JCTC Journal of Chemical Theory and Computation

Vibrational Raman Spectra from the Self-Consistent Charge Density Functional Tight Binding Method via Classical Time-Correlation Functions

Steve Kaminski,*,† Michael Gaus,‡ Prasad Phatak,§ David von Stetten,† Marcus Elstner,‡ and Maria Andrea Mroginski*,†

Technische Universität Berlin, Institut für Chemie, Max-Volmer-Laboratorium, Sekr. PC 14, Strasse des 17. Juni 135, D-10623 Berlin, Germany, Universität Karlsruhe, Institut für Theoretische Chemie, Kaiserstrasse 12, D-76131 Karlsruhe, Germany, and Department of Chemistry, Indiana University, 800 E Kirkwood Avenue, Bloomington, Indiana 47405

Received December 9, 2009

Abstract: The Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB) method has been extended for the calculation of vibrational Raman spectra employing the Fourier Transform of Time-Correlation Function (FTTCF) formalism. As Witek and co-workers have already shown for a set of various organic molecules, the minimal basis SCC-DFTB approach performs surprisingly good in terms of polarizability calculations. Therefore, we were encouraged to use this electronic structure method for the purpose of Raman spectra calculations via FTTCF. The molecular polarizability was accessed via second order numeric derivatives of the SCC-DFTB energy with respect to the components of an external electric field "on-the-fly" during a molecular dynamics (MD) simulation. The finite electric field approach delivers Raman spectra that are in overall good agreement for most of 10 small organic model compounds examined in the gas phase compared to a standard Normal Mode Analysis (NMA) approach at the same (SCC-DFTB) and at a higher level of theory (BLYP aug-cc-pVTZ). With the use of reparametrized SCC-DFTB repulsive potentials, a distinct improvement of the Raman spectra from the SCC-DFTB/FTTCF protocol of conjugated hydrocarbons has been observed. Further QM/MM test calculations of L-phenylalanine in aqueous solution revealed larger deviations concerning vibrational frequencies and relative intensities for several stretching and bending modes in the benzene ring as compared to experimental results. Our SCC-DFTB/FTTCF approach was also tested against a hybrid method, in which polarizability calculations at the B3YLP 6-31G(d) level were performed on a trajectory at the SCC-DFTB level. We found that our SCC-DFTB/FTTCF protocol is not only much more efficient but in terms of the resulting Raman spectra also of similar accuracy compared to the hybrid approach. In our opinion, the more accurately calculated polarizabilities at the B3YLP 6-31G(d) level cannot compensate for the usually insufficient sampling of phase space when employing high level QM methods in a FTTCF framework.

Introduction

Vibrational infrared and Raman spectroscopy are among the most important experimental techniques currently used to obtain information on structures and chemical states of a variety of chemical systems. In order to increase the amount of information obtained about the system under investigation, experimental spectra are often compared to those generated from electronic structure methods. A clear assignment of the observed vibrational bands to specific intramolecular motions is one of the most important contributions of computational methods to the interpretation of experimental data. Moreover, given that accurate intra- and intermolecular force fields are available, computational vibrational spectroscopy is able to give detailed insight to molecular structures and their interactions with the environment.

* Corresponding author e-mail: steve.kaminski@chem.tu-berlin.de (S.K.), andrea.mroginski@tu-berlin.de (M.A.M.).

† Technische Universität Berlin.

‡ Universität Karlsruhe.

§ Indiana University.

10.1021/ct900660x
Published on Web 02/24/2010

For such calculations, density functional methods[1,2] (DFT) are usually employed since they offer accurate results, at least for systems in the gas phase, at a reasonable computational cost. With a given electronic structure method, the vibrational spectra of molecular systems can be accessed via several rather different methodologies: the so-called Normal Mode Analysis (NMA),[3] Fourier transform of time-correlation functions (FTTCF),[4] and methods based on a principal component analysis (PCA).[5]

NMA is by far the most frequently applied method in the field of computational chemistry. According to the NMA approach, the vibrational frequencies are obtained by the diagonalization of the Hessian matrix (second derivatives of the energy with respect to the atomic displacements) at an equilibrium geometry of the molecule. The spectral intensities are independently calculated through the first spatial derivatives of either the molecular dipole moment (infrared) or the molecular polarizability (Raman). The third feature of a vibrational band, its shape, is not accessible from the NMA approach since the calculations are formally performed at 0 K.

The second methodology for calculating vibrational spectra refers to the so-called Fourier transform of time-correlation functions (FTTCF). Based on Fermi's golden rule,[4] linear response theory[4,10,11] delivers expressions for a practical calculation of infrared and Raman spectra from dipole ($\boldsymbol{\mu}$) and polarizability ($\boldsymbol{\alpha}$) time correlation functions given by

$$\text{Infrared: } I(\omega) \propto \int_{-\infty}^{\infty} \langle \boldsymbol{\mu}(t) \cdot \boldsymbol{\mu}(0) \rangle \, \mathrm{e}^{-i\omega t} \, \mathrm{d}t \qquad (1)$$

$$\text{Raman: } I(\omega) \propto \int_{-\infty}^{\infty} \langle \boldsymbol{\alpha}(t) \cdot \boldsymbol{\alpha}(0) \rangle \, \mathrm{e}^{-i\omega t} \, d\mathrm{t} \qquad (2)$$

Once a time series of $\boldsymbol{\mu}$ or $\boldsymbol{\alpha}$ is available for a molecular system, the complete infrared or Raman spectrum containing frequency, intensity, and band shape information can be directly obtained from evaluating the expressions 1 and 2, respectively. These expressions can be applied under the assumption that the linear response approximation is valid; i.e., the perturbation of the system due to the applied field is small.

Besides the FTTCF formalism, another established methodology exists to extract molecular vibrations from classical trajectories. The principal component analysis, also often termed quasiharmonic analysis or essential dynamics, is based on a statistical analysis of (mass weighted) atomic fluctuations. The cross-correlator of such atomic displacements, called the covariance matrix, is usually diagonalized to yield its eigenvalues and eigenvectors, comparable to those obtained from the Hessian matrix in a standard normal-mode analysis. Thus, the method provides vibrational frequencies and normal modes of the regarded system.

The PCA technique is usually employed to identify large correlated motions in macromolecules, such as proteins,[6−8] with related modes in the far-infrared region (FIR). Wheeler and co-workers[9] were the first to evaluate the performance of this statistical method in the calculation of vibrational frequencies of small molecules in the mid infrared region (MIR).

However, for this type of calculation, Schmitz and Tavan[24] found that PCA-based methods exhibit significant limitations compared to the NMA and FTTCF formalisms.

The NMA technique depends on the harmonic approximation which assumes a quadratic potential energy expression at local minima of the potential energy surface. The necessity of finding minima on the potential surface is one of the major drawbacks of the NMA approach. Especially for large flexible molecules it becomes a nontrivial task to identify all its equilibrium conformations which are sampled at finite temperature and contribute to an experimental vibrational spectrum. These problems can be avoided by employing the FTTCF method. Here, the time series of $\boldsymbol{\mu}$ or $\boldsymbol{\alpha}$ are collected from a MD trajectory at finite temperature, where most of the equilibrium conformations sampled during a sufficiently long simulation contribute to the calculated spectrum.

Despite its drawbacks, the NMA approach is more widely used in the field of computational vibrational spectroscopy than the FTTCF method. One of the reasons is that, in order to generate a single spectrum, FTTCF requires an ensemble of structures which is computationally very demanding when using high level electronic structure methods like DFT. Nevertheless, numerous studies were performed concerning the calculation of infrared and Raman spectra via FTTCF of a rather small chemical system.[12−25] For many of these studies, the Car−Parrinello[26] molecular dynamics approach was used. Another reason is related to the assignment of vibrational bands to intramolecular motions. While being a straightforward task in the framework of NMA, it became only recently available[27,28] in terms of a systematic approach for the FTTCF formalism.

To benefit from the advantages of FTTCF, it would be necessary to combine this methodology with a computationally more efficient electronic structure method. An established approximative quantum chemical method is the Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB) method derived by Elstner and co-workers.[29] One of its strengths is the accurate calculation of molecular geometries comparable to higher levels of theory.[30] At the same time, the performance of SCC-DFTB comes at a considerably reduced computational cost, i.e., about 3 orders of magnitude less compared to standard DFT methods. This allows a treatment of much larger chemical systems as, e.g., biomolecules for which SCC-DFTB has already shown to deliver accurate results compared to higher level methods.[31,32]

Concerning spectroscopic properties, Witek and co-workers have already shown that vibrational frequencies[33,34] as well as Raman intensities[35] in the framework of a normal-mode analysis are satisfactorily described by SCC-DFTB.

The combination of SCC-DFTB with FTTCF has also been successfully employed for several simple chemical systems[22,36,37] in infrared studies. Furthermore, SCC-DFTB has recently been shown to deliver accurate vibrational infrared spectra of molecules in a complex protein environment.[38] Albeit these infrared studies using SCC-DFTB in combination with FTTCF are promising, this approach has to our knowledge not been used so far for the calculation of Raman spectra. In our opinion, however, the FTTCF approach could be valuable for the interpretation of various experimental

Raman studies, especially on biomolecular systems containing large floppy molecules. The size of such systems requires a combination of FTTCF with an efficient quantum chemical method such as SCC-DFTB.

Therefore, the aim of this work was the extension of the SCC-DFTB method to calculate vibrational Raman spectra via the FTTCF formalism. The implemented methodology will be tested in detail on several small organic model compounds in the gas phase. Due to the importance of an accurate prediction of condenced phase spectra, the calculated spectrum of L-phenylalanine in aqueous solution will be also compared to experimental results.

Although beyond the scope of the present work, the long-term goal is to find out whether this methodology can be helpful to interpret Raman spectra of biomolecules (especially cofactors in proteins) in a complex protein/solvent environment.

Theoretical Approach

The SCC-DFTB Formalism. The SCC-DFTB approach is an approximate quantum chemical method for which an extensive description is given elsewhere in the literature.[29] Like semiempirical methods, SCC-DFTB benefits from several approximations such as avoiding the calculations of one- and two-electron integral expressions as well as taking only valence electrons explicitly into account. Therefore, a computational speedup of about 3 orders of magnitude compared to DFT is achieved.

