Rationalising the vibrational spectra of biomolecules using atomistic simulations

David M. Benoit

Nachwuchsgruppe Theorie – SFB 569, Albert-Einstein-Allee 11, University of Ulm, D-89081 Ulm, Germany

TABLE OF CONTENTS

1. ABSTRACT

This review presents an account of the recent developments in the use of atomistic simulations to predict vibrational spectra of biomolecules. I give an overview of the concepts used in the various theoretical vibrational models and discuss their relative merits and weaknesses. The issue of anharmonicity in vibrational dynamics is examined in particular detail, owing to its crucial influence on simulations of vibrational spectra for flexible biosystems. The performance of each technique is illustrated by typical applications that show the latest work in this field on systems of biological interest, with a particular focus on the vibrational dynamics of the amide group in peptides.

2. INTRODUCTION

Molecular vibrations open a window on the intricate oscillatory dynamics caused by the subtle relationship between inter- and intra-molecular forces in biological molecules. Indeed, vibrational spectra not only provide information on the nature of the chemical groups present in a biomolecule, but are also sensitive to the strength of the various chemical bonds occurring between atoms and to the molecular conformation.

There have been considerable efforts in the last decade to record high-resolution spectra of biologically relevant molecules both in the gas phase and in loosely solvated environments[1–4]. The measured vibrational

spectra are a mine of information on the actual dynamics of the biomolecule observed. However, due to the complexity of the measured vibrational spectra arising from the large number of close-lying transitions, theoretical models are often necessary to understand and rationalise the observed data.

Parallel to these developments, advances in the field of theoretical vibrational spectroscopy over a similar period have allowed a more accurate interpretation of vibrational spectra of biomolecules and provided an understanding of their structural and dynamic features. Such insights have only been possible by moving away from assignments based on empirical group frequencies and using accurate atomistic simulation methods such as quantum-chemical calculations and molecular-dynamics simulations.

The application of atomistic simulations to compute the vibrational spectra of biological molecules requires an input from two different fields of computational chemistry. On the one hand, there can be no simulation without an accurate description of the forces at play between each of the atoms that make up the biomolecule. These forces shape the energetic landscape of the biosystem, and the roughness of the energetic landscape, or potential energy surface (PES), dictates the ease with which the molecule can vibrate in a given direction. On the other hand, the actual simulation of the vibrational motion of a biomolecule requires a physical model of vibrations, be it a classical model akin to a set of balls and springs, or a quantum model of nuclear motion.

In this review, I will concentrate on the recent advances in theoretical vibrational spectroscopy with a particular focus on biomolecular applications, since this is an area that is rarely investigated in mainstream computational chemistry of large systems. However, I will begin by presenting a very brief perspective on some of the challenges involved in the calculation of PES for biomolecules.

2.1. Potential energy surfaces for biomolecules

Due to the size of the systems usually studied in atomistic simulations of biological matter, the most common approach used to describe the PES of the biomolecules studied is to describe the atom–atom interaction potential using a classical force field. These force fields neglect the electronic degrees of freedom and use a set of fitted parameters to describe the various stretching, bending and out-of-plane terms of the intra-molecular part of the PES and the non-bonded interaction terms of the inter-molecular part of the PES. There are a number of very successful atomic force fields that are routinely used to perform molecular-dynamics simulations of biological systems, such as AMBER[5] or CHARMM[6]. For very large systems where a full atomistic description is not possible, coarser-grained models can be used, for example Scheraga's united residue model (UNRES)[7].

In recent years, however, the development of efficient quantum-mechanical methods such as density-fitting density functional theory (df-DFT), local *ab initio* methods (LMP2, LCCSD(T))[8], divide-and-conquer schemes[9, 10] and linear-scaling semi-empirical methods[11] has led to an increasing number of studies that use a quantum-mechanical description of the atom–atom interactions in biomolecules. This trend towards taking explicit account of the electronic degrees of freedom of the system has led to a higher level of accuracy for the PES and has been nurtured by the constant advances in the field of computer science offering ever increasing computational power. Nonetheless, a high-level quantum-mechanical description is still confined to medium-sized biomolecules (about 500–1000 atoms). In order to bridge the system-size gap between the force-field approach and a quantum-chemical treatment, Warshel and Karplus[12] first introduced in 1972 a mixed quantum-mechanical/molecular-mechanics (QM/MM) method, which has been extended over the years to allow a reliable investigation of biosystems. In the QM/MM approach, one part of the biomolecule (the catalytic active site or any other region of interest) is described using quantum-mechanical electronic-structure theory while the rest of the biosystem is efficiently described using a classical force field. For example, Cui and Karplus[13] have implemented an efficient method to compute vibrational frequencies of the active site of myoglobin within the QM/MM framework.

Simulation of the vibrational spectrum of a biomolecule offers a particularly attractive way of assessing the quality of the computed PES. Indeed, vibrational frequencies are very sensitive to the shape of the PES around the minimum region and in particular to its curvature, which is related to the strength of the chemical bond. This fact is used experimentally to provide, for example, an empirical measure of hydrogen-bond strength by observing the relative shift of the OH-stretch frequency in the complex compared to that of the free molecule[14]. Assuming an exact vibrational calculation for a given PES, the availability of experimental vibrational spectra of biomolecules enables a one-to-one comparison with the predicted values and any deviation can thus be traced back to imperfections of the PES used. For this reason, vibrational calculations are often used to benchmark or help in adjusting the parameters of potential energy surfaces. For example, Bowman and Xantheas[15] have used accurate vibrational calculations to adjust a model potential for the $Cl^- H_2O$ complex and showed that the PES obtained is of equivalent quality to a PES derived from high-level *ab initio* MP4 calculations.

Therefore, the purpose of computing accurate vibrational spectra of biomolecules using atomistic simulations is twofold: Firstly, the predictions help disentangle the observed spectra and, secondly, the comparison between simulated and measured spectra provides a powerful tool to evaluate the quality of the computed PES for the biosystem studied.

3. HARMONIC APPROXIMATION FOR LARGE MOLECULAR SYSTEMS

The approximation of harmonic motion is conceptually simple and offers a convenient framework to describe the small-amplitude vibrations of molecules. Due to this simplicity, the harmonic model is often the first step of any vibrational treatment of molecular vibrations and, in the case of large biomolecules, the most used method for routine rationalisation of the observed IR and Raman spectra.

3.1. Harmonic model

The basic assumption of this model is that the potential energy surface, $V(\mathbf{q})$, can be accurately described by a second-order expansion around the energy minimum (equilibrium structure, denoted by **0**), that is

$$V(\mathbf{q}) = V(\mathbf{0}) + (\nabla V)_{\mathbf{0}} \cdot \mathbf{q} + \frac{1}{2} \left[{}^t\mathbf{q} \cdot (\nabla^2 V)_{\mathbf{0}} \cdot \mathbf{q} \right] \quad (1)$$

where $\mathbf{q}$ is a displacement vector from the point **0**. For most applications, the structure is at a stationary point and therefore $(\nabla V)_{\mathbf{0}} = 0$; note however that this is not a requirement of the model (see, for example, the instantaneous normal-mode analysis[16, 17] or derived formalisms[18], which are often performed at $(\nabla V)_{\mathbf{0}} \neq 0$). The remaining second-order derivative of the potential energy surface is known as the force-constant matrix or Hessian matrix. The vibrational frequencies for the resulting quadratic PES can then be obtained from the classical equation of motion, or by solving the nuclear Schrödinger equation analytically. The solutions of the classical equations of motion are usually obtained by diagonalisation of the mass-weighted Hessian matrix, leading to a series of eigenvalues (the fundamental vibrational frequencies) and corresponding eigenvectors, which are also called normal modes of vibration[19, 20]. Alternatively, the analytical solutions of the nuclear Schrödinger equation for the same potential lead to the harmonic vibrational-energy levels and corresponding harmonic vibrational wave functions.

