Local Mode Analysis: Decoding IR Spectra by Visualizing Molecular Details

M. Massarczyk,^{†,⊥} T. Rudack,^{†,‡,§,||,⊥} J. Schlitter,^{*,†} J. Kuhne,[†] C. Kötting,[†] and K. Gerwert^{*,†,§}

[†]Department of Biophysics, Ruhr-University, 44801 Bochum, Germany

[‡]Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Champaign, Illinois 61801, United States

[§]Chinese Academy of Sciences-Max-Planck Partner Institute for Computational Biology (PICB), Shanghai Institutes for Biological Sciences (SIBS), Shanghai 200031, China

Supporting Information

ABSTRACT: Integration of experimental and computational approaches to investigate chemical reactions in proteins has proven to be very successful. Experimentally, time-resolved FTIR difference-spectroscopy monitors chemical reactions at atomic detail. To decode detailed structural information encoded in IR spectra, QM/MM calculations are performed. Here, we present a novel method which we call local mode analysis (LMA) for calculating IR spectra and assigning spectral IR-bands on the basis of movements of nuclei and partial charges from just a single QM/MM trajectory. Through LMA the decoding of IR spectra no longer requires several simulations or optimizations. The novel approach correlates the motions of atoms of a single



simulation with the corresponding IR bands and provides direct access to the structural information encoded in IR spectra. Either the contributions of a particular atom or atom group to the complete IR spectrum of the molecule are visualized, or an IR-band is selected to visualize the corresponding structural motions. Thus, LMA decodes the detailed information contained in IR spectra and provides an intuitive approach for structural biologists and biochemists. The unique feature of LMA is the bidirectional analysis connecting structural details to spectral features and vice versa spectral features to molecular motions.

INTRODUCTION

For a deep understanding of molecules, the knowledge of their precise atomic structure and dynamics is required.¹ In particular, to understand the reaction mechanism, structural insights at the sub-Ångstrom level are essential, since such changes are crucial for catalysis.² Changes in molecular structure are monitored at atomic resolution by vibrational spectroscopy. A change in bond length of 0.001 Å corresponds to a shift of the vibrational frequency of 1 cm^{-1} in the case of phosphates.³ However, the overlay of many geometrical changes and the coupling between vibrations prevent obtaining the structural details directly from the vibrational spectrum without a thorough theoretical analysis. Therefore, complementary vibrational spectroscopy and quantum mechanical calculations are frequently used in literature, in particular for small molecules. Already in 1968 Warshel showed an extensive comparison of calculated and experimental vibrational frequencies.⁴ Tavan and Schulten⁵ used theoretical spectra for the evaluation of the retinal structure by comparing them to the experimental work from Gerwert and Siebert.⁶

For the simulation of large biomolecules like proteins, it is necessary to partition the system into one region that is treated by accurate quantum mechanics (QM) and its surrounding that is treated by classical molecular mechanics (MM). In contrast to the early QM calculations, QM/MM simulations of proteins less frequently report on calculated vibrational modes. This is a pity, since experimental vibrational spectroscopy is an excellent method to complement multiscale calculations.

FTIR difference spectroscopy monitors absorbance changes of all groups that change during protein function, while the unchanged background absorbance is canceled out. Thus, a QM/MM calculation with only the reactive part of the protein treated by QM reflects the FTIR spectrum effectively.^{7–11} Time-resolved FTIR has recently revealed reaction steps and accompanying structural transitions of important reaction cycles or courses in biomolecules like enzymes down to the nanosecond scale. For example, the molecular reaction mechanism of the light-driven proton pump bacteriorhodopsin and the GTPase mechanism of the Ras protein are elucidated in molecular detail with high spatial-temporal resolution.^{11,12}

One reason for the omission of vibrational analysis in the current literature is that for proteins at room temperature many

Special Issue: Klaus Schulten Memorial Issue

Received:September 15, 2016Revised:January 16, 2017Published:January 16, 2017

conformations are populated. In turn for these systems a single geometry optimization with subsequent normal-mode analysis¹³ cannot describe the experimental spectrum adequately. Instead normal-mode analysis (NMA) needs to be performed for a sample of configurations selected from a MM simulation, which are then refined by a QM/MM simulation.^{9,10} This combination allows evaluation of long, classical trajectories that reflect heterogeneous broadening, but simplifies molecular motions to harmonic vibrations. Further, depending on the complexity of the system NMA for a large sample of configurations might be needed to reflect the experimentally observed heterogeneous broadening.¹⁴

Another approach to obtain the vibrational spectrum, which is more computer time-consuming but reflects better the heterogeneous broadening, is either by a direct calculation from the MD simulation via the Fourier transform (FT) of the total electric dipole moment or directly as the FT of its autocorrelation function. Both of these approaches lead to the same spectrum¹⁵ and were applied to several molecular systems including proteins.^{7,16–18} However, the assignment of IR-bands to nuclear motions is not as straightforward as in normal-mode analysis. While for the latter, isotopic bandshifts can be calculated without regress to quantum mechanics, for FT spectra the time-consuming QM/MM simulation must be repeated to reveal an isotopic effect. Therefore, usually only one complete spectrum without any assignments is calculated by this method. However, the IR-band assignments remain a prerequisite for the effective evaluation of experimental spectra. Recently, two methods have been developed that aim for IRband assignment through computational models by extraction of effective normal modes from the trajectory¹⁹ or by analysis of the vibrational density of states (VDOS).^{20,21}