The SCC-DFB energy based on a second-order expansion of the DFT energy with respect to density fluctuations relative to a chosen reference density is given by

$$E = \sum_{i\mu\nu} c_{i\mu} c_{i\nu} H^0_{\mu\nu} + \frac{1}{2} \sum_{A,B} \Delta q_A \Delta q_B \gamma_{AB} + E^{\text{rep}} \qquad (3)$$

where $c_{i\mu}$ are the coefficients for the minimal basis representation of confined pseudoatomic orbitals $\psi_i = \Sigma_\mu c_{i\mu} \phi_\mu$ and $H^0_{\mu\nu}$ the Hamilton matrix which depends only on the reference density. The induced charge on each atom A is denoted as Δq_A; γ_{AB} is a distance dependent function describing charge interactions, and E^{rep} means a sum of two-centered core potentials. The coefficients $c_{i\mu}$ are determined by solving the Kohn−Sham equations and transforming them into a set of algebraic equations,

$$\sum_\nu c_{i\nu} (H_{\mu\nu} - \varepsilon_i S_{\mu\nu}) = 0 \qquad (4)$$

with the charge self-consistent Hamiltonian

$$H_{\mu\nu} = \langle \phi_\mu | \hat{H}^0 | \phi_\nu \rangle + \frac{1}{2} S_{\mu\nu} \sum_C \Delta q_C (\gamma_{AC} + \gamma_{BC}) \qquad (5)$$

The overlap matrix elements $S_{\mu\nu}$ and the $H^0_{\mu\nu}$ are calculated using the PBE functional[39] and tabulated for a dense mesh of interatomic distances.

As it will be discussed in one of the following sections in more detail, we need an expression for the SCC-DFTB energy of a molecular system interacting with an external electric field to obtain a Raman spectrum via the FTTCF formalism. For this purpose, we follow Elstner[40] with the addition of an extra term to eq 3, describing the interaction of the field with the induced Mulliken charges of the system, as

$$E^{\text{field}} = E - \sum_A \Delta q_A \sum_{j=1}^{3} D_j x_{jA} \qquad (6)$$

Here, D_j is the Cartesian component of the electric field and x_A denotes the Cartesian coordinate of atom A. The charge self-consistent Hamiltonian (eq 5) becomes

$$H^{\text{field}}_{\mu\nu} = H_{\mu\nu} - \frac{1}{2} S_{\mu\nu} \sum_{j=1}^{3} D_j x_{jA} \qquad (7)$$

The FTTCF Formalism for Raman Spectra. The evaluation of the FTTCF formalism for Raman spectra, discussed only in brief, starts with the Raman differential scattering cross section in the quantum mechanical framework[4,10] given by

$$\lambda^4 \frac{d^2\sigma}{d\omega d\Omega} = \sum_i \sum_f \rho_i |\langle f | \varepsilon^s \mathbf{P} \cdot \varepsilon^0 | i \rangle|^2 \delta(\omega_{fi} - \omega) \qquad (8)$$

where $|i\rangle$ and $|f\rangle$ denote the wave functions of the initial and final states of the system and ρ_i is the probability for the system to be found in the initial state i. On the left side, λ denotes the wavelength of the scattered radiation and the absorption frequency ω_{fi} is proportional to the energy levels of the final (E_f) and initial states (E_i) via $\omega_{fi} = (E_f - E_i)/\hbar$. $\mathbf{P}$ is the Cartesian polarizability tensor which can be split up into an isotropic and an anisotropic component, such as

$$\mathbf{P} = \mathbf{P}_{\text{iso}} + \mathbf{P}_{\text{aniso}} = \begin{pmatrix} \alpha_{xx} & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & \alpha_{yy} & \alpha_{yz} \\ \alpha_{zx} & \alpha_{zy} & \alpha_{zz} \end{pmatrix} \qquad (9)$$

For eq 8 the associated linear-response equations derived by Gordon[4,10,11] are given by

$$\left(\lambda^4 \frac{d^2\sigma}{d\omega d\Omega}\right)_{\text{iso}} = \int_{-\infty}^{\infty} \left\langle \frac{1}{3} tr[\mathbf{P}_{\text{iso}}(t) \mathbf{P}_{\text{iso}}(0)] \right\rangle e^{-i\omega t} \, dt \qquad (10)$$

$$\left(\lambda^4 \frac{d^2\sigma}{d\omega d\Omega}\right)_{\text{aniso}} = \int_{-\infty}^{\infty} \langle tr[\mathbf{P}_{\text{aniso}}(t) \mathbf{P}_{\text{aniso}}(0)] \rangle \, e^{-i\omega t} \, dt \qquad (11)$$

where tr denotes the trace of a matrix. These are valuable expressions for a practical calculation of Raman spectra. Basically, only the isotropic and anisotropic polarizability tensors $\mathbf{P}_{\text{iso}}(t)$ and $\mathbf{P}_{\text{aniso}}(t)$ as a function of time are needed. The corresponding rotational invariants[11] can be defined as

$$\alpha_{\text{iso}} = \frac{1}{3}[(\alpha_{xx} + \alpha_{yy} + \alpha_{zz})] \qquad (12)$$

$$\alpha^2_{\text{aniso}} = \frac{1}{3}[(\alpha_{xx} - \alpha_{yy})^2 + (\alpha_{yy} - \alpha_{zz})^2 + (\alpha_{zz} - \alpha_{xx})^2 + (\alpha_{yy} - \alpha_{zz})^2 + 6(\alpha^2_{xy} - \alpha^2_{yz} + \alpha^2_{yz})] \qquad (13)$$

The two invariants measuring the isotropy and anisotropy of the electronic polarizability are connected to $\mathbf{P}_{\text{iso}}$ and $\mathbf{P}_{\text{aniso}}$ via

$$\frac{1}{3}tr[\mathbf{P}_{iso}\mathbf{P}_{iso}] = \alpha_{iso} \quad (14)$$

$$tr[\mathbf{P}_{aniso}\mathbf{P}_{aniso}] = \alpha^2_{aniso} \quad (15)$$

The isotropic and anisotropic components of the scattering cross sections apply, in particular, to experimental spectra where the scattered light is measured parallel or perpendicular to the plane of the polarized incident laser beam[11]

$$\left(\lambda^4\frac{d^2\sigma}{d\omega d\Omega}\right)_{\parallel} = \left(\lambda^4\frac{d^2\sigma}{d\omega d\Omega}\right)_{iso} + \frac{2}{15}\left(\lambda^4\frac{d^2\sigma}{d\omega d\Omega}\right)_{aniso} \quad (16)$$

$$\left(\lambda^4\frac{d^2\sigma}{d\omega d\Omega}\right)_{\perp} = \frac{1}{10}\left(\lambda^4\frac{d^2\sigma}{d\omega d\Omega}\right)_{aniso} \quad (17)$$

The ratio of these expressions leads to an important observable in Raman spectroscopy, the depolarization ratio ρ,

$$\rho = \frac{\left(\lambda^4\frac{d^2\sigma}{d\omega d\Omega}\right)_{\perp}}{\left(\lambda^4\frac{d^2\sigma}{d\omega d\Omega}\right)_{\parallel}} \quad (18)$$

To summarize, the vibrational Raman spectrum of a chemical system can be obtained by generating an ensemble of structures (indicated by $\langle ... \rangle$ in eqs 10 and 11), e.g., through molecular dynamics simulations and simultaneous electronic structure calculations at each time step to compute the polarizability tensor components (eqs 12 and 13: α_{xx}, α_{yy}, ...).

However, care must be taken since a single simulation usually covers only a small amount of phase space and the simulation time is usually also insufficiently small to reach an equal distribution of energy among the normal modes, leading to an erroneous intensity pattern in the spectra as pointed out, e.g., by Hornicek and co-workers.[14] More efficient than producing one rather long single simulation is the generation of multiple shorter independent trajectories, each starting with a set of randomly reassigned atomic velocities. Averaging over a certain number of spectra generally yields reliable relative intensities.

Additional errors concerning spectral intensities arise from classically (Newtonian) derived trajectories and time-correlation functions for which Berens and co-workers[11] proposed a quantum correction factor for Raman spectra,

$$Q_{hc} = \frac{\beta\hbar\omega}{1 - \exp(-\beta\hbar\omega)} \quad (19)$$

where β is equal to $(kT)^{-1}$, k denotes the Boltzmann constant, and ω is the vibrational frequency.

Implementation of Polarizability Calculations. In the previous section, the time dependent molecular polarizability $\mathbf{P}_{iso/aniso}(t)$ was identified as the essential quantity for the evaluation of the linear response eqs 10 and 11. The components of $\mathbf{P}$ (α_{xx}, α_{yy}, ...) can be practically evaluated by considering the perturbation of a molecular system exposed to an external electric field $\mathbf{F}$. The energy of the perturbed system can be expressed in a Taylor series relative to the field-free energy[41] as

$$E(\mathbf{F}) = E(0) + \sum_i \left(\frac{\partial E}{\partial F_i}\right)_0 F_i + \frac{1}{2}\sum_{ij}\left(\frac{\partial^2 E}{\partial F_i \partial F_j}\right)_0 F_i F_j ... \quad (20)$$

Assuming the harmonic approximation, the Taylor expansion is truncated after the second term which is the response property of interest, the components of the molecular polarizability,

$$\alpha_{ij} = -\left(\frac{\partial^2 E}{\partial F_i \partial F_j}\right) \qquad \text{with } i,j = x,y,z \quad (21)$$

The required second order derivatives of the energy with respect to the electric field components can be performed numerically using the following expressions[42] for diagonal

$$\alpha_{ii} = \frac{\partial^2 E}{\partial F_{ii}^2} = \frac{1}{F^2}(E_{i0} - 2E_{00} + E_{-i0}) + O(F^2) \quad (22)$$

and off-diagonal components of the polarizability tensor

$$\alpha_{ij} = \frac{\partial^2 E}{\partial F_{ij}} = \frac{-1}{2F^2}(E_{i0} + E_{-i0} + E_{0j} + E_{0-j} - 2E_{00} - E_{ij} - E_{-i-j}) + O(F^2) \quad (23)$$

where the expression for the error $O(F^2)$ can be found elsewhere.[43] The numeric step size, here equal to the electric field strength, must be chosen carefully. Too strong applied fields, on the one hand, may hamper the SCF convergence in electronic structure calculations, while a too weak perturbation may lead to numeric errors because of small energy differences. To test the numeric stability of the derived values of α_{ij}, we follow the approach of Magdó[58] by doubling the electric field strength and evaluating α_{ij} again via

$$\alpha_{ij} = \frac{1}{3}\left(4\frac{\partial^2 E}{\partial F_i \partial F_j} - \frac{\partial^2 E}{\partial 2F_i \partial 2F_j}\right) \quad (24)$$

A reduced numeric error ($O(F^4)$ instead of $O(F^2)$) is obtained using eq 24, whereas the required number of energy evaluations in the presence of an electric field increases from 12 to 24.

Vibrational Mode Assignment. Martinez and co-workers developed a methodology[27,28] for the assignment of vibrational bands to intramolecular motions from FTTCF calculations. In the following section, we will briefly summarize the most important aspects.

