100) surface. The full valence photoelectron spectrum of adenine is composed from several overlapping excitations.[62] The present work is aimed to show feasibility of spectra simulations for nanosystems; thus only ionization from the highest occupied molecular orbital (HOMO) has been considered. The Si(100) surface has been modeled by a cluster of 119 silicon atoms, shown in Figure 9, resulting in a total system with 636 normal modes. For computation of geometry structures and frequencies the ONIOM QM/MM scheme has been adopted. The photoelectron spectra have been calculated for both isolated adenine molecule and adenine@Si(100), putting into evidence spectrum changes upon adsorption. Both spectra are plotted in the range of 8.0−8.7 eV roughly corresponding to the first band of valence shell photoelectron spectrum. Figure 10 shows the spectra both in the absolute energy scale and in a relative scale where the 0−0 transition is set to zero. It can be seen that adsorption on Si surface yields a small red shift of the excitation origin, while new vibronic transitions corresponding to intermolecular vibrations modulate the spectrum line shape. It is interesting to analyze the number of combinations for each class C_n for such a large system, which is directly related (see section 2.5) to the number of transitions to compute, and investigate the efficiency of the adopted selection procedure. Table 1 lists ${}_NC_n$ for isolated adenine and adenine@Si(100) and the spectrum intensity

achieved with MAXINT set to the default value 10^8. It is noteworthy that in both cases, either an isolated molecule with 39 normal modes or a macrosystem with over 600, almost all spectrum intensity (about 98%) has been recovered at an equivalent computational cost. For the cluster, the default value of MAXINT is not sufficient to consider the whole initial pool even for only three simultaneously excited modes (C_3 class). This particular case shows the ability of the a priori strategy to select only the relevant transitions and discard the less probable ones.

Conclusions

A general approach for simulation of absorption and emission vibronic spectra has been implemented and applied to a variety of molecular systems, showing the high flexibility of the developed computational tool. The integration of all procedures within the same computational package allows for the fully automatic computation of vibrationally resolved optical spectra. Despite the fact that our computational scheme has been tailored for large systems, it can be utilized as well to generate high quality spectra for small systems, when nonadiabatic and anharmonic couplings are negligible, since it does not imply unnecessary approximations. Moreover, good quality spectra can be effectively computed even for large systems with hundreds of normal modes, paving the route to spectroscopic studies of systems of direct biological and/or technological interest.

It has been demonstrated that a very good agreement between the computed and experimental vibrationally resolved REMPI spectra can be achieved, when good accuracy geometries and force fields are determined for the ground and excited electronic states, for example, for the case of anisole. It is worthwhile highlighting that the remarkable overall agreement, also as far as band positions are concerned, has been achieved only through the correction of the frequencies for anharmonicity. The accuracy and effectiveness of our a priori procedure for the selection of the relevant transitions has been shown by comparison of the photoelectron spectra computed for isolated adenine and for adenine adsorbed on silicon surface. Despite significant difference in the systems size, in both cases the electronic transition is localized on adenine molecule. The implemented scheme indeed effectively chooses the relevant transitions, providing an equivalent computational cost for spectra computations for isolated molecule and adenine@Si(100). Spectrum quality has been discussed on the example of chlorophyll c2 showing that line shape converges much faster than spectrum intensity. The former, which provides all necessary information for most important transitions, is usually the actual interesting property. Thus, reliable computations for large systems can be performed with relatively small computational cost. Overall, we hope that the tool here presented to compute vibrationally resolved optical spectra, together with its integration into a computational chemistry package, will routinely allow detailed analyses of UV−vis spectra, improving their interpretation and understanding.

Acknowledgment. This work was supported by Italian MIUR. The large scale computer facilities of the VILLAGE network (http://village.unina.it) are kindly acknowledged. CINECA and the Wroclaw Centre for Networking and Supercomputing are acknowledged for access to the MOLPRO code and related computer time.