The main difficulty involved in these calculations revolves around the evaluation of the Hessian matrix and its diagonalisation. For a system containing N atoms, the Hessian matrix requires $3N \times 3N$ second derivatives, which can either be evaluated by numerical differences[21] or differentiation of energy gradients[22, 23], or computed analytically[24]. However, since the computational scaling of a Hessian matrix calculation is at least N^2, it can quickly become very time consuming to compute the full Hessian matrix for large biomolecules.

In this case, the cost of computing a large number of numerical steps ($3N$ at least) has to be weighed against the computational effort needed to perform an analytical calculation of the Hessian matrix. For classical force fields the analytical route to the Hessian matrix is the most straightforward since all potential energy terms have an analytical expression. However, for a given *ab initio* or semi-empirical approach, the computation of analytical second derivatives is not always simple. Recent advances

in the coupled–perturbed Hartree–Fock (CPHF) methodology[25] enable more efficient ways of computing analytical Hessians for Hartree–Fock-based methods (all traditional semi-empirical methods, simple HF and some DFT implementations). Some *ab initio* packages also offer efficient implementations of second derivatives for MP2 (for example in CADPAC[26] or Gaussian03[27]), but the vast majority of *ab initio* programs rely on numerical-difference techniques to obtain the Hessian matrix due to their general applicability and straightforward implementation[22, 23]. Moreover, the recent advent of computational clustering and the efficient realisation of persistent computational grids[28] have led to a series of successful parallel approaches to the numerical computation of the Hessian matrix for large systems[29], exploiting the fact that each energy point needed for the numerical derivative calculation is independent of the other and thus very well suited to a distributed approach.

The diagonalisation of large Hessian matrices, which is necessary in order to obtain the harmonic fundamental frequencies and normal modes of a biosystem, can be a considerable computational burden. Indeed, in spite of being symmetrical, the size of such matrices for systems of tens of thousands of atoms often requires specially designed algorithms, such as iterative diagonalisation[30, 31] or a divide-and-conquer approach[32, 33], to cope with the size of the problem. Consequently, a number of schemes have been derived to either bypass the computation of the full Hessian matrix using subspace iterations[34] and/or normal-mode selection[34, 35], or compute only part of the Hessian matrix[36, 37]. In practice, the computation of Hessian matrices for moderately sized molecules is now commonplace for most quantum-chemistry programs.

3.2. Scaling procedures

From an empirical point of view, the harmonic model often leads to a large discrepancy between the predicted and observed fundamental vibrational frequencies (sometimes over $200\ \text{cm}^{-1}$ for a single frequency). Moreover, the degree of agreement varies greatly with the nature of the underlying PES used, whether computed using high-level *ab initio* methods or a classical force field.

One marked difference between the harmonic model and the real vibrational transition is the asymmetry of the underlying potential energy surface. For example, a simple diatomic molecule (as shown in Figure 1) is reasonably described by a symmetric harmonic potential around its energy minimum, but the approximation worsens as the molecule deviates from its minimum geometry. An analysis of the potential energy curve for such a diatomic molecule, which is well described by a Morse oscillator[38], leads to the following conclusions. The vibrational-energy levels predicted by a harmonic approximation to the PES are equally spaced by $\hbar\nu$, while in a real system (or for a Morse oscillator) the energy-level separation decreases as the vibrational quantum number increases, until the dissociation limit is reached. Moreover, the harmonic model always leads to symmetric wave
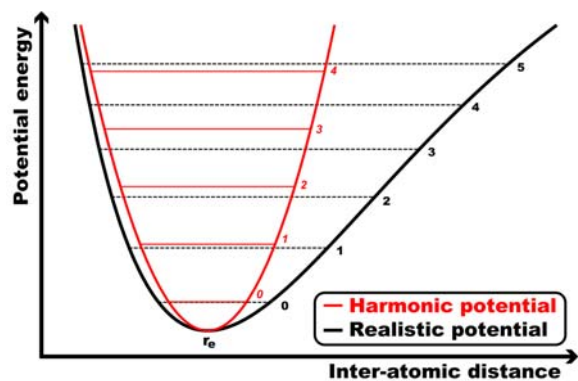
Figure 1. Harmonic and realistic potential energy curve around the energy minimum (r_e) of a diatomic molecule. Note that in this particular case the harmonic vibrational-energy levels (horizontal red lines, numbered in italics from 0 to 4) overestimate the realistic vibrational-energy levels (dashed black lines, numbered from 0 to 5).

functions, which do not always correspond to the real wave functions. The separation between the first two energy levels of a Morse oscillator is given by:

$$\nu_{1 \leftarrow 0}^{\text{Morse}} = \omega_e - 2\omega_e x_e \qquad (2)$$

where $\omega_e x_e$ is the anharmonicity constant, and is in most cases a positive number. We see from expression (2) that the fundamental vibrational frequencies predicted by the more realistic Morse model are in general lower than the harmonic prediction ω_e. This observation has led to the introduction of scaling factors in order to improve the agreement between predicted harmonic frequencies and the observed fundamental value. These scaling factors vary according to the level of theory used to compute the PES and range typically from 0.8 to 1.0 (see Reference[39] for a comprehensive list). Note that this analysis does not constitute a theoretical justification for the use of scaling factors but merely tries to give an understanding of the difference between harmonic and anharmonic treatments of vibrations. This approach, also called the global scaling factor method, has the advantage of being very straightforward, albeit requiring a pre-existing calibration for the level of theory used, and is currently applied in the G2 and G3 thermochemistry models to account for anharmonicity in the vibrational partition function[40, 41].

However, we need to be aware that this scaling analysis is based on the potential energy curve of a generic diatomic molecule, which has mostly positive $\omega_e x_e$. Negative anharmonicity constants can occur in polyatomic molecules, particularly for vibrational modes that describe non-dissociative motion, such as bending vibrations. In this case, the "real" vibrational frequency (or $1 \leftarrow 0$ level separation) is likely to be greater than the harmonic results and thus invalidates the global scaling hypothesis. In general, a careful investigation of the nature of the vibrational motion is therefore needed and the scaling factor needs to be adapted accordingly.

One of the main approximations of the global scaling factor technique is the use of a common degree of anharmonicity for all oscillators in a polyatomic molecule. This approximation becomes more severe as the size of the molecule increases. Indeed, a single global scaling factor applied to the computed harmonic frequencies does not always lead to the best agreement for all types of frequency, and several authors have suggested the use of multiple scaling factors to remedy this problem[42]. The most successful method to date is the scaled quantum-mechanical (SQM) force-field technique suggested by Pulay *et al.*[43, 44]. In this technique, the elements of the force-constant matrix F (or mass-weighted Hessian matrix), rather than the final frequencies, are scaled using:

$$F_{ij}^{\text{scaled}} = \sqrt{\lambda_i \lambda_j} F_{ij} \qquad (3)$$

where λ_i and λ_j are the scaling factors for each coordinate. However, instead of having a single scaling factor for all coordinates or a separate factor for each coordinate, the coordinates that correspond to similar types of vibration are scaled jointly. Rauhut and Pulay report that about 10 different factors give enough flexibility to correct Hartree–Fock harmonic spectra for most organic molecules[44].

Lin *et al.*[45] have taken a step in the opposite direction and used experimental harmonic frequencies to adjust computed DFT harmonic frequencies. They have developed an empirical exchange and correlation functional, named EDF2, that attempts to reproduce the correct curvature of the potential energy surface around the minimum. They perform a multi-functional fitting of the computed harmonic frequencies calculated using both local and gradient-corrected (GGA) exchange and correlation functionals. The procedure yields a re-parametrised functional that leads to computed harmonic frequencies close to those obtained using high-level *ab initio* methods, such as coupled cluster (CCSD(T)), but at a much lower computational cost.