The time dependent atomic velocities as available from a molecular dynamics trajectory are used as a key quantity of the methodology. The power spectra of velocity autocorrelation functions (vibrational density of states) have been used earlier to assign vibrational bands to atomic motions.[12] However, these power spectra were highly delocalized in frequency space, making an unambigious band assignment for larger molecules a difficult task. The methodology of Martinez and co-workers provides so-called effective normal modes, i.e., linear combinations of atomic displacements

constructed in such a way that their corresponding power spectra are as localized as possible in frequency. To achieve this task, the following generalized eigenvalue equation must be solved:

$$\mathbf{Z}^{-1}\mathbf{K}^{(2)} = \mathbf{Z}^{-1}\mathbf{K}^{(0)}\Lambda \tag{25}$$

in which Λ is a diagonal matrix containing the vibrational frequencies. The solution matrix $\mathbf{Z}$ performs the transformation between a set of j initial velocities $\dot{x}_j$ in either Cartesian or internal coordinate space into the final set of k effective normal mode velocities $\dot{q}_k$ via

$$\dot{q}_k(t) = \mathbf{Z}^{-1}\dot{x}_j(t) \tag{26}$$

Power spectra P_k^q generated from these effective normal mode velocities $\dot{q}_k$ via

$$P_k^q(\omega) = \int_{-\infty}^{\infty} \langle \dot{q}_k(t) \cdot \dot{q}_k(0) \rangle \, \mathrm{e}^{-i\omega t} \, \mathrm{d}t \tag{27}$$

appear as localized bands in frequency. The matrices $\mathbf{K}^{(n)}$, necessary to set up eq 25, are calculated via

$$K_S^{(n)} = \frac{\beta}{2\pi} \int_{-\infty}^{\infty} |\omega|^n P_{kl}^x(\omega) \, \mathrm{d}\omega \tag{28}$$

Here, P_{kl}^x denotes the power spectra of all auto- and cross-correlation functions from the initial set of atomic velocities in Cartesian or internal coordinate space,

$$P_{kl}^x(\omega) = \int_{-\infty}^{\infty} \langle \dot{x}_k(t) \cdot \dot{x}_l(0) \rangle \, \mathrm{e}^{-i\omega t} \, \mathrm{d}t \tag{29}$$

From the matrices $\mathbf{K}$ and $\mathbf{Z}$, an expression for the potential energy distribution (PED) can be defined,[44] i.e., the contribution of the internal coordinate i to the potential energy of the jth normal mode,

$$\mathrm{PED}_{ij} = \sum_{k}^{3N-6} \frac{Z_{ji} Z_{ki} K_{jk}^{(2)}}{\lambda_{ii}} \tag{30}$$

Computational Details

In order to test the performance of the implementation described in the previous section, we have chosen a set of 10 small organic molecules for *in vacuo* calculations: water (H_2O), butadiene (C_4H_6), ethanethiol (C_2H_6S), benzene (C_6H_6), methylacetate ($C_3H_6O_2$), maleimide ($C_4H_3NO_2$), *n*-methylacetamide (C_3H_7NO), pentane (C_5H_{12}), trimethylamine (C_3H_9N), and glycerol ($C_3O_3H_8$). In addition, the methodology was also tested for a single L-phenylalanine residue in aqueous solution.

In order to generate vibrational Raman spectra for these compounds via the FTTCF formalism, molecular dynamics simulations were performed using an extended SCC-DFTB program of Elstner and co-workers interfaced with the CHARMM[45] 32b2 software. The MD simulation protocol is described as follows for *in vacuo* simulations. In a first step, we performed an initial geometry optimization using CHARMM's adopted basis Newton−Raphson (ABNR) minimizer until a gradient threshold of 1×10^{-5} au was reached. The system was then heated and equilibrated by means of a 12.5 ps MD simulation. Subsequently, these simulations were continued for a further 400 ps under constant temperature conditions (300 K). During this production run, the fluctuating polarizability tensor elements (α_{ij}) needed to generate Raman spectra were collected at each time step.

In the case of phenylalanine, the amino acid in its zwitterionic form (as usual at neutral pH) was placed in a cubic box of TIP3P[46] water (1485 molecules) of dimensions $35 \times 35 \times 35$ Å^3. Periodic boundary conditions were applied, and electrostatic interactions beyond a cutoff radius of 12 Å were neglected by employing an atom based force-shifting function. All calculations were performed in a QM/MM framework, with phenylalanine as the QM part, as implemented in the SCC-DFTB/CHARMM interface.[47] After an initial minimization of the complete system for 1000 steps using CHARMM's ABNR minimizer, a subsequent MD simulation for the purpose of heating and equilibration was performed for 50 ps. For the following production runs (altogether 400 ps) in the NPT ensemble, the Anderson-Hoover equations[48] for constant pressure and temperature as implemented in CHARMM were employed.

Subsequent to the MD simulations, the Fourier transform of the polarizability autocorrelation was computed and further processed with a Blackman[49] filter to increase the signal-to-noise ratio. In a final step, the spectra were treated with a quantum correction factor (see eq 19) improving the intensity pattern as suggested by Berens and co-workers for Raman spectra generated from classical trajectories. Following eqs 10, 11, and 19, the Raman spectrum is finally calculated as

$$I_\omega = Q_{hc} \cdot \left[\left(\lambda^4 \frac{d^2\sigma}{d\omega d\Omega} \right)_{\mathrm{iso}} + \left(\lambda^4 \frac{d^2\sigma}{d\omega d\Omega} \right)_{\mathrm{aniso}} \right] \tag{31}$$

In accordance with Nyquist's theorem,[50] the spectral resolution after a Fourier transform is inversely proportional to the product of the simulation length and the chosen time step. Due to the rather low spectral density of the chosen small organic compounds for *in vacuo* calculations, a resolution of 8 cm^{-1} was assumed to be sufficient. To be comparable with the experimental results, however, spectra of phenylalanine in solution were calculated at a resolution of 4 cm^{-1}. Hence, with a chosen time step of 0.5 fs, 8192 and 16 384 (phenylalanine) simulation steps were necessary for the generation of a single spectrum using a Fast Fourier Transform (FFT) routine. For reasons mentioned in the theory section, 100 independent MD simulations were performed per model compound, and their resulting spectra were averaged to obtain a reliable intensity pattern.

In order to evaluate the performance of the implemented SCC-DFTB/FTTCF procedure, additional Raman spectra were computed at the same level of theory (SCC-DFTB) but using the NMA technique as implemented into an SCC-DFTB standalone code by Witek et al.[33,35] In fact, the accuracy of this theoretical approach should be tested against experimental spectra of molecules in the gas phase. Unfortunately, experimental spectroscopic data recorded in the gas phase are very difficult to obtain. Thus, most of the measurements are performed in solution. Comparison to

experimental data, however, will be done for the SCC-DFTB/FTTCF calculations of L-phenylalanine in water.

It should be mentioned that not only relative intensities will differ between the FTTCF and NMA approach but also vibrational frequencies, since for FTTCF they depend on the integration time step. The chosen value of 0.5 fs leads to a frequency-dependent blue shift in the spectrum according to[23]

$$\Delta\omega = \omega^3 \Delta t^2/24 \tag{32}$$

The chosen step size (Δt) of 0.5 fs is believed to be a good compromise between accuracy and computational cost.

In addition, we compared the performance of SCC-DFTB/FTTCF with Raman spectra calculated at higher levels of theory. Density functional theory (DFT) provides efficient and accurate access to vibrational spectra. It has been shown[51] that reliable Raman intensities depend mostly on the size and quality of the basis set in use. Dunning's augmented triple-ζ basis set[52] (aug-cc-pVTZ) performs well in this respect. Since recent work of Riley et al.[53] also revealed very good performance in the prediction of vibrational frequencies (40 cm^{-1} error) for the combination of the BLYP[54,55] functional together with Dunning's aug-cc-pVTZ basis set, we have chosen this setup in combination with an NMA approach for comparison to our SCC-DFTB/FTTCF results.

Besides our implementation, it is also possible to generate a Raman spectrum via the FTTCF formalism out of a hybrid approach. Here, we followed the procedure of Yu and Cui[22] by performing single point calculations (to obtain the molecular polarizability) at the DFT level on snapshots sampled from a trajectory computed at the SCC-DFTB level. Here, the DFT calculations were done using a computationally less demanding combination of functional and basis set (B3LYP/6-31 g(d)), since numerous single-point calculations are required (8192) to generate a spectrum. Therefore, they were restricted to only 2 of our 10 model compounds: butadiene and maleimide. With such a setup, we tried to estimate the impact of two parameters on the relative Raman intensities: namely, the size of the basis set and the number of independent trajectories used for spectral averaging. All DFT based calculations were performed using Gaussian 03.[56]

Since the output of either SCC-DFTB/NMA or DFT-BLYP/NMA calculations consists of Raman activities A_k rather then intensities I_k, they were multiplied with a frequency dependent factor to be comparable to the SCC-DFTB/FTTCF results

$$I_k = \frac{(\omega_0 - \omega_k)^4}{\omega_k\left(1 - \exp\left[-\frac{h\omega_k}{K_b T}\right]\right)} \cdot A_k \tag{33}$$

Here, K_b denotes the Boltzmann constant, ω_k is the vibrational frequency of mode k, and the incident laser line ω_0 was chosen at a value of 1064 nm. The temperature has been chosen to be 300 K.

An important aspect concerning the interpretation of the calculated spectra is the assignment of Raman active bands to intramolecular motions. In the framework of NMA, the eigenvectors of the Hessian matrix, referring to vibrational motions, are calculated by default. Since the *in vacuo* SCC-DFTB/FTTCF calculations will be compared to calculations using the NMA approach, the eigenvectors will be used for a qualitative assignment of bands to internal motions.

The solution spectrum of phenylalanine from SCC-DFTB/FTTCF calculations has been compared to experimental data. Therefore, effective normal modes as described previously were evaluated, and the potential energy distribution of selected modes were estimated via eq 30. The atomic displacements of phenylalanine were expressed in the internal coordinate space. For this purpose, a set of 63 nonredundant internal coordinates for the amino acid were defined from bond length, bond angles, and proper and improper (out-of-plane) torsion angles following the rules of Pulay et al.[57]

Experimental Procedure

The vibrational Raman spectrum of L-phenylalanine in water ($\sim$0.18 M) was measured at two temperatures (298 and 133 K) and pH $\sim$ 7 using a Bruker RFS 100/S Fourier-transform spectrometer. The excitation line was at 1064 nm, and the spectrum has been recorded at a resolution of 4 cm^{-1}.