References

(1) Barone, V.; Polimeno, A. *Chem. Soc. Rev.* **2007**, *36*, 1724.

(2) Barone, V.; Improta, R.; Rega, N. *Acc. Chem. Res.* **2008**, *41*, 605.

(3) Clary, D. C. *Science* **2006**, *314*, 5797.

(4) Hellman, A.; Baerends, E. J.; Biczysko, M.; Bligaard, T.; Christensen, C. H.; Clary, D. C.; Dahl, S.; van Harrevelt, R.; Honkala, K.; Jonsson, H.; Kroes, G. J.; Luppi, M.; Manthe, U.; Norskov, J. K.; Olsen, R. A.; Rossmeisl, J.; Skulason, E.; Tautermann, C. S.; Varandas, A. J. C.; Vincent, J. K. *J. Phys. Chem. B* **2006**, *110*, 17719.

(5) Jacquemin, D.; Preat, J.; Wathelet, V.; Fontaine, M.; Perpete, E. *J. Am. Chem. Soc.* **2006**, *128*, 2072.

(6) Altoe, P.; Bernardi, F.; Garavelli, M.; Orlandi, G.; Negri, F. *J. Am. Chem. Soc.* **2005**, *127*, 3952.

(7) *Computational Molecular Spectroscopy*; Jensen, P., Bunker, P. R., Eds.; John Wiley & Sons: London, 2000.

(8) Lami, A.; Petrongolo, C.; Santoro, F. In *Conical Intersections, Electronic Structure, Dynamics and Spectroscopy*; Domcke, W., Yarkony, R., Koeppel, H., Eds.; World Scientific Publishing Co.: Singapore, 2004; pp 699−739.

(9) Grimme, S. *Rev. Comput. Chem.* **2004**, *20*, 153.

(10) Christiansen, O. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2942.

(11) Bowman, J.; Carter, S.; Huang, X. *Int. Rev. Phys. Chem.* **2003**, *22*, 533.

(12) Vaara, J. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5399.

(13) Brancato, G.; Barone, V.; Rega, N. *Theor. Chem. Acc.* **2007**, *117*, 1001.

(14) Barone, V.; Brustolon, M.; Cimino, P.; Polimeno, A.; Zerbetto, M.; Zoleo, A. *J. Am. Chem. Soc.* **2006**, *128*, 15865.

(15) Rega, N. *Theor. Chem. Acc.* **2006**, *116*, 344.

(16) Franck, J. *Trans. Faraday Soc.* **1926**, *21*, 536.

(17) Condon, E. *Phys. Rev.* **1926**, *28*, 1182.

(18) Condon, E. *Phys. Rev.* **1928**, *32*, 858.

(19) Scalmani, G.; Frish, M. J.; Menucci, B.; Tomasi, J.; Cammi, R.; Barone, V. *J. Chem. Phys.* **2006**, *124*, 094107.

(20) Furche, F.; Ahlrichs, R. *J. Chem. Phys.* **2004**, *121*, 12772.

(21) Köhn, A.; Hättig, C. *J. Chem. Phys.* **2003**, *119*, 5021.

(22) Kemper, M.; Van Dijk, J.; Buck, H. *Chem. Phys. Lett.* **1978**, *53*, 121.

(23) Berger, R.; Fischer, C.; Klessinger, M. *J. Phys. Chem. A* **1998**, *102*, 7157.

(24) Dierksen, M.; Grimme, S. *J. Chem. Phys.* **2005**, *122*, 244101.

(25) Santoro, F.; Improta, R.; Lami, A.; Bloino, J.; Barone, V. *J. Chem. Phys.* **2007**, *126*, 084509.

(26) Santoro, F.; Improta, R.; Lami, A.; Bloino, J.; Barone, V. *J. Chem. Phys.* **2007**, *126*, 184102.

(27) Jankowiak, H.-C.; Stuber, J. L.; Berger, R. *J. Chem. Phys.* **2007**, *127*, 234101.