4. BEYOND THE HARMONIC APPROXIMATION

While there is no doubt that the harmonic approximation is a very valuable tool to understand the vibrational spectra of biomolecules, deviations from the harmonic model occur frequently due to the size and intrinsic complexity of biosystems. Several authors[46–48] have highlighted the importance of anharmonicity in biomolecules using techniques that go beyond the harmonic approximation.

4.1. Molecular-dynamics simulations

A direct way of going beyond the harmonic approximation is to extract the vibrational spectrum from molecular-dynamics (MD) simulations. During such simulations, the motion of each atom at a given temperature is integrated using Newton's laws, leading to a simulation of the time evolution of the system, and thus particle–particle correlation functions can be easily recorded. The fluctuation–dissipation theorem (or Green–Kubo relations) connects auto-correlation functions to macroscopic observables such as diffusion coefficients or vibrational

spectra. Two correlation functions are of interest for the simulation of vibrational spectra, namely the velocity auto-correlation function (VACF) and the electric dipole auto-correlation function (DACF). It is easy to understand that a Fourier transform of the DACF is related to the vibrational spectrum since the selection rules of infrared absorption are linked to electric dipole-moment fluctuations. The VACF is also connected to vibrational motion, and a Taylor expansion shows that this correlation function can be expressed as:

$$\langle \mathbf{v}_i(t) \cdot \mathbf{v}_i(0) \rangle = \langle v_i^2 \rangle \left(1 - \frac{1}{2}\omega_i^2 t^2 + \dots \right) \quad (4)$$

where $\langle \dots \rangle$ is an ensemble average over the MD simulation, $\mathbf{v}_i(t)$ is the velocity of particle i at time t, and ω_i is the vibrational frequency associated with particle i oscillating in the mean force field of the other particles. The main advantage of this approach is that the MD protocol can be applied to a very large number of systems, using either classical force fields or *ab initio* potentials, and offers a direct approach to simulate solvent effects for vibrational properties. Moreover, since no assumption is made about the harmonic nature of the PES, the vibrational spectra computed using MD correlation functions are intrinsically anharmonic and can also be computed at different system temperatures. Since the time-correlation method relies on signal-processing techniques, there are some constraints on the nature of the MD trajectory. The length of the simulation has a direct influence on the spectral resolution of the obtained spectrum and a 30 ps-long trajectory is usually necessary to obtain a 1 cm^{-1} resolution[49]. Additionally, due to the nature of the procedure generally used in MD to integrate Newton's equations of motion, the size of the integration time step can cause an overestimation of the vibrational frequencies obtained with the Fourier transform methods. For example, Schmitz and Tavan estimated that a time step of 0.5 fs can lead to a shift close to 12 cm^{-1} for C–H stretches[50].

There are also drawbacks to the time-correlation approach, mainly stemming from the statistical nature of this vibrational treatment. The vibrational wave function is not available, and vibrational overtones and combination bands cannot be simulated using this approach; moreover normal modes of vibration are absent. To remedy this issue, another method associated with MD simulations is the instantaneous normal mode (INM) approach of Seeley and Keyes[16]. With this technique, a harmonic analysis is performed at regular intervals in the course of a molecular-dynamics simulation of the time evolution of the biosystem. This leads to a set of harmonic frequencies and normal modes for each selected time step that are then used to produce a time-averaged spectrum of the system. With this approach, a normal-mode description of the vibrations of the biosystem is preserved, but at the expense of the anharmonicity of the vibrational description.

4.2. Perturbative approaches to anharmonicity

Second-order perturbation theory of the vibrational Hamiltonian (VPT2) provides a convenient analytical formalism that can be used to obtain approximate

anharmonic solutions to the vibrational problem. With this approach, higher-order terms are added to the second-order expansion of the PES given in equation (1) to obtain:

$$
\begin{aligned}
V(\mathbf{q}) = {} & V(\mathbf{0}) + (\nabla V)_0 \cdot \mathbf{q} + \frac{1}{2!}\left[{}^t\mathbf{q} \cdot (\nabla^2 V)_0 \cdot \mathbf{q}\right] \\
& + \frac{1}{3!}\{[{}^t\mathbf{q} \cdot (\nabla^3 V)_0 \cdot \mathbf{q}] \cdot \mathbf{q}\} + \frac{1}{4!}\{{}^t\mathbf{q} \cdot [{}^t\mathbf{q} \cdot (\nabla^4 V)_0 \cdot \mathbf{q}] \cdot \mathbf{q}\} + \dots
\end{aligned}
\quad (5)
$$

The PES expansion is often limited to fourth-order derivatives to reduce the computational effort of the procedure, although there is no theoretical limit to this expansion scheme. In most cases, the minimum of the PES is shifted so that $V(\mathbf{0}) \equiv 0$, and the analysis is performed for $(\nabla V)_0 = 0$. The vibrational Hamiltonian is then decomposed into a standard harmonic-oscillator Hamiltonian, H_0, and a series of anharmonic contributions to the potential energy term, V_3, V_4, …:

$$
\begin{aligned}
H = {} & H_0 + V_3 + V_4 + \dots \\
= {} & \left\{ -\frac{1}{2}\sum_{j=1}^{N}\frac{\partial^2}{\partial q_j^2} + \frac{1}{2!}\left[{}^t\mathbf{q} \cdot (\nabla^2 V)_0 \cdot \mathbf{q}\right] \right\} \\
& + \left\{ \frac{1}{3!}\{[{}^t\mathbf{q} \cdot (\nabla^3 V)_0 \cdot \mathbf{q}] \cdot \mathbf{q}\} \right\} + \left\{ \frac{1}{4!}\{{}^t\mathbf{q} \cdot [{}^t\mathbf{q} \cdot (\nabla^4 V)_0 \cdot \mathbf{q}] \cdot \mathbf{q}\} \right\}
\end{aligned}
\quad (6)
$$

Taking into account the fact that H_0 can be solved analytically and assuming that the anharmonic contributions are small, it is possible to use standard Rayleigh–Schrödinger perturbation theory to calculate the effects of V_n on the harmonic result (see also References [51, 52]). The vibrational-energy levels corrected to second order for a PES expanded in terms of fourth-order derivatives are given by the approximate formula[51]:

$$E_i = \sum_i \omega_i \left(n_i + \frac{1}{2} \right) + \sum_{i \geq j} \chi_{ij} \left(n_i + \frac{1}{2} \right) \left(n_j + \frac{1}{2} \right) \quad (7)$$

where χ_{ij} is a function of the cubic and quartic force constants and is approximately equal to[52]:

$$\chi_{ij} = \frac{1}{4\omega_i\omega_j}\left(\eta_{iijj} - \sum_k^N \frac{\eta_{iik}\eta_{jjk}}{\omega_k^2} + \sum_k^N \frac{2(\omega_i^2 + \omega_j^2 - \omega_k^2)\eta_{ijk}^2}{[(\omega_i + \omega_j)^2 - \omega_k^2][(\omega_i - \omega_j)^2 - \omega_k^2]} \right) \quad (8)$$

It has been shown experimentally[53] that such second-order perturbation expansion leads to an excellent agreement with the observed anharmonic transitions.

Along with an expression for the vibrational-energy levels, perturbation theory leads to anharmonic corrections for the harmonic vibrational wave function. A knowledge of the corrected wave function enables the calculation of vibrationally-averaged properties such as dipole moment, rotational constants[54, 55], and NMR shifts[54]. The relatively simple expression of the second-order perturbation results has led to the implementation of this method in the Gaussian03 suite of *ab initio* programs. However, the high-order derivatives of the PES are often an obstacle to the use of this approach to anharmonicity corrections. Given that second derivatives of the PES (i.e. the Hessian matrix) can already present certain challenges for large molecules, the computation of third and fourth derivatives is a very time-consuming operation. At present

there are no packages that implement analytical high-order derivatives for *ab initio* electronic-structure theory, and most of the time the derivatives required by the perturbation-theory treatment are computed using numerical differences. Several schemes have been developed to minimise the impact of such calculations (see Reference [52] for a short survey). Moreover, the accuracy of the computed derivatives is sensitive to the size of the numerical step used in the numerical differentiation procedure, and several authors have suggested different recommendations[52].