Results and Discussion

Numeric Stability of Calculated Polarizabilities. For all 10 model compounds, MD simulations were performed to find appropriate numeric step sizes (the applied field strength) for the computation of the polarizability tensor components. Magdó et al.[58] suggested values of about 0.004 au for polarizability calculations on linear tetrapyrroles at local minima and in the framework of a NMA approach. According to the FTTCF methodology, however, the computation of polarizabilities are performed on structures extracted out of a MD trajectory, which are usually not in a local minimum. Therefore, we tested the step sizes 0.0005, 0.001, 0.005, 0.01 and 0.05 au. For all model compounds except glycerol, methylacetate, and maleimide, the effects were negligible, meaning that the spectra could not be visually distinguished from each other. For the mentioned compounds, however, the smallest step size (0.001) leads to slight artificial baseline drifts, while the spectra resulting from the two higher step sizes are almost identical. Raman spectra of methylacetate and glycerol, estimated using different step sizes, are shown in Figure 1. A pronounced baseline effect together with a decreasing signal-to-noise ratio is clearly visible for the spectrum obtained using a step size of 0.0005 au. Large step sizes (0.05 au) on the contrary only have quite small influences on the overall spectral shape.

Whereas these calculations have been performed using a simple numeric differentiation scheme (eq 21), additional tests for all model compounds were done using a more accurate approach as shown in eq 24. A comparison between the two methodologies revealed a visually negligible difference for the resulting spectra. Therefore, to keep the computational effort low, all simulations in this work have been performed using eq 21.

Stability of Spectral Intensities. As pointed out in the theory section, only an average over spectra from several

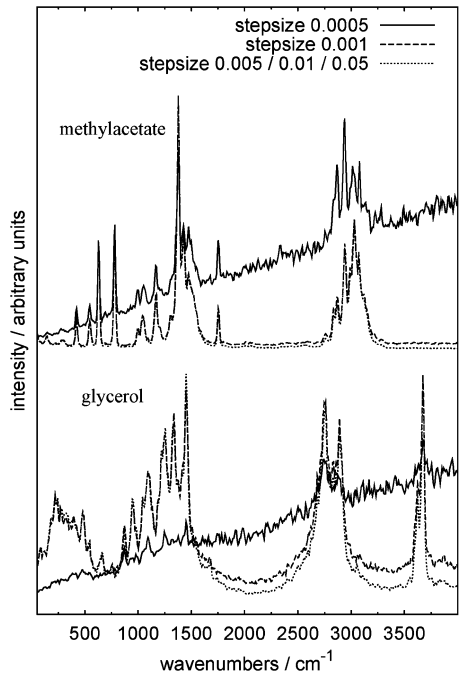

Figure 1. Raman spectra of methylacetate and glycerol in dependence of numeric step sizes for the evaluation of the polarizability tensor elements. The resulting Raman spectra for step sizes of 0.005, 0.01, and 0.05 au can hardly be distinguished and are therefore represented by a single dotted line.

independent MD simulations will give reliable relative band intensities. Figure 2 exemplifies for two model compounds (pentane, *n*-methylacetamide) that spectra generated from single trajectories (bottom spectra) are by no means representative. The top spectra, on the other hand, show that in all of the spectral regions, changes in intensities between 50 and 100 on the average remain very small, so that convergence in the intensity pattern has been achieved.

Raman Spectra in the Gas Phase. Among the 10 model compounds for which *in vacuo* calculations have been performed, five of them have been selected for a more detailed discussion. For the remaining model compounds, only the spectra will be shown. All spectra derived from the time-correlation-function formalism as implemented in this work (termed SCC-DFTB/FTTCF) result from simulations with a step size of 0.005 au and from an average of 100 single spectra.

For the purpose of a better visual comparison with the SCC-DFTB/FTTCF results, the line spectra from the NMA approach have been convoluted using Lorentzian functions with a half-width of 10 cm^{-1} and a peak maximum at the position of the calculated vibrational frequencies. In all vibrational spectra shown, the most intense band was scaled to unity.

Water. For the smallest model compound, the resulting spectra in Figure 3 (top left) show excellent agreement when comparing the NMA and FTTCF formalisms for the SCC-DFTB method in terms of vibrational frequencies. Compared to BLYP aug-cc-pVTZ, the highest frequency mode, i.e., the asymmetric O−H stretching, is substantially blue-shifted (244 cm^{-1}), whereas the other two modes are in good agreement with SCC-DFTB results. Concerning the relative intensities, very good agreement is found between SCC-DFTB/FTTCF and BLYP aug-cc-pVTZ, wheras SCC-DFTB/NMA shows a distinctly different pattern.

Glycerol. Analogous to water, the calculated vibrational frequencies compare well (see Figure 4 /bottom right) between SCC-DFTB/FTTCF and SCC-DFTB/NMA. This is also true for most spectral regions when incorporating BLYP aug-cc-pVTZ into the comparison. The largest spectral shift (∼160 cm^{-1}) can be found for the double band feature in the C−H stretching region near 3000 cm^{-1}. A striking feature of the glycerol spectrum from the SCC-DFTB/FTTCF approach is its broad background. It is the only model compound where intramolecular hydrogen bonds appear with a continuous variation of donor−acceptor distances (O−H···O) during the MD simulations, which is most likely responsible for the observed broadening. In terms of the spectral shape, the SCC-DFTB/NMA approach is in closer agreement to BLYP aug-cc-pVTZ in the high frequency region above 3500 cm^{-1}. Here, the O−H stretching vibrations are overestimated in the SCC-DFTB/FTTCF spectrum. Below 500 cm^{-1}, O−H wagging modes at 249 and 262 cm^{-1} are overestimated for the SCC-DFTB/NMA approach.

Ethanethiol. For the only sulfur compound, good agreement in band positions is found for all three compared methodologies in the region between 500−1500 cm^{-1}, as shown in Figure 4. The largest deviation comparing SCC-DFTB and DFT occurs for the S−H stretching mode at 2538 cm^{-1} (SCCDFTB/NMA) which is blue-shifted by 130 cm^{-1}. The intensity pattern, however, varies strongly between SCC-DFTB/FTTCF and SCC-DFTB/NMA for three distinct modes. These are the ones at 139, 684, and 3065 cm^{-1} (referring mainly to S−H wagging, S−C stretching, and C−H stretching vibrations), which are over- (139 cm^{-1}) and underestimated (684 and 3065 cm^{-1}), respectively, for the SCC-DFTB/NMA approach compared to the other two methodologies. In the high frequency region above 2500 cm^{-1}, the spectral shapes of both SCC-DFTB approaches are not in accordance with BLYP aug-cc-pVTZ.

Maleimide. While SCC-DFTB/FTTCF shows good agreement concerning band positions (Figure 3) over the whole spectral range compared to SCC-DFTB/NMA and BLYP aug-cc-pVTZ, its spectral shape is clearly distinct from BLYP aug-cc-pVTZ, especially the N−H stretching mode at ∼3500 cm^{-1}, just as the region below 1200 cm^{-1} is overestimated with respect to the most pronounced band at 1771 cm^{-1}. On the contrary, the SCC-DFTB/NMA compares well concerning relative intensities to BLYP aug-cc-pVTZ over the entire spectral range.

***n*-Methylacetamide.** Both SCC-DFTB approaches deliver a very similar vibrational spectrum concerning band positions and the overall spectral shape, as illustrated in Figure 5. The agreement with BLYP aug-cc-pVTZ in terms of band positions is also very good, except in the C−H stretching region around 3000 cm^{-1}. Overestimated intensities for both SCC-DFTB methods can be found in the broad spectral feature near 1450 cm^{-1} mainly belonging to C−H methly deformations. In the C−H stretching region, the bands from

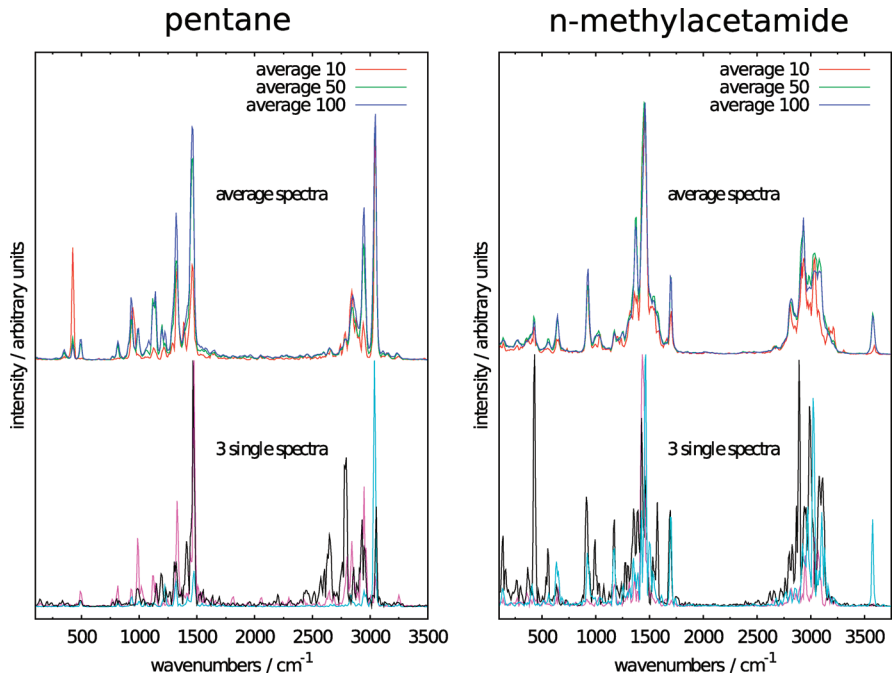

Figure 2. Raman spectra from the SCC-DFTB/FTTCF formalism of two model compounds. The top spectra illustrate the convergence of the overall spectral shape as the number of single spectra for averaging increases. On the bottom, three spectra from single independent trajectories are shown to illustrate spectral variations.

both SCC-DFTB approaches are underestimated compared to BLYP aug-cc-pVTZ.

Raman Spectra in Solution. In Figure 6, experimental and calculated Raman spectra of L-phenylalanine in water are shown. Since the experimental spectrum recorded at room temperature, which is principally the correct one to compare to our calculations, exhibits an intense background below 750 cm^{-1}, it has been measured again at ~133 K. The observed background is most likely due to couplings of the solutes' vibrations to the librational motions of water which are largely suppressed at 133 K. As a result, in the low temperature spectrum, a better resolution of the solutes' vibrations in this region is observed. Therefore, we will refer to this 133 K spectrum in the following discussion.

Concerning intensity fluctuations arising from the MD simulation, small variations are found between spectra generated out of 50 and 100 independent trajectories. This is basically the same observation as for the *in vacuo* calculations, indicating a sufficient sampling of the solute in aqueous solution.