(28) Eckart, C. *Phys. Rev.* **1935**, *47*, 552.

(29) Duschinsky, F. *Acta Physicochim. URSS* **1937**, *7*, 551.

(30) Sharp, T. E.; Rosenstock, H. M. *J. Chem. Phys.* **1964**, *41*, 3453.

(31) Chang, J.-L. *J. Chem. Phys.* **2008**, *128*, 174111.

(32) Manneback, C. *Physica* **1951**, *17*, 1001–1010.

(33) Doktorov, E. V.; Malkin, I. A.; Man'ko, V. I. *J. Mol. Spectrosc.* **1977**, *64*, 302.

(34) Ruhoff, P. T. *Chem. Phys.* **1994**, *186*, 355.

(35) Johnson, P. M.; Xu, H.; Sears, T. J. *J. Chem. Phys.* **2006**, *125*, 164330.

(36) Hazra, A.; Nooijen, M. *Int. J. Quantum Chem.* **2003**, *95*, 643.

(37) Santoro, F.; Improta, R.; Lami, A.; Bloino, J.; Barone, V. *J. Chem. Phys.* **2008**, *128*, 224311.

(38) Bloino, J.; Biczysko, M.; Crescenzi, O.; Barone, V. *J. Chem. Phys.* **2008**, *128*, 244105.

(39) Barone, V. *J. Chem. Phys.* **2005**, *122*, 014108.

(40) Becke, D. *J. Chem. Phys.* **1993**, *98*, 5648.

(41) Moeller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

(42) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

(43) Kendall, R.; Dunning Jr., T.; Harrison, R. *J. Chem. Phys.* **1992**, *96*, 6769.

(44) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.

(45) Barone, V.; Cimino, P. *Chem. Phys. Lett.* **2008**, *454*, 139.

(46) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669.

(47) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419.

(48) Rappè, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A., III.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024.

(49) Hampel, C.; Peterson, K.; Werner, H.-J. *Chem. Phys. Lett.* **1992**, *190*, 1.

(50) Deegan, M. J. O.; Knowles, P. J. *Chem. Phys. Lett.* **1994**, *227*, 321.

(51) Korona, T.; Werner, H.-J. *J. Chem. Phys.* **2003**, *118*, 300.

(52) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Rauhut, G.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Hampel, C.; Hetzer, G.; Lloyd, A. W.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pitzer, R.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T. *MOLPRO*, version 2006.1; http://www.molpro.net.

(53) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; 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.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian Development Version*, revision G.01; Gaussian, Inc.: Wallingford, CT, 2006.

(54) Eisenhardt, C. G.; Pietraperzia, G.; Becucci, M. *Phys. Chem. Chem. Phys.* **2001**, *3*, 1407.

(55) Hoffmann, L. J. H.; Marquardt, S.; Gemechu, A. S.; Baumgärtel, H. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2360.

(56) Bopp, J. C.; Roscioli, J. R.; Johnson, M. A.; Miller, T. M.; Viggiano, A. A.; Villano, S. M.; Wren, S. W.; Lineberger, W. C. *J. Phys. Chem. A* **2007**, *111*, 1214.

(57) Borrelli, R. *Chem. Phys. Lett.* **2007**, *445*, 84.

(58) Grimsrud, E. P.; Chowdhury, S.; Kebarle, P. *J. Chem. Phys.* **1985**, *83*, 1059.

(59) Vassiliev, S.; Bruce, D. *Photosynth. Res.* **2008**, *97*, 75.

(60) Noguchi, T. *Plant Cell Physiol.* **2002**, *43*, 1112.

(61) Brancato, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2006**, *125*, 164515.

(62) Trofimov, A. B.; Schirmer, J.; Kobychev, V. B.; Potts, A. W.; Holland, D. M. P.; Karlsson, L. *J. Phys. B* **2006**, *39*, 305.

CT8004744