One of the main drawbacks of the second-order perturbational approach is the occurrence of degenerate or near-degenerate vibrational states in the spectrum of the system studied, which can cause the perturbation expansion to diverge or become singular. This is due to the fact that the terms needed for the second-order correction depend on one over the energy difference between the vibrational states. Unfortunately, as the size of the system increases, the likelihood of accidental degeneracies (leading to Fermi-type resonances) increases and particular care has to be taken to circumvent problems in the perturbation corrections, using, for example, degenerate perturbation theory[51]. Nevertheless, the applications of vibrational perturbation theory in this formalism remain limited to the smaller biological molecules, such as uracil[56] and alanine[57], with the exception of small peptides[58].

A very recent re-formulation of the perturbative approach to anharmonicity by Lin, Gilbert and Gill[52] offers a possible improvement compared to the large computational effort incurred for standard VPT2 calculations. Instead of calculating the energy corrections to the harmonic vibrational-energy levels using perturbation theory, they shift the origin of the harmonic oscillator wave function to account for anharmonicity. The justification of this approach is that the main difference between harmonic and anharmonic vibrational wave functions is that the anharmonic one is usually shifted. With their approach, the value of the wave-function shift is adjusted so that the energy difference between the ground and first excited vibrational states (the fundamental transition frequency) approximately matches the second-order perturbation result of equations (7) and (8). Lin *et al.* observe that, since they do not use second-order perturbation theory, their method does not suffer from the drawbacks of VPT2 with regard to accidental degeneracy. This transformation of the harmonic wave function along with a more modest computational requirement for high-order derivatives of the PES allows an efficient estimation of the VPT2 results and has been implemented in the Q-Chem package[59]. The preliminary tests performed by Lin, Gilbert and Gill on small organic molecules show that this method constitutes an attractive alternative to the canonical second-order perturbation-theory treatment of anharmonicity. Moreover, due to the efficiency of their scheme, this method could be applied successfully to moderately sized biological systems.

4.3. CC-VSCF towards solving the full-dimensional Schrödinger equation

One way of overcoming the shortcomings of the perturbative approach to anharmonicity is to use a variational method to solve the vibrational Hamiltonian. However, the fully coupled Hamiltonian for systems larger than a few atoms is too complex to be solved directly and some approximations have to be used. The vibrational self-consistent field (VSCF) method is a variational alternative to the perturbation-theory approach and is currently the leading method for the approximate computation of anharmonic frequencies of large systems. The idea of using a self-consistent-field approach to solve vibrational problems was first formulated in 1978 by Carney *et al.*[60] and Bowman[61]. The main concept behind the VSCF approach has its origins in the mean-field theory developed for electronic-structure calculations; namely replacing the explicit correlation present in a many-particle system with a series of single-particle problems coupled through an effective potential that is dependent on all the other degrees of freedom.

With this method[61–65], the wave function of the system, Ψ_k, for each vibrational state k is assumed to be separable and is described by a product of single-mode wave functions, ϕ_i^k, one for each normal coordinate Q_i:

$$\Psi_k(Q_1, \ldots, Q_n) = \prod_i^n \phi_i^k(Q_i) \qquad (9)$$

Thus, the vibrational Hamiltonian can be rewritten as a collection of single-mode equations of the form:

$$\left(-\frac{1}{2}\frac{\partial^2}{\partial Q_j^2} + V_{j,k}^{\text{eff.}}(Q_j)\right)\phi_j^k(Q_j) = \epsilon_j^k \phi_j^k(Q_j) \qquad (10)$$

In order to achieve the separation of the vibrational Hamiltonian into single-mode equations, we have introduced an effective potential, $V_{j,k}^{eff}(Q_j)$, for each state k:

$$V_{j,k}^{\text{eff.}}(Q_j) = \left\langle \prod_{l \neq j}^n \phi_l^k(Q_l) \middle| V(Q_1, \ldots, Q_n) \middle| \prod_{l \neq j}^n \phi_l^k(Q_l) \right\rangle \qquad (11)$$

The solutions of each of these one-dimensional equations (one for each mode j) can then be used as basis functions (ϕ_j^k) for the vibrational wave function of the system Ψ_k for each state, as described in (9). Since the effective potential depends implicitly on the wave function, the set of equations needs to be solved iteratively until self-consistency is achieved. The total energy of the system is then given by the sum of the single-mode energies corrected for the double counting caused by the use of an effective potential:

$$E_k = \sum_{j=1}^n \epsilon_j^k - (n-1)\left\langle \prod_{j=1}^n \phi_j^k(Q_j) \middle| V(Q_1, \ldots, Q_n) \middle| \prod_{j=1}^n \phi_j^k(Q_j) \right\rangle \qquad (12)$$

The main computational difficulty in solving equation (10) comes from the evaluation of the multi-dimensional integral, defined in equation (11), which is needed in order to evaluate the mean-field potential. To overcome this difficulty, Jung and Gerber[64] suggested using an *n*-body representation of the potential energy surface:

$$V(\mathbf{Q}) = \sum_{j=1}^{n} V_{j}^{(1)}(Q_j) + \sum_{i=1}^{n}\sum_{j>i}^{n} V_{ij}^{(2)}(Q_i,Q_j) + \sum_{i=1}^{n}\sum_{j>i}^{n}\sum_{k>j}^{n} V_{ijk}^{(3)}(Q_i,Q_j,Q_k) + \cdots \quad (13)$$

Chaban *et al.* have shown that including in (13) only terms up to $V_{ij}^{(2)}(Q_i,Q_j)$, that is, using a pairwise approximation of the PES, is sufficient to give reasonable vibrational frequencies for large molecular systems[66]. This approximation simplifies calculations for large molecular systems, since the VSCF equations now involve at most two-dimensional integrals. These integrals can be computed efficiently by using a grid representation, which requires the values of the one-mode and two-mode representations of the PES and the value of the normal-mode wave function at the grid points.

For large biomolecules, the calculations needed to construct a global potential energy surface *ab initio* often become prohibitive. Therefore, classical force fields, such as AMBER, have been used by Roitberg *et al.* to successfully compute the many-dimensional anharmonic wave function of the bovine pancreatic trypsin inhibitor (BPTI) using VSCF[47]. However, further studies on glycine have shown that the classical force fields commonly used in biosimulations are not sufficiently accurate to provide a reliable potential energy surface for anharmonic corrections and that *ab initio*-derived PES should be used instead[67].

Unfortunately, the fitting of multi-dimensional functions to a set of *ab initio* points to obtain an analytical PES cannot always be guaranteed to converge for large systems. An alternative to the use of an analytical PES is the direct method that was developed by Chaban *et al.*[66] and implemented, for example, in the Gamess-us[68] electronic package. With this approach, the PES needed for the VSCF calculation is constructed progressively by performing a series of *ab initio* energy calculations for a set of regularly spaced grid points along the normal-mode coordinates. This approach has shown to be accurate[66], and can thus be used to benchmark the accuracy of a potential energy surface, obtained at a given level of *ab initio* theory, through the computation of the vibrational spectrum for a given molecule. However, the direct calculation of a PES for large molecular systems requires *ab initio* calculations at a very large number of grid points and can be computationally demanding for molecular systems with a large number of normal modes, such as biomolecules. In order to accelerate the computation of the PES, Benoit[65, 69] suggested reducing systematically the number of mode–mode couplings included in the PES, thus leading to faster VSCF calculations (Fast-VSCF method). The direct method does not require the computation of a global *n*-dimensional PES or any fitting to an analytical

function, which makes it very suitable for large molecular systems.