The overall spectral pattern obtained from SCC-DFTB/FTTCF calculations on L-phenylalanine fits the experiment to an extent that makes it possible to qualitatively assign most of the spectral regions, as illustrated by dashed vertical lines in Figure 6. Especially the experimental line shapes are mostly well reproduced. However, several regions in the calculated spectrum deviate significantly in terms of vibrational frequencies and/or relative intensities as compared to the experiment. These regions are labeled 1−5 in Figure 6 and will be discussed in more detail. For the purpose of an assignment of the labeled bands to vibrational motions, effective normal modes have been calculated for selected bands as shown in Figure 7. Here, the colored spectra denote the normal modes localized in frequency and therefore are helpful for an assignment of Raman active vibrational bands.

In spectral region 1 in Figure 6, the bands are overestimated by the SCC-DFTB/FTTCF calculations compared to the experiment and blue-shifted by ~80 cm^{-1}. From a PED analysis (eq 30), these bands are related to C−H out-of-plane motions of the benzene ring as well as C−C backbone stretchings and C−C−C bendings of the benzene ring.

The intensity of the single band marked as number 2 in Figure 6, which can be assigned to the most prominent one in the experiment, is significantly underestimated by the SCC-DFTB/FTTCF calculations, and its vibrational frequency is blue-shifted by ~55 cm^{-1}. The related motions are mainly C−C−C bendings of the benzene ring.

The relative intensities between the three Raman active bands in region 3 are well reproduced by the SCC-DFTB/FTTCF calculations. The intensity of this spectral region compared to the neighboring ones, hovewer, is overestimated and blue-shifted by ~90 cm^{-1}. The most prominent band in the SCC-DFTB/FTTCF spectrum can be assigned to C−C stretchings and C−C−H bending motions in the benzene ring. The second and third bands are mainly composed of C−C−H bending motions in the benzene ring as well as backbone C−C and C−N stretchings.

The spectral region number 4 is characterized by a blue-shifted (~60 cm^{-1}) broad feature in the SCC-DFTB/FTTCF calculations. Its intensity is overestimated, and the related motions are C−C−H backbone bending and C−C as well as C−O backbone stretching motions.

The highest frequency modes in the spectrum in Figure 6 corresponding to region 5 are extremely shifted to higher wavenumbers (~200 cm^{-1}) as compared to the experiment. This double band feature is dominated by C−C stretching motions in the benzene ring. This is not surprising since it

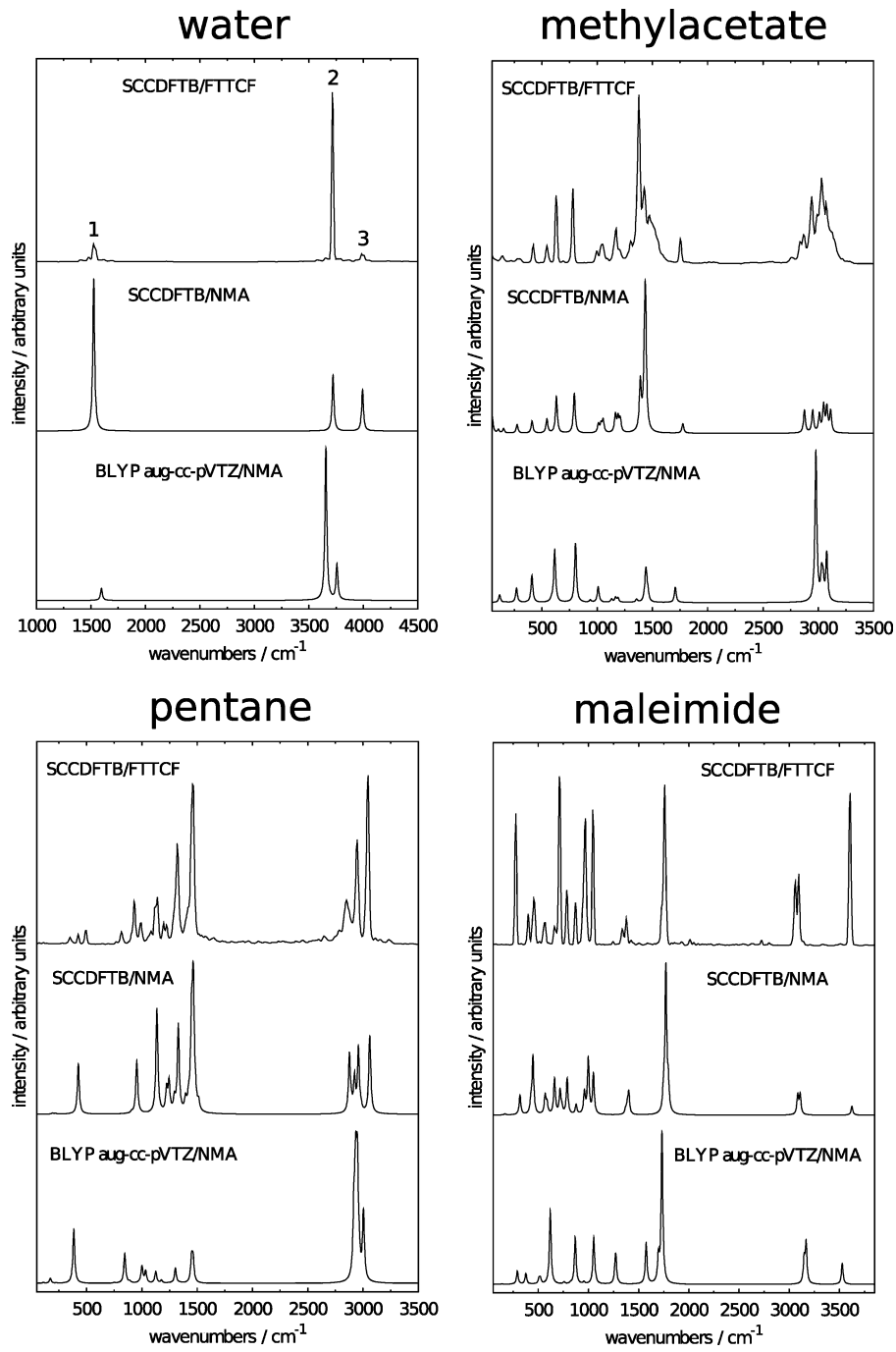

Figure 3. Vibrational Raman spectra of 4 model compounds from different methodologies (FTTCF vs NMA) and different levels of theory (SCC-DFTB vs DFT) for comparison. Raman active modes of water for which depolarization ratios were estimated are marked.

is well-known for the non-self-consistent DFTB method that for benzene in the gas phase the C−C stretching mode with symmetry E_{2g} is overestimated by more than 200 cm^{-1}.[59] For the self-consistent scheme, this shortcoming is not eliminated, as shown for benzene (mode 5 in Figure 4). For SCC-DFTB, this band appears at 1826 cm^{-1}, whereas for BLYP it is found at 1571 cm^{-1}.

SCC-DFTB Repulsive Potentials for Vibational Spectra Calculations. The vibrational Raman spectra from SCC-DFTB/FTTCF calculations presented so far in this work have shown to be in overall good agreement with higher level theoretical methods and experimental data. However, vibrational bands referring to C−C stretching and bending motions in conjugated π systems (butadiene, benzene, phenylalanine) show large frequency shifts up to 200 cm^{-1}. Overpolarization effects in conjugated systems, a known problem of SCC-DFTB,[60] are responsable for such errors.

To overcome such problems, the improvement of SCC-DFTB repulsive pair potentials for a better prediction of a variety of molecular properties is in progress. Małolepsza and co-workers[61] developed a set of pair potentials which substantially improve calculated vibrational frequencies for modes where hydrogen and carbon atoms are involved. Furthermore, Gaus and co-workers[62] recently presented a procedure for an automatized parametrization of repulsive potentials for several molecular properties.

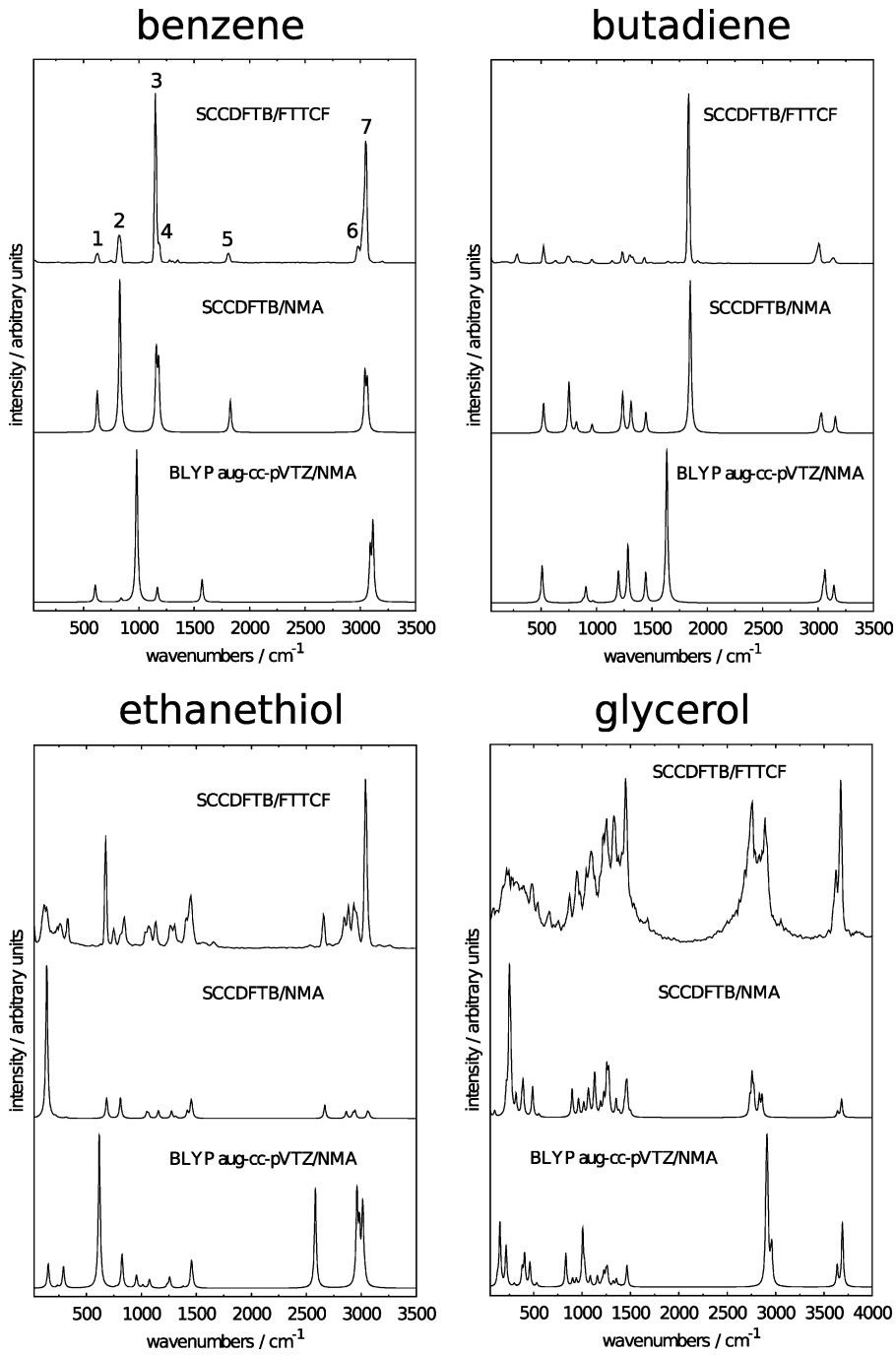

Figure 4. Raman spectra of 4 model compounds from different methodologies (FTTCF vs NMA) and different levels of theory (SCC-DFTB vs DFT) for comparison. Raman active modes of benzene for which depolarization ratios were estimated are marked.