A simple VSCF scheme does not explicitly take into account mode–mode coupling, since each mode is treated independently in the effective field of the other modes. A simple way of correcting this lack of explicit correlation between modes in the VSCF approach is to apply a perturbation correction to the SCF solutions. This method, called CC-VSCF (for correlation-corrected VSCF), has been extensively described by Norris *et al.*[70] and, assuming that the effect of the non-separability of modes is small, uses the following perturbation potential:

$$V_{k}^{\text{pert.}}(Q_1,\ldots,Q_n) = V(Q_1,\ldots,Q_n) - \sum_{j=1}^{n} V_{j,k}^{\text{eff.}}(Q_j) \quad (14)$$

In this expression, the perturbation potential is formulated as the difference between the true potential energy of the system and the effective VSCF potential. The energy correction is then given by an expectation value of the perturbing potential, in a similar fashion to the usual Moller–Plesset perturbation theory:

$$E_{k}^{\text{corr.}} = \sum_{s \neq t}^{n} \frac{\langle \Psi_s | V_{k}^{\text{pert.}} | \Psi_t \rangle \langle \Psi_t | V_{k}^{\text{pert.}} | \Psi_s \rangle}{\tilde{E}_{s}^{(0)} - \tilde{E}_{t}^{(0)}} \quad (15)$$

where $\tilde{E}_{s}^{(0)}$ and $\tilde{E}_{t}^{(0)}$ are the sum of the single-mode energies (ε_j) for the states *s* and *t* of the unperturbed VSCF solutions. Matsunaga *et al.*[71] developed a perturbation theory that can handle degenerate cases in a satisfactory manner, named DPT2-VSCF (for VSCF including second-order energy corrections obtained through degenerate perturbation theory). The main advantage of the perturbative correction scheme is its low computational cost, which makes it suitable for large molecular systems. However, this scheme, too, can lead to non-physical corrections, as reported by Chaban[66] and Christiansen[72].

5. APPLICATIONS

The various techniques presented earlier have several applications in the domain of biologically relevant molecules. In the following, I will focus on some particular applications that illustrate the capabilities of these methods.

In the field of biosystems containing transition metals, Kozlowski *et al.*[73] have used the SQM technique to investigate the vibrational spectra of various porphyrins. Using a PES computed at the B3LYP/6-31G(d) level of theory, they show that this approach is able to reproduce the observed experimental fundamental frequencies accurately, with a mean absolute deviation smaller than 5 cm^{-1}. Moreover, their study was able to dispel doubts concerning the attribution of some observed frequencies of the free-base porphine and thus demonstrated that it had D_{2h} symmetry. They also extended their study to the spectroscopy of metalloporphyrins, investigating the Zn-, Mg-[74], and Ni-porphyrins[75]. For the first two, they

report an excellent agreement with the experimental data and an average deviation in the range of 4–7 cm^{-1}. Their SQM analysis provides further evidence of the planarity of these two prophyrins. Their investigation of the Ni-porphyrin allowed the reassignment of several vibrational bands and shows that an out-of-plane distortion of the porphyrin cycle, while lowering the symmetry of the complex, has only a limited effect on the vibrational spectrum.

Large-amplitude motions of protein systems are of great interest since they are related to the overall protein flexibility, conformational preferences of the peptide chain and protein–protein and protein–ligand interactions. In order to investigate the flexibility and conformational state of bacteriorhodopsin, Whitmire *et al.* used the normal modes obtained from a harmonic frequency calculation and compared the results to experimental far-infrared spectra[76]. They used the CHARMM classical force field and an iterative diagonalisation of the computed Hessian matrix to obtain the low-frequency part of the vibrational spectra of the wild-type and D96N mutant bacteriorhodopsin. They observed a good agreement between the harmonic predictions and the measured spectra but remarked that the different absorption patterns obtained for wild-type and D96N mutant rhodopsin are not reproduced by the simple harmonic analysis. They suggest that anharmonicity might account for the different FIR spectra of the two types of protein and show that the low-frequency absorption profile indicated a greater flexibility of the wild-type compared to the mutant rhodopsin.

Protein flexibility also plays a role in the mechanism of binding-site recognition by ligands. Sanejouand[77] performed a harmonic analysis of the low-frequency modes of human CD4, a receptor for the human immunodeficiency virus (HIV). He showed the existence of a hinge-bending mechanism for the two N-terminal domains of CD4, which influences the extent to which it binds to Gp120, an envelope protein of HIV. Using the CHARMM force field to describe the potential energy surface of the protein, he observed that the low-frequency modes of CD4 describe a concerted motion of two sizeable domains about a hinge axis. The link between this bending mechanism and the binding activity is also consistent with the experimental evidence of Gp120 affinity obtained from mutant CD4 and monoclonal-antibody studies.

The amide region (amide bands) of the vibrational spectra of proteins provides a wealth of information on their secondary structure. Accurate modelling of this group of vibrational frequencies remains relatively complex due to the usual size of biosystems, the number of vibrational resonances between the vibrating groups, and the presence of hydrogen-bonding interactions creating the secondary structure. The amide-I band originates mainly from vibrations of the carbonyl groups on the protein backbone, and thus a partial Hessian approach, such as the one used recently by Besley and Metcalf[37], can be used to reduce the computational effort. Their study on a set of model peptides shows that this technique leads to only a small deviation (10 – 20 cm^{-1}) from the full Hessian calculation

as long as the amide carbon, oxygen and nitrogen atoms are included. For agitoxin, a protein containing 38 residues, they observe a good qualitative agreement between their harmonic calculations[37, 78] and the amide signal observed experimentally.

The anharmonicity of the various amide bands has been investigated in detail by Wang and Hochstrasser[58], using second-order vibrational perturbation theory (VPT2) and a B3LYP/6-31+G(d,p) PES. In their study of small peptides, they observe that the amide-A band (originating from NH-stretch vibrations) possesses a significant single-mode anharmonicity, about 10 times that of the amide-I band, and that a realistic description of the experimental frequencies has to include anharmonic corrections. They also show that the anharmonic frequencies predicted by VPT2 are in very reasonable agreement with the experimental gas-phase observations and note that subtle vibrational shifts due to hydrogen bonding are also well reproduced by this method. They demonstrate, by systematically computing the anharmonic coupling between the different types of amide modes in each peptide, that the position of the amide-A band is fairly insensitive to mode–mode coupling, while that of the amide-I band is strongly affected by the presence of low-frequency amide vibrations and peptide backbone motion. Kaledin and Bowman[79] have performed a full-dimensional anharmonic calculation of the vibration of N-methyl acetamide, a model compound for the amide-band vibration, using a correlated VSCF approach. They show that it is possible to obtain a reliable theoretical description of the amide-I, -II and -II bands by including mode–mode coupling for a PES computed at the MP2/aug-cc-pVTZ level of *ab initio* theory. Gregurick *et al.*[80] have also used a correlated VSCF approach, with a PES computed using the AMBER force field, to investigate the vibrational spectra of micro-solvated dipeptides. Their study shows that the predicted anharmonic frequencies are usually higher (blue-shifted) than those obtained by a harmonic analysis. The anharmonic terms in the PES tend to rigidify the system, and the frequency shifts are shown to be as high as 300 cm^{-1} for some modes. Moreover, Gregurick *et al.* managed to relate the vibrational frequencies to conformational changes in the peptide–water complex, highlighting the effects of hydrogen bonding on the amide-band vibrations. In the second part of their study, they use a non-solvated tri-alanine peptide in an anti-parallel beta-sheet conformation to test the accuracy of the force field used. They note that the anharmonic frequencies are in good agreement with the observed experimental spectra for most amide bands, often improving over the harmonic frequencies obtained from a purpose-built empirical force field. However, they observe discrepancies for some low-intensity amide modes (I, III and V) and suggest that the AMBER force field should be partially revised to account for these defects.

Roitberg *et al.* have performed the largest anharmonic calculations on a biological system to date using the VSCF method. In two landmark papers[47, 81] they compute the anharmonic vibrational wave function for a number of vibrational modes of BPTI, using various

approximations, in order to assess the influence of anharmonicity on protein properties. For the classical potential used (MOIL package[82]), they observe that single-mode anharmonicity has a stronger effect on the low-frequency modes, with some frequencies changing by a factor of four, and renders the modes stiffer. Moreover, mode–mode anharmonicity affects these frequencies even further: Roitberg *et al.* suggest that the mixed fourth-order derivatives of the PES, rather than the third-order derivatives, are the main contributors to these corrections. They also highlight that anharmonicity plays an important role in the calculation of the atomic mean square displacement, and thus of the Debye–Waller factor. They observe substantial deviations for low-frequency torsional motions, mainly due to the stiffening of the system caused by the anharmonicity corrections.

One of the first analyses of anharmonicity in peptides was performed by Levy *et al.*[46] using classical molecular dynamics. They compute the mean square fluctuations of decaglycine in its alpha-helix conformation for several temperatures using the CHARMM force field. They observe that the fluctuations predicted at room temperature by a harmonic model are smaller than the data obtained from X-ray analyses by at least a factor of two. The results obtained from full molecular-dynamics simulations lead to a much better agreement with experiment, thus indicating that anharmonicity plays an important role in the dynamics of proteins. Recently, Grégoire *et al.*[83] have used Car–Parrinello molecular-dynamics simulations with a BLYP/PW PES to compute the vibrational spectra of protonated peptides in the gas phase. They highlight that this type of approach enables the simultaneous exploration of the various peptide conformations accessible by the system at room temperature. The method can therefore account for the presence of several different contributions to the observed vibrational spectrum that arise from a number of low-lying minima. In the case of protonated peptides, where proton migration can occur rapidly and has a strong influence on the overall molecular structure, they show that a molecular-dynamics approach leads to a good agreement between predicted and observed vibrational bands, but that the transition intensities can be overestimated. They note that accounting for both anharmonic effects and entropic contributions to the structural equilibrium are crucial to an accurate simulation of these flexible biological systems.

6. PERSPECTIVES

The use of atomistic simulations to predict and understand the vibrational dynamics of biological systems is still an area in expansion. The insight gained from accurate theoretical modelling leads to a deeper understanding of the structural and dynamical interplay in many important areas of biochemistry and biophysics. The techniques presented in this review demonstrate that it is possible nowadays to simulate vibrational spectra of sizeable systems of biological interest with an accuracy close to 10 cm^{-1} for fundamental bands. To obtain such agreement, the simple harmonic model has to be complemented with a more accurate description of

vibrational dynamics accounting for single-mode anharmonicity and mode–mode coupling. However, such refinement of the vibrational treatment comes at a price: the complexity of the anharmonic models renders them very demanding on computational resources. Nevertheless, with the advances in computer technology and the constant efforts to develop more efficient methodologies, the use of atomistic simulations to obtain accurate vibrational spectra is set to become a valuable routine tool in biosciences. Moreover, for the theoretical community, the direct link between potential energy surfaces and experimental data through vibrational spectra offers a unique possibility of assessing the quality of the theoretical interaction models used for systems of biological interest.

7. ACKNOWLEDGMENTS

8. REFERENCES

1. E. G. Robertson and J. P. Simons: Getting into shape: Conformational and supramolecular landscapes in small biomolecules and their hydrated clusters. *Phys. Chem. Chem. Phys*. 3, 1 (2001)

2. I. Hünig and K. Kleinermanns: Conformers of the peptides glycine-tryptophan, tryptophan-glycine and tryptophan-glycine-glycine as revealed by double resonance laser spectroscopy. *Phys. Chem. Chem. Phys*. 6, 2650 (2004)

3. W. Chin, F. Piuzzi, J.-P. Dognon, I. Dimicoli, and M. Mons: Gas-phase models of gamma turns: Effect of side-chain/backbone interactions investigated by IR/UV spectroscopy and quantum chemistry. *J. Chem. Phys*. 123, 084301 (2005)

4. S. R. Mercier, O. V. Boyarkin, A. Kamariotis, M. Guglielmi, I. Tavernelli, M. Cascella, U. Rothlisberger, and T. R. Rizzo: Microsolvation Effects on the Excited-State Dynamics of Protonated Tryptophan. *J. Am. Chem. Soc*. 128, 16938 (2006)

5. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman: A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc*. 117, 5179 (1995)

6. A. D. MacKerell Jr., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus: All-Atom Empirical Potential for Molecular Modeling and

Dynamics Studies of Proteins. *J. Phys. Chem. B* 102, 3586 (1998)

7. A. Liwo, S. Oldziej, M. Pincus, R. Wawak, S. Rackovsky, and H. A. Scheraga: A united-residue force field for off-lattice protein-structure simulations. I. Functional forms and parameters of long-range side-chain interaction potentials from protein crystal data. *J. Comput. Chem.* 18, 849 (1996)

8. J. E. Subotnik, A. Sodt, and M. Head-Gordon: A near linear-scaling smooth local coupled cluster algorithm for electronic structure. *J. Chem. Phys.* 125, 074116 (2006)

9. D. G. Fedorov and K. Kitaura: Theoretical Development of the Fragment Molecular Orbital (FMO) Method. In: Modern Methods for Theoretical Physical Chemistry of Biopolymers. Eds: E. Starikov, J. Lewis, and S. Tanaka, *Elsevier*, (2006)

10. I. Nakanishi, D. G. Fedorov, and K. Kitaura: Molecular recognition mechanism of FK506 binding protein: An all-electron fragment molecular orbital study. *Proteins* 68, 145 (2007)

11. V. M. Anisimov, V. L. Bugaenko, and V. V. Bobrikov: Validation of Linear Scaling Semiempirical LocalSCF Method. *J. Chem. Theory Comput.* 2, 1685 (2006)

12. A. Warshel and M. Karplus: Calculation of ground and excited state potential surfaces of conjugated molecules. I. Formulation and parametrization. *J. Am. Chem. Soc.* 94, 5612 (1972)

13. Q. Cui and M. Karplus: Molecular properties from combined QM/MM methods. I. Analytical second derivative and vibrational calculations. *J. Chem. Phys.* 112, 1133 (2000)

14. R. M. Badger: The Relation Between the Energy of a Hydrogen Bond and the Frequencies of the O-H Bands. *J. Chem. Phys.* 8, 288 (1940)

15. J. M. Bowman and S. S. Xantheas: "Morphing" of ab initio-based interaction potentials to spectroscopic accuracy: Application to $Cl^-(H_2O)$. *Pure Appl. Chem.* 76, 29 (2004)

16. G. Seeley and T. Keyes: Normal-mode analysis of liquid-state dynamics. *J. Chem. Phys.* 91, 5581 (1989)

17. B. C. Xu and R. M. Stratt: Liquid theory for band structure in a liquid. II. p orbitals and phonons. *J. Chem. Phys.* 92, 1923 (1990)

18. A. Kidera: Smart Monte Carlo simulation of a globular protein. *Int. J. Quant. Chem.* 75, 207 (1999)

19. W. D. Gwinn: Normal Coordinates: General Theory, Redundant Coordinates, and General Analysis Using Electronic Computers. *J. Chem. Phys.* 55, 477 (1971)

20. E. B. Wilson Jr., J. C. Decius, and P. C. Cross: Molecular Vibrations, *Dover Publications*, New York (1980)

21. S. F. Boys, G. B. Cook, C. M. Reeves, and I. Shavitt: Automatic Fundamental Calculations of Molecular Structure. *Nature* 178, 1207 (1956)