To investigate the effect of parametrized pair potentials on vibrational frequencies, we recalculated Raman spectra of benzene and butadiene (since optimized parameters only exist for carbon and hydrogen so far) in the gas phase with the SCC-DFTB/FTTCF protocol. The results (termed new-sk (Slater−Koster) parameters) shown in Figure 8 have been compared to SCC-DFTB/FTTCF calculations with the standard set of repulsive potentials (top spectra). Vibrational bands significantly affected by the new parameters in use were marked with symbols (triangle, star, circle). For benzene, three strongly shifted bands can be observed as shown in Figure 8. The resulting spectrum is in much closer agreement to the one from high level density functional methods (first two spectra from the bottom). This is also true for butadiene. Here, the intense C═C stretching mode (marked with a triangle) is red-shifted by approximately 200 cm^{-1} and compares well to the results obtained from DFT calculations.

The Raman spectra from SCC-DFTB/FTTCF calculations have been compared to a variety of DFT methods with different basis sets and density functionals (first two spectra from the bottom) involved. With the new repulsive pair potentials for carbon and hydrogen atoms, most of the observed Raman active bands obtained from the SCC-DFTB/FTTCF calculations are in terms of frequency positions and intensities within the range of scattering observed from different high level DFT methods.

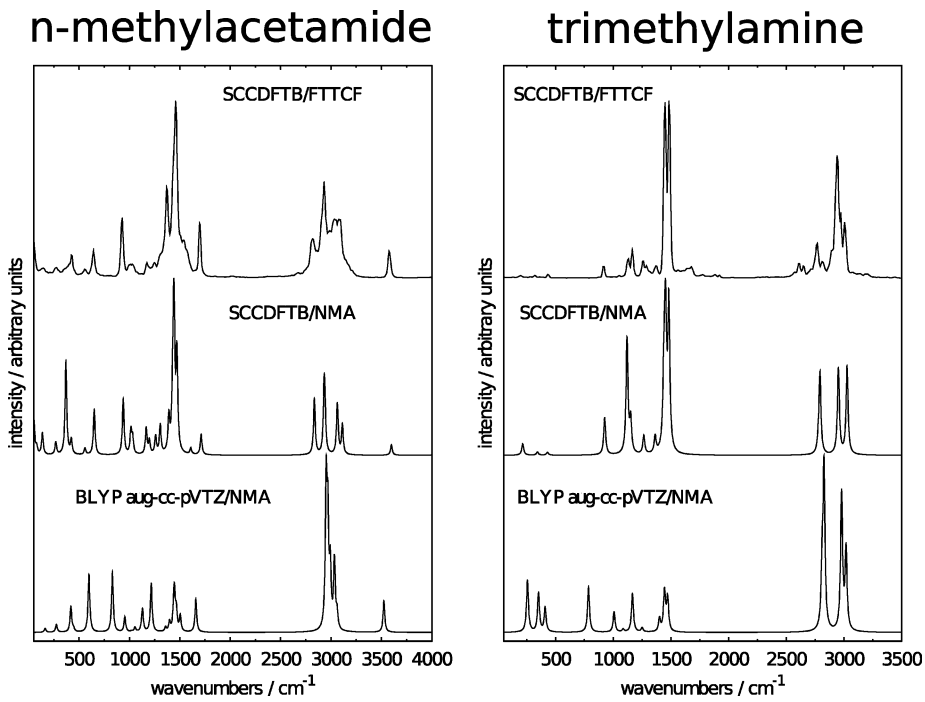

Figure 5. Raman spectra of 2 model compounds from different methodologies (FTTCF vs NMA) and different levels of theory (SCC-DFTB vs DFT) for comparison.

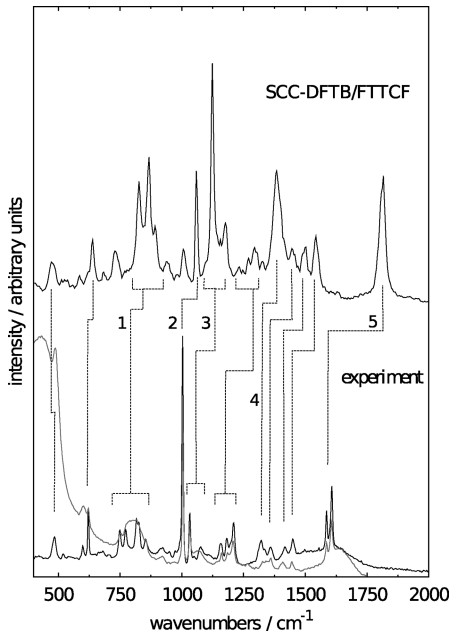

Figure 6. Raman spectra of phenylalanine in water from SCC-DFTB/FTTCF calculations (top) and experimental measurements (bottom) at 298 K (light gray) and 133 K (black), respectively. The top spectrum results from an average of 100 independent MD simulations. Dashed lines indicate the qualitative assignment of various spectral regions of which the numbered ones will be discussed in the text.

Further improvements concerning vibrational spectra of molecules containing various functional groups can be expected, since the work on repulsive potentials for other pairs of elements is in progress.

Depolarization Ratios. Depolarization ratios are important observables in Raman spectroscopy. By employing eqs 16 and 17, depolarization ratios for individual vibrational modes

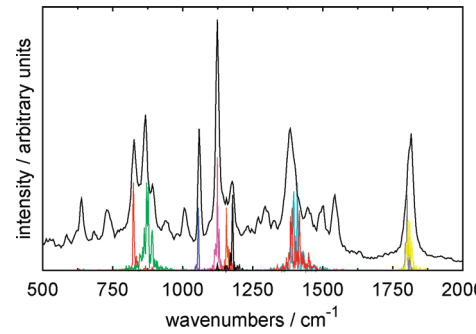

Figure 7. Calculated vibrational Raman spectrum of phenylalanine in solution from SCC-DFTB/FTTCF calculations averaged over 100 independent simulations (black). Colored spectra denote localized effective normal modes referring to spectral regions numbered in Figure 6. Relative intensities of effective normal modes were manually adjusted for a better illustration.

are accessible via the FTTCF formalism and shall be compared here to results obtained from SCC-DFTB and BLYP aug-cc-pVTZ calculations for polarized laser light following the NMA approach. The respective calculations have been performed for the most prominent Raman active modes (marked in Figures 3 and 4) of water and benzene since they show the most simple spectral pattern of all tested compounds, making an unambiguous mode assignment straightforward. Due to the high symmetry of benzene (D_{6h}), several of its Raman active modes are degenerated. Since they are equal concerning their vibrational frequencies and depolarization ratios, each couple of degenerated modes was treated as a single band and not counted twice.

In the case of water and for the symmetric and asymmetric O−H stretching vibrations (modes 2 and 3), the depolarization ratios from SCC-DFTB/FTTCF are in good agreement with the results obtained from BLYP aug-cc-pVTZ calcula-

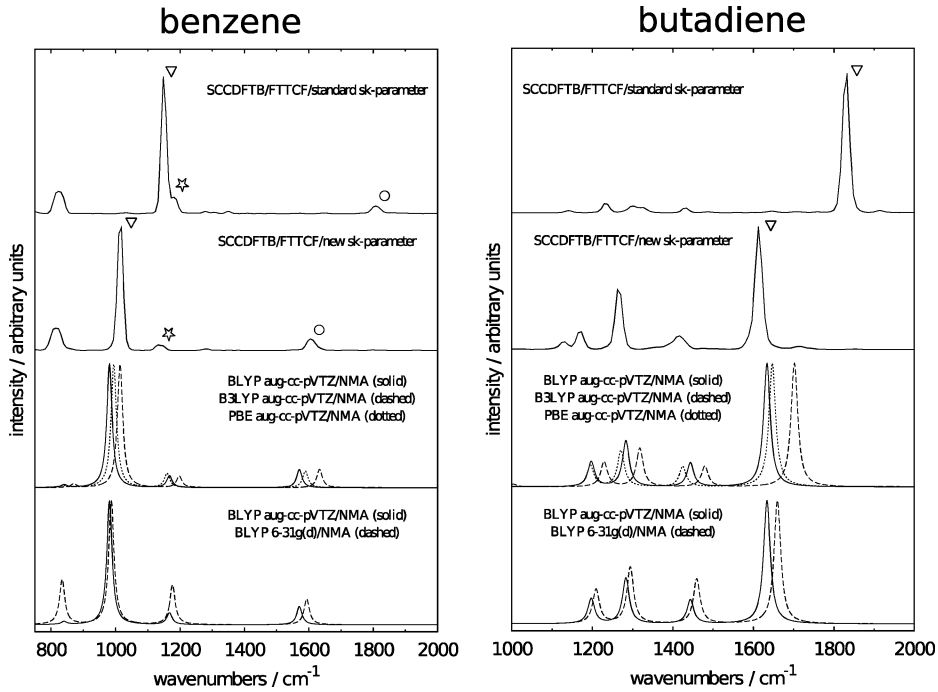

Figure 8. Calculated Raman spectra of butadiene and benzene in the gas phase. The first two spectra from the top of the graph result from SCC-DFTB/FTTCF calculations employing different Slater−Koster parameters for C−C and C−H repulsion. The symbols (triangle, star, circle) illustrate the frequency shift of several vibrational bands. The two graphs from the bottom show Raman spectra from DFT/NMA calculations done with different combinations of density functionals and basis sets.