22. P. Pulay: Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules. I. Theory. *Mol. Phys.* 17, 197 (1969)

23. J. McIver Jr. and A. Komornickiy: Structure of transition states in organic reactions. General theory and an application to the cyclobutene-butadiene isomerization using a semiempirical molecular orbital method. *J. Am. Chem. Soc.* 94, 2625 (1972)

24. D. M. Bishop and M. Randic: Ab Initio Calculation of Harmonic Force Constants. *J. Chem. Phys.* 44, 2480 (1966)

25. Y. Alexeev, M. W. Schmidt, T. L. Windus, and M. S. Gordon: A parallel distributed data CPHF algorithm for analytic Hessians. *J. Comput. Chem.* 28, 1685 (2007)

26. R. D. Amos, I. L. Alberts, J. S. Andrews, S. M. Colwell, N. C. Handy, D. Jayatilaka, P. J. Knowles, R. Kobayashi, K. E. Laidig, G. Laming, A. M. Lee, P. E. Maslen, C. W. Murray, J. E. Rice, E. D. Simandiras, A. J. Stone, M.-D. Su and D. J. Tozer, CADPAC: The Cambridge Analytic Derivatives Package Issue 6, University of Cambridge (1995)

27. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian 03, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004.

28. A. S. P. Gomes, A. Merzky, and L. Visscher: A Framework for Execution of Computational Chemistry Codes in Grid Environments. *Lecture Notes in Computer Science* 3993, 97 (2006)

29. J. Neugebauer, M. Reiher, C. Kind, and B. Hess: Quantum chemical calculation of vibrational spectra of

large molecules - Raman and IR spectra for Buckminsterfullerene. *J. Comput. Chem.* 23, 895 (2002)

30. L. Mouawad and D. Perahia: Diagonalization in a mixed basis: A method to compute low-frequency normal modes for large macromolecules. *Biopolymers* 33, 599 (1993)

31. B. R. Brooks, D. Janezic, and M. Karplus: Harmonic analysis of large systems. I. Methodology. *J. Comput. Chem.* 16, 1522 (1995)

32. P. Durand, G. Trinquier, and Y. H. Sanejouand: A new approach for determining low-frequency normal modes in macromolecules. *Biopolymers* 34, 759 (1994)

33. F. Tama, F. X. Gadea, O. Marques, and Y. H. Sanejouand: Building-block approach for determining low-frequency normal modes of macromolecules. *Proteins* 41, 1 (2000)

34. F. Filippone and M. Parrinello: Vibrational analysis from linear response theory. *Chem. Phys. Lett.* 345, 179 (2001)

35. M. Reiher and J. Neugebauer: A mode-selective quantum chemical method for tracking molecular vibrations applied to functionalized carbon nanotubes. *J. Chem. Phys.* 118, 1634 (2003)

36. J. D. Head: Computation of vibrational frequencies for adsorbates on surfaces. *Int. J. Quant. Chem.* 65, 827 (1997)

37. N. A. Besley and K. A. Metcalf: Computation of the amide I band of polypeptides and proteins using a partial Hessian approach. *J. Chem. Phys.* 126, 035101 (2007)

38. P. Morse: Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels. *Phys. Rev.* 34, 57 (1929)

39. A. P. Scott and L. Radom: Harmonic Vibrational Frequencies: An Evaluation of Hartree-Fock, Moller-Plesset, Quadratic Configuration Interaction, Density Functional Theory, and Semiempirical Scale Factors. *J. Phys. Chem.* 100, 16502 (1996)

40. L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople: Gaussian-2 theory for molecular energies of first- and second-row compounds. *J. Chem. Phys.* 94, 7221 (1991)

41. L. Curtiss, K. Raghavachari, P. Redfern, V. Rassolov, and J. Pople: Gaussian-3 (G3) theory for molecules containing first and second-row atoms. *J. Chem. Phys.* 109, 7764 (1998)

42. M. D. Halls, J. Velkovski, and H. B. Schlegel: Harmonic frequency scaling factors for Hartree-Fock, S-VWN, B-LYP, B3-LYP, B3-PW91 and MP2 with the Sadlej pVTZ electric property basis set. *Theor. Chem. Acc.* 105, 413 (2001)

43. P. Pulay, G. Fogarasi, G. Pongor, J. E. Boggs, and A. Vargha: Combination of theoretical ab initio and experimental information to obtain reliable harmonic force constants. Scaled quantum mechanical (QM) force fields for glyoxal, acrolein, butadiene, formaldehyde, and ethylene. *J. Am. Chem. Soc.* 105, 7037 (1983)

44. G. Rauhut and P. Pulay: Transferable Scaling Factors for Density Functional Derived Vibrational Force Fields. *J. Phys. Chem.* 99, 3093 (1995)

45. C. Y. Lin, M. W. George, and P. M. W. Gill: EDF2: A Density Functional for Predicting Molecular Vibrational Frequencies. *Aust. J. Chem.* 57, 365 (2004)

46. R. M. Levy, D. Perahia, and M. Karplus: Molecular dynamics of an alpha-helical polypeptide: Temperature dependence and deviation from harmonic behavior. *Proc. Nat. Acad. Sci. USA* 79, 1346 (1982)

47. A. Roitberg, R. B. Gerber, R. Elber, and M. A. Ratner: Anharmonic wave functions of proteins: quantum self-consistent field calculations of BPTI. *Science.* 268, 1319 (1995)

48. D. M. Leitner: Vibrational Energy Transfer in Helices. *Phys. Rev. Lett.* 87, 188102 (2001)

49. M. Schmitz and P. Tavan: On the art of computing the IR spectra of molecules in the condensed phase. In: Modern Methods for Theoretical Physical Chemistry of Biopolymers. Eds: E. Starikov, J. Lewis, and S. Tanaka, *Elsevier*, (2006)

50. M. Schmitz and P. Tavan: Vibrational spectra from atomic fluctuations in dynamics simulations. II. Solvent-induced frequency fluctuations at femtosecond time resolution. *J. Chem. Phys.* 121, 12247 (2004)

51. V. Barone: Anharmonic vibrational properties by a fully automated second-order perturbative approach. *J. Chem. Phys.* 122, 014108 (2005)

52. C. Y. Lin, A. T. B. Gilbert, and P. M. W. Gill: Calculating molecular vibrational spectra beyond the harmonic approximation. *Theor. Chem. Acc.* 120, 23 (2008)

53. D. E. Reisner, R. W. Field, J. L. Kinsey, and H.-L. Dai: Stimulated emission spectroscopy: A complete set of vibrational constants for X-tilde 1A1 formaldehyde. *J. Chem. Phys.* 80, 5968 (1984)

54. K. Ruud, P.-O. Astrand, and P. R. Taylor: An efficient approach for calculating vibrational wave functions and zero-point vibrational corrections to molecular properties of polyatomic molecules. *J. Chem. Phys.* 112, 2668 (2000)

55. V. B. M. E. Alikhani: Hydrogen-bonding between the hydrogen peroxide molecule and the hydroperoxy radical (H_2O_2–HO_2): the global minimum. *Chem. Phys. Lett.* 391, 134 (2004)

56. V. Barone, G. Festa, A. Grandi, N. Rega, and N. Sanna: Accurate vibrational spectra of large molecules by density functional computations beyond the harmonic approximation: the case of uracil and 2-thiouracil. *Chem. Phys. Lett.* 388, 279 (2004)

57. P. Danecek, J. Kapitan, V. Baumruk, L. Bednarova, V. Kopecky, and P. Bour: Anharmonic effects in IR, Raman, and Raman optical activity spectra of alanine and proline zwitterions. *J. Chem. Phys.* 126, 224513 (2007)