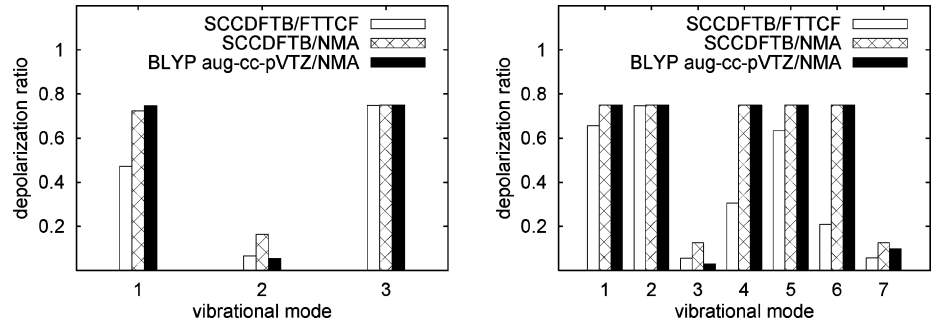

Figure 9. Calculated depolarization ratios for the most prominent Raman active modes (marked in Figures 3 and 4) of water (left) and benzene (right) from three different approaches as indicated in the graphs.

tions as illustrated in Figure 9. Concerning the H−O−H bending vibration (mode 1), the respective ratio from SCC-DFTB/FTTCF calculations is underestimated by approximately 25% in comparison to BLYP aug-cc-pVTZ, whereas SCC-DFTB/NMA perfectly agrees with BLYP aug-cc-pVTZ for modes 1 and 3, while mode 2 is overestimated by a factor of 3.

For benzene, depolarization ratios for all vibrational modes except number 3 are in a very good agreement comparing SCC-DFTB/NMA with BLYP aug-cc-pVTZ. The skeletal breathing mode number 3 is overestimated by the SCC-DFTB/NMA method. The SCC-DFTB/FTTCF calculations on the other hand reveal significantly reduced ratios for modes 4 and 6, referring to C−H wagging and C−H stretching vibrations, respectively.

Raman Spectra via DFT Polarizability Calculations. For all previously presented Raman spectra in the framework of our SCC-DFTB/FTTCF method, both the trajectory as well as the molecular polarizabilities were calculated at the SCC-DFTB level of theory. While SCC-DFTB is known to yield molecular structures in close agreement to higher level methods,[30] calculated polarizabilities are less accurate compared to high level methods due to the minimal basis set employed by SCC-DFTB. Therefore, as an alternative to our SCC-DFTB/FTTCF protocol, we figured out the performance of an approach, called the hybrid method, in which the trajectory and the polarizabilities were calculated on different levels of theory. Whereas the trajectory was still generated at the SCC-DFTB level, subsequent single-point calculations on snapshots of the trajectory were performed using higher level methods to obtain the molecular polarizabilitiy. We used such a hybrid approach to estimate the impact of the two parameters, i.e., the quality and size of the basis set for polarizability calculations on the one hand and the amount of phase space sampling on the other hand, on the overall spectral pattern.

Within our SCC-DFTB/FTTCF protocol using "on-the-fly" calculations of molecular polarizabilities, it is a computationally feasible task to generate a large ensemble of trajectories to guarantee for a sufficient phase pace sampling,

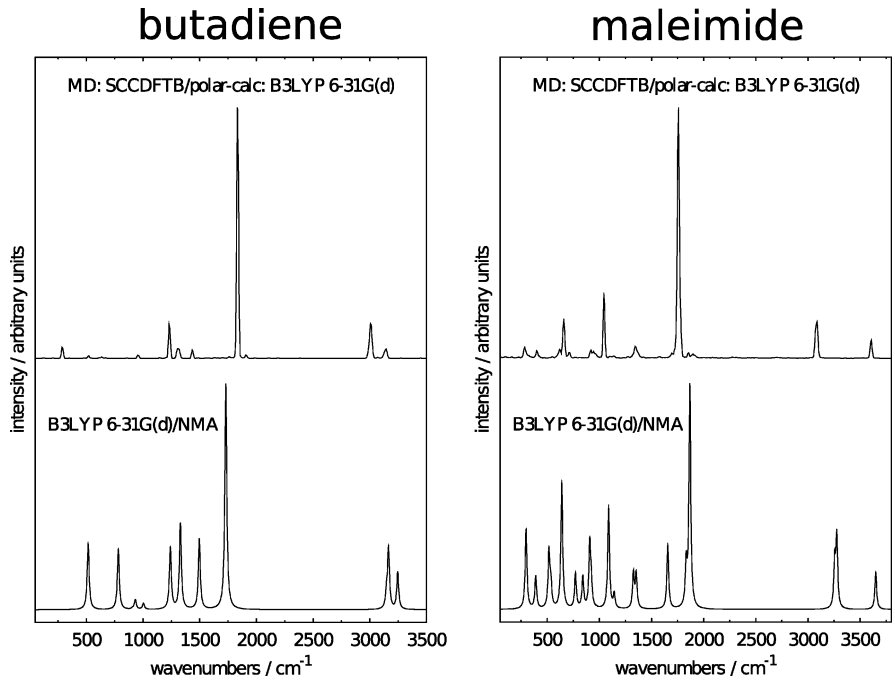

Figure 10. Comparison of vibrational Raman spectra of butadiene and maleimide (each from two different approaches as explained in the text). The top spectra result from an average of spectra from 10 independent simulations.

important for a reliable spectral pattern as already shown in Figure 2. This is not true any longer when employing density functional methods, such as DFT/B3LYP 6-31G(d), for the calculation of polarizabilities. For a butadiene molecule in the gas phase, the computational effort on a conventional desktop PC to generate a Raman spectrum out of a single trajectory (8192 simulation steps/0.5 fs time step/8 cm^{-1} spectral resolution) is as follows. Within the use of our SCC-DFTB/FTTCF protocol, the concurrent generation of the trajectory and the polarizabilities takes approximately 3 min on a single processing core. For the polarizability calculations itself, about 135 h of computing time is needed when employing DFT/B3LYP 6-31G(d) on a single CPU.

Due to the substantially increased computational cost, DFT/B3LYP 6-31G(d) single-point calculations were necessarily done on snapshots from a smaller ensemble of SCC-DFTB trajectories. Vibrational Raman spectra averaged over 10 single spectra were generated for two chosen model compounds. We used butadiene and maleimide as test cases, for which the resulting spectra are shown in Figure 10 and compared to a related B3LYP 6-31G(d)/NMA spectrum. Further on, we will denote the spectra derived from B3LYP 6-31G(d) single-point calculations of SCC-DFTB/FTTCF snapshots as B3LYP 6-31G(d)/polar.

Concerning butadiene, in the region above 1000 cm^{-1}, the B3LYP 6-31G(d)/polar calculations deliver a spectrum in good agreement to its B3LYP 6-31G(d)/NMA analogue, as shown in the left picture of Figure 10. The respective bands mainly refer to C−C and C−H stretching vibrations. However, the spectral pattern below 1000 cm^{-1} is not reproduced satisfactorily. Here, from the B3LYP 6-31G(d)/polar protocol, the intensities of the vibrational bands are strongly underestimated.

The observations made for butadiene are also valid for maleimide. The most pronounced C═O stretching vibration at 1833 cm^{-1} (B3LYP 6-31G(d)/NMA) as well as the C−H and N−H stretchings in the region above 3000 cm^{-1} are in a very good agreement compared to the B3YLP 6-31G(d)/NMA spectrum, whereas the SCC-DFTB/FTTCF calculations (Figure 3) show a very different spectral pattern in this region. The relative intensities below 1500 cm^{-1} are again not satisfactory within the B3LYP 6-31G(d)/polar approach compared to the results from NMA calculations, indicating that a sufficient sampling has not been achieved after an overall simulation time of 40 ps.

The fact that a 40 ps MD simulation is not necessarily sufficient in terms of the Raman intensity pattern is also illustrated for pentane in Figure 2. The spectrum colored in red (40 ps) differs in the region below 1500 cm^{-1} significantly from the ones in green (200 ps) and blue (400 ps). Nonella and co-workers[25] made a similar observation by comparing infrared Spectra from an FTTCF and NMA approach of *p*-benzoquinone in aqueous solution. They concluded that the 17.5 ps QM/MM MD simulation, although consuming considerable computational resources, was too short for the computation of a reliable vibrational spectrum.

Larger spectral deviations in the lower frequency region can also be observed by the comparison of the B3LYP 6-31G(d)/NMA spectra to their analogues (BLYP aug-cc-pVTZ) in Figures 3 and 4. The spectral pattern in this region is therefore highly sensitive to the level of theory on the one hand, as well as, in the FTTCF framework, to the number of independent trajectories used for spectral averaging on the other hand.

From our test calculations, we conclude that a hybrid approach to calculate Raman spectra via the FTTCF formalism yields results which are not clearly superior to the ones from our SCC-DFTB/FTTCF protocol, at least not below 1500 cm^{-1}. The generation of spectra via the hybrid approach is, however, computationally much more demanding and

therefore, in contrast to the SCC-DFTB/FTTCF method, restricted to rather small chemical system.

Summary and Conclusions

The SCC-DFTB electronic structure method has been extended for "on-the-fly" calculations of molecular polarizabilities via a finite electric field approach to access vibrational Raman spectra in the framework of the FTTCF formalism. FTTCF, in contrast to a standard Normal Mode Analysis, incorporates anharmonic motions as well as effects from the fluctuating environment at a finite temperature in the pattern of a vibrational spectrum.

The numeric differentiation step size for the calculation of polarizabilitiy tensor elements has been shown to be important, and a value of 0.005 au or larger was sufficient to avoid obvious numeric errors for the set of tested molecules. Furthermore, the stability of the employed numeric differentiation scheme has been verified to be sufficient by comparison to a more sophisticated method proposed by Magdó and co-workers.

Vibrational Raman spectra generated via the SCC-DFTB/FTTCF formalism were in good agreement for a set of 10 model compounds examined in the gas phase and compared to an NMA approach at the same (SCC-DFTB) and at a higher level of theory (BLYP aug-cc-pVTZ). Especially the consensus of vibrational frequencies for both methodologies (NMA vs FTTCF) at the SCC-DFTB level suggests that an integration time step of 0.5 fs is not too coarse and is a good compromise between accuracy and computational effort.

Compared to high level BLYP aug-cc-pVTZ calculations, the largest deviations concerning vibrational frequencies using SCC-DFTB were found for C−C stretchings in conjugated systems as well as C−H stretching modes in general. We were able to correct for such errors by employing a set of optimized repulsive pair potentials for SCC-DFTB. Recalculated Raman spectra for butadiene and benzene with the SCC-DFTB/FTTCF protocol were in very good agreement compared to high level desity functional methods. With ongoing progress in parametrizing SCC-DFTB repulsive potentials for other pairs of elements, further improvements in vibrational spectra calculations can be expected.

With respect to the relative band intensities compared to spectra from high level methods, none of the two techniques (neither NMA nor FTTCF) could be identified to be clearly superior to the other one in combination with the SCC-DFTB method. The situation varies rather strongly in terms of the specific compound and the spectral region under consideration. Concerning the FTTCF formalism, the quality of the spectral pattern depends strongly on the number of independent MD trajectories taken into account. For the model compounds tested in the gas phase, the spectral pattern was shown to adequately converge from an average of ∼50 single spectra, taken from a ∼200 ps simulation.

Besides the *in vacuo* calculations, Raman spectra of L-phenylalanine in solution in a QM/MM framework have been calculated using our SCC-DFTB/FTTCF implementation. Here, the same behavior concerning the intensity convergence through spectral averaging was observed as compared to the *in vacuo* calculations. However, this may not generally be true for a solute in a polar solvent. Strong intermolecular interactions, such as hydrogen bonds, could hamper an efficient sampling of phase space, making a larger number of independent trajectories necessary.

The overall spectral pattern of phenylalanine from SCC-DFTB/FTTCF calculations is in an acceptable agreement with the experiment, especially concerning the line shapes. Nevertheless, for several spectral regions, significant deviations concerning the relative band intensities were observed. Furthermore, due to well-known limitations of the SCC-DFTB approach, the C−C stretching modes in the benzene ring are strongly overestimated as indicated by a shift of $\sim$200 cm^{-1} to higher wavenumbers as compared to the experiment.

Depolarization ratios estimated via the FTTCF formalism were in a good overall agreement with the other two methodologies, although SCC-DFTB/NMA compares altogether better with ratios obtained from high level BLYP aug-cc-pVTZ calculations.

Finally, in this work, we tried to estimate whether an acceptable Raman spectrum in the FTTCF framework could be obtained by replacing our SCC-DFTB/FTTCF method with a hybrid approach. Here, the calculation of molecular polarizabilities on snapshots obtained from a SCC-DFTB trajectory were performed on the B3YLP 6-31G(d) level. In fact, the spectra resulting from insufficiently sampled 40 ps trajectories unsatisfactorily reproduced NMA calculations in the frequency range below 1500 cm^{-1}. The spectral pattern for two test cases in this region were characterized by a reduced number of intense Raman active bands compared to B3YLP 6-31G(d)/NMA. In our opinion, the hybrid approach, which incorporates high level methods for the calculation of polarizabilities, is therefore not an appropriate alternative to our SCC-DFTB/FTTCF method, since the quality of the Raman spectral pattern depends too strongly on appropriate phase space sampling. The more accurate polarizabilities obtained with B3YLP compared to SCC-DFTB cannot compensate for the effect of insufficient sampling. With the use of the highly efficient SCC-DFTB method, a sufficiently large ensemble of trajectories can be generated even for large molecules in solution, whereas for large chemical systems, the computational effort needed by a hybrid approach only allows for the generation of single short trajectories, from which the resulting Raman intensities are not representative.

Acknowledgment. The authors thank Rodolphe Vuilleumier for the supply of the software package for calculating effective normal modes. Financial support by the DFG Cluster of Excellence "UniCat" is acknowledged.

References

(1) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864.

(2) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

(3) Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular vibrations: The theory of infrared and Raman vibrational spectra*; Dover Publications: Mineola, NY, 1980; pp 11−33.

(4) McQuarrie, D. A. *Statistical Mechanics*; University Science Books: Herndon, VA, 2000; pp 467−506.

(5) Jolliffe, I. T. *Principal Component Analysis*; Springer: New York, 2002; p 299.

(6) Kitao, A.; Go, N. *Curr. Opin. Struct. Biol.* **1999**, *9*, 164.

(7) Balsera, M. A.; Wriggers, W.; Oono, Y.; Schulten, K. *J. Phys. Chem.* **1996**, *100*, 2567.

(8) Amadei, A.; Linssen, A. B.; Berendsen, H. J. *Proteins* **1993**, *17*, 412.

(9) Wheeler, R. A.; Dong, H.; Boesch, S. E. *ChemPhysChem* **2003**, *4*, 382.

(10) Gordon, R. G. *Adv. Magn. Reson.* **1968**, *3*, 1.

(11) Berens, P. H.; White, S. R.; Wilson, K. R. *J. Chem. Phys.* **1981**, *75*, 515.

(12) Gaigeot, M. P.; Vuilleumier, R.; Sprik, M.; Borgis, D. *J. Chem. Theory Comput.* **2005**, *1*, 772.

(13) Gaigeot, M. P.; Sprik, M. *J. Phys. Chem. B* **2003**, *107*, 10344.

(14) Hornicek, J.; Kaprálová, P.; Bour, P. *J. Chem. Phys.* **2007**, *127*, 4502.

(15) Pagliai, M.; Cavazzoni, C.; Cardini, G.; Erbacci, G.; Parrinello, M.; Schettino, V. *J. Chem. Phys.* **2008**, *128*, 224514.

(16) Putrino, A.; Parrinello, M. *Phys. Rev. Lett.* **2002**, *88*, 176401.

(17) Lammers, S.; Meuwly, M. *J. Phys. Chem. A* **2007**, *111*, 1638.

(18) Schultheis, V.; Reichold, R.; Schropp, B.; Tavan, P. *J. Phys. Chem. B* **2008**, *112*, 12217.

(19) Asvany, O.; Kumar, P. P.; Redlich, B.; Hegemann, I.; Schlemmer, S.; Marx, D. *Science* **2005**, *309*, 1219.

(20) Kaczmarek, A.; Shiga, M.; Marx, D. *J. Phys. Chem. A* **2009**, *113*, 1985.

(21) Padma Kumar, P.; Marx, D. *Phys. Chem. Chem. Phys.* **2006**, *8*, 573.

(22) Yu, H.; Cui, Q. *J. Chem. Phys.* **2007**, *127*, 234504.

(23) Schmitz, M.; Tavan, P. *J. Chem. Phys.* **2004**, *121*, 12247.

(24) Schmitz, M.; Tavan, P. *J. Chem. Phys.* **2004**, *121*, 12233.

(25) Nonella, M.; Mathias, G.; Tavan, P. *J. Phys. Chem. A* **2003**, *107*, 8638.

(26) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.

(27) Martinez, M.; Gaigeot, M. P.; Borgis, D.; Vuilleumier, R. *J. Chem. Phys.* **2006**, *125*, 144106.

(28) Vuilleumier, R. *Mol. Phys.* **2007**, *105*, 2857.

(29) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260.

(30) Otte, N.; Scholten, M.; Thiel, W. *J. Phys. Chem. A* **2007**, *111*, 5751.

(31) Elstner, M. *Theo. Chem. Acc.* **2006**, *116*, 316.

(32) Riccardi, D.; Schaefer, P.; Yang, Y.; Yu, H.; Ghosh, N.; Prat-Resina, X.; Koenig, P.; Li, G.; Xu, D.; Guo, H.; Elstner, M.; Cui, Q. *J. Phys. Chem. B* **2006**, *110*, 6458.

(33) Witek, H. A.; Irle, S.; Morokuma, K. *J. Chem. Phys.* **2004**, *121*, 5163.

(34) Witek, H. A.; Morokuma, K. *J. Comput. Chem.* **2004**, *25*, 1858.

(35) Witek, H. A.; Morokuma, K.; Stradomska, A. *J. Chem. Phys.* **2004**, *121*, 5171.

(36) Meuwly, M.; Müller, A.; Leutwyler, S. *Phys. Chem. Chem. Phys.* **2003**, *5*, 2663.

(37) Fouqueau, A.; Meuwly, M. *J. Chem. Phys.* **2005**, *123*, 244308.

(38) Phatak, P.; Ghosh, N.; Yu, H.; Cui, Q.; Elstner, M. *Proc. Nat. Acad. Sci.* **2008**, *105*, 19672.

(39) Perdew, J. P.; Burke, K.; Wang, Y. *Phys. Rev. B* **1996**, *54*, 16533.

(40) Elstner, M. Ph.D. Thesis, University of Paderborn, Paderborn, Germany, 1998.

(41) Koch, W.; Holthausen, M. C.; Holthausen, M. C. *A chemist's guide to density functional theory*, 2nd ed.; Wiley-Vch: Weinheim, Germany, 2000; pp 177−178.

(42) Abramowitz, M.; Stegun, I. A. *Handbook of mathematical functions with formulas, graphs, mathematical tables*; Courier Dover Publications: Mineola, NY, 1964; p 884.

(43) Van Dujineveldt-Van De Rijdt, J. G. C. M.; Van Dujineveldt, F. B. *J. Mol. Struct.* **1976**, *35*, 263.

(44) McCarthy, W. J.; Lapinski, L.; Nowak, M. J.; Adamowicz, L. *J. Chem. Phys.* **1998**, *108*, 10116.

(45) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D. *J. Comput. Chem.* **1983**, *4*, 187.

(46) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(47) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569.

(48) Lamoureux, G.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 3025.

(49) Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids*; Oxford University Press: New York, 1990; p 208.

(50) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical recipes: The art of scientific computing*, 3rd ed.; Cambridge University Press: Cambridge, U. K., 2007; p 605.

(51) Halls, M. D.; Schlegel, H. B. *J. Chem. Phys.* **1999**, *111*, 8819.

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

(53) Riley, K. E.; Op't Holt, B. T.; Merz, K. M., Jr. *J. Chem. Theory Comput.* **2007**, *3*, 407.

(54) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(55) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

(56) 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 03*, Revision C02; Gaussian, Inc.: Wallingford, CT, 2004.

(57) Pulay, P.; Fogarazi, G.; Pang, F.; Boggs, J. E. *J. Am. Chem. Soc.* **1979**, *101*, 2550.

(58) Magdó, I.; Nemeth, K.; Mark, F.; Hildebrandt, P.; Schaffner, K. *J. Phys. Chem. A* **1999**, *103*, 289.

(59) Porezag, D.; Frauenheim, T.; Köhler, T.; Seifert, G.; Kaschner, R. *Phys. Rev. B* **1995**, *51*, 12947.

(60) Wanko, M.; Hoffmann, M.; Frauenheim, T.; Elstner, M. *J. Comput.-Aided Mol. Des.* **2006**, *20*, 511.

(61) Małolepsza, E.; Witek, H. A.; Morokuma, K. *Chem. Phys. Lett.* **2005**, *412*, 237.

(62) Gaus, M.; Chou, C. P.; Witek, H.; Elstner, M. *J. Phys. Chem. A* **2009**, *113*, 10321.

CT900660X