58. J. Wang and R. M. Hochstrasser: Anharmonicity of Amide Modes. *J. Phys. Chem. B* 110, 3798 (2006)

59. Y. Shao, L. Fusti-Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. Distasio Jr., R. C. Lochan, T. Wang, G. J. O. Beran, N. A. Besley, J. M., Herbert, C. Y. Lin, T. Van Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock III, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer III, J. Kong, A. I. Krylov, P. M. W. Gill, M. Head-Gordon: Advances in methods and algorithms in a modern quantum chemistry program package. *Phys. Chem. Chem. Phys.* 8, 3172 (2006)

60. G. C. Carney, L. L. Sprandel, and C. W. Kern: Variational Approaches to Vibration-Rotation Spectroscopy for Polyatomic Molecules. *Adv. Chem. Phys.* 37, 305 (1978)

61. J. M. Bowman: Self-consistent field energies and wavefunctions for coupled oscillators. *J. Chem. Phys.* 68, 608 (1978)

62. J. M. Bowman, K. M. Christoffel, and F. L. Tobin: Application of SCF-SI theory to vibrational motion in polyatomic molecules. *J. Phys. Chem.* 83, 905 (1979)

63. J. M. Bowman: The self-consistent-field approach to polyatomic vibrations. *Acc. Chem. Res.* 19, 202 (1986)

64. J. O. Jung and R. B. Gerber: Vibrational wave functions and spectroscopy of $(H_2O)n$, n=2,3,4,5: Vibrational self-consistent field with correlation corrections. *J. Chem. Phys.* 105, 10332 (1996)

65. D. M. Benoit: Fast vibrational self-consistent field calculations through a reduced mode–mode coupling scheme. *J. Chem. Phys.* 120, 562 (2004)

66. G. M. Chaban, J. O. Jung, and R. B. Gerber: Ab initio calculation of anharmonic vibrational states of polyatomic

systems: Electronic structure combined with vibrational self-consistent field. *J. Chem. Phys.* 111, 1823 (1999)

67. R. B. Gerber, G. M. Chaban, S. K. Gregurick, and B. Brauer: Vibrational spectroscopy and the development of new force fields for biological molecules. *Biopolymers* 68, 370 (2003)

68. M. Schmidt, K. Baldridge, J. Boatz, S. Elbert, M. Gordon, J. Jensen, S. Koseki, N. Matsunaga, K. Nguyen, S. Su, T. L. Windus, M. Dupuis, J. A. Montgomery: General atomic and molecular electronic structure system. *J. Comput. Chem.* 14, 1347 (1993)

69. D. M. Benoit: Efficient correlation-corrected vibrational self-consistent field computation of OH-stretch frequencies using a low-scaling algorithm. *J. Chem. Phys.* 125, 244110 (2006)

70. L. S. Norris, M. A. Ratner, A. E. Roitberg, and R. B. Gerber: Moller–Plesset perturbation theory applied to vibrational problems. *J. Chem. Phys.* 105, 11261 (1996)

71. N. Matsunaga, G. M. Chaban, and R. B. Gerber: Degenerate perturbation theory corrections for the vibrational self-consistent field approximation: Method and applications. *J. Chem. Phys.* 117, 3541 (2002)

72. O. Christiansen: Moller–Plesset perturbation theory for vibrational wave functions. *J. Chem. Phys.* 119, 573 (2003)

73. P. M. Kozlowski, A. A. Jarzecki, and P. Pulay: Vibrational Assignment and Definite Harmonic Force Field for Porphine. 1. Scaled Quantum Mechanical Results and Comparison with Empirical Force Field. *J. Phys. Chem.* 100, 7007 (1996)

74. A. A. Jarzecki, P. M. Kozlowski, P. Pulay, B.-H. Ye, and X.-Y. Li: Scaled quantum mechanical and experimental vibrational spectra of magnesium and zinc porphyrins. *Spectrochim. Acta, Part A* 53, 1195 (1997)

75. P. M. Kozlowski, T. S. Rush III, A. A. Jarz¸ecki, M. Z. Zgierski, B. Chase, C. Piffat, B.-H. Ye, X.-Y. Li, P. Pulay, and T. G. Spiro: DFT-SQM Force Field for Nickel Porphine: Intrinsic Ruffling. *J. Phys. Chem. A* 103, 1357 (1999)

76. S. E. Whitmire, D. Wolpert, A. G. Markelz, J. R. Hillebrecht, J. Galan, and R. R. Birge: Protein flexibility and conformational state: A comparison of collective vibrational modes of wild-type and D96N bacteriorhodopsin. *Biophys. J.* 85, 1269 (2003)

77. Y. H. Sanejouand: Normal-mode analysis suggests important flexibility between the two N-terminal domains of CD4 and supports the hypothesis of a conformational change in CD4 upon HIV binding. *Protein Eng.* 9, 671 (1996)

78. N. A. Besley: Computing protein infrared spectroscopy with quantum chemistry. *Phil. Trans. R. Soc. A* 365, 2799 (2007)

79. A. L. Kaledin and J. M. Bowman: Full Dimensional Quantum Calculations of Vibrational Energies of N-Methyl Acetamide. *J. Phys. Chem. A* 111, 5593 (2007)

80. S. K. Gregurick, E. Fredj, R. Elber, and R. B. Gerber: Vibrational Spectroscopy of Peptides and Peptide-Water Complexes: Anharmonic Coupled-Mode Calculations. *J. Phys. Chem. B* 101, 8595 (1997)

81. A. Roitberg, R. B. Gerber, and M. A. Ratner: A Vibrational Eigenfunction of a Protein: Anharmonic Coupled-Mode Ground and Fundamental Excited States of BPTI. *J. Phys. Chem. B* 101, 1700 (1997)

82. R. Elber, A. Roitberg, C. Simmerling, R. Goldstein, G. Verkhivker, H. Li, C. Keasar, J. Zang, and A. Ulitsky: MOIL: A program for simulations of macromolecules. *Comp. Phys. Commun.* 91, 159 (1995)

83. G. Grégoire, M. P. Gaigeot, D. C. Marinica, J. Lemaire, J. P. Schermann, and C. Desfrançois: Resonant infrared multiphoton dissociation spectroscopy of gas-phase protonated peptides. Experiments and Car–Parrinello dynamics at 300 K. *Phys. Chem. Chem. Phys.* 9, 3082 (2007)

Abbreviations: PES: potential energy surface; DFT: density functional theory; (L)MP2: (local) Moller-Plesset perturbation theory; (L)CCSD(T): (local) coupled-cluster singles doubles and perturbative triples; QM/MM: quantum mechanics/molecular mechanics; HF: Hartree–Fock; CPHF: coupled–perturbed Hartree–Fock; GGA: generalised gradient approximation; MD: molecular dynamics; VACF: velocity auto-correlation function; DACF: dipole auto-correlation function; INM: instantaneous normal mode; VPT2: Second-order perturbation theory of the vibrational Hamiltonian; NMR: nuclear magnetic resonance; VSCF: vibrational self-consistent field; BPTI: bovine pancreatic trypsin inhibitor; CC-VSCF: correlation-corrected vibrational self-consistent field; DPT2-VSCF: vibrational self-consistent field including second-order energy corrections obtained through degenerate perturbation theory; FIR: far infrared; HIV: human immunodeficiency virus

Key Words Anharmonicity, Vibrational spectrum, vibrational self-consistent field, Potential energy surface, biomolecules, Harmonic approximation, Vibrational perturbation theory, Molecular dynamics, Amide band, Review

Send correspondence to: David M. Benoit, Nachwuchsgruppe Theorie – SFB 569, Albert-Einstein-Allee 11, University of Ulm, D-89081 Ulm, Germany, Tel: 49-731-502-2940 Fax: 49-731-502-2933, E-mail: david.benoit@uni-ulm.de