Here we report on a direct approach for interpreting theoretical IR spectra based on just a single QM/MM simulation. The novel method called local mode analysis (LMA) combines the advantages of both normal-mode analysis and FT of the total electric dipole moment, namely reflecting heterogeneous broadening of the spectrum while maintaining a straightforward way of IR-band assignment to local vibrational modes. LMA achieves this combination by following two approaches: Molecular vibrations can be attributed to the spectrum by local mode decomposition (LMD) or vice versa absorbance bands can be attributed to vibrations by a band-pass filter. By applying LMA to bulk water and solvated trans-Nmethylacetamide (t-NMA) we prove that it is possible to localize and visualize spectral features. Thereby, the FTIR difference spectra can be decoded and the structural details are revealed and visualized.

THEORETICAL APPROACH

Why Do We Need LMA? Experimental IR spectra are very structured and bear plenty of detailed information about molecular vibrations that is not immediately accessible. This also applies in principal to theoretical spectra derived from QM/MM simulations because they result from the total electric dipole moment including smeared-out electron distribution. Interpretation of IR spectra always requires localization of spectral features in two respects: first, to identify the impact of a molecular group on the spectrum, and second, to find the spatial origin of a spectral feature in the molecule under investigation. Experimentally both tasks can be solved by repeated isotopic labeling measurements that can be tedious, time-consuming, and expensive. Through LMA each of the

tasks is solved separately using two economical approaches: LMD (Figure 1) and band-pass filter (Figure 2).



Figure 1. Local mode decomposition (LMD). LMD visualizes the contribution of structural motions of single atoms or groups to the total IR spectrum which is illustrated at the example of solvated trans-N-methylacetamide (t-NMA). Employing Fourier transformation (FT) to the autocorrelation function of the total wave function dipole moment $\boldsymbol{\mu}_{\text{OM}}(t)$ (top) is a common method to calculate IR absorption spectra (A). LMD however calculates the atomic moments $\mathbf{s}_k(t) =$ $\mathbf{q}_{k}(t) \cdot \mathbf{r}_{k}(t)$ for each atom (B) as the product of positions $\mathbf{r}_{k}(t)$ and partial charges $\mathbf{q}_k(t)$ which sum up to the total dipole moment $\mathbf{m}_{PC}(t)$ (C), which allows for calculating partial spectra for atoms, groups or vibrational modes, to localize their influence in the spectrum (middle). More detailed partial spectra of t-NMA can be found in Figure 6. Comparing the spectrum based on $\mu_{OM}(t)$ (A) with the one based on $\mathbf{m}_{PC}(t)$ (C) is used as quality check to prove that the used partial charge model (here Mulliken) and positions out of the QM calculation is an equally good method to obtain absorption spectra as using directly the quantum mechanically calculated dipole moment $\mu_{OM}(t)$.

LMA enables extraction of desired information from the QM trajectory of a single simulation. This is due to the fact that the QM dipole moment can be approximated by means of positions and partial charges calculated from the original trajectory, which yields practically the same spectrum. As a consequence, the calculation of partial spectra of atoms or groups of atoms by LMD allows for easily assigning the contributions of each atom to an IR-band and giving suggestions for targeted isotopic labeling or mutagenesis experiments to validate the hypothesis coming from the computational model. Therefore, LMD replaces expensive human work time by cheaper computational time.

Conversely to LMD, the use of a band-pass filter attributes to a selected IR-band the underlying localized molecular vibrations. They are visualized and localized as part of the dynamics in the form of animated local vibrational modes, by extracting only the atom motions out of the QM/MM trajectory that are related to the selected IR-band. Therefore,



Figure 2. Band-pass filter. The band-pass filter visualizes spectral features in the structural motion which is illustrated at the example of solvated *trans-N*-methylacetamide. Through reverse Fourier transformation (FT) any specific IR-band can be attributed to the underlying molecular vibration by isolating the spectral region of interest in the spectrum and obtaining the frequency-restricted motion. This motion is revealed by the change in the position vector $r_k(t)$ of all N atoms comprising the system. Exemplarily shown are the changes of the x, y, and z coordinates over time of the nitrogen atom (blue), the carbon atom (light green) of the methyl group bound to the nitrogen atom, and one of its protons (gray). The resulting frequency-restricted trajectory is visualized by animated molecular graphics and is comparable to the visualization within normal-mode analysis.

the band-pass filter is a direct way to decode dynamical and structural information from experimental spectra via a computational model. For a broader scientific audience of structural biologists, visualization of the underlying structural dynamics offers an insight into complex experimental spectra which otherwise can only be interpreted by spectroscopy experts.

Local Mode Decomposition. The electromagnetic absorption spectrum of a fluctuating dipole moment $\mu(t)$ is known from the literature^{15,18,22} by

$$A_{\rm QM}(\omega) \propto \omega^2 F_{\omega}[\langle \boldsymbol{\mu}(0) \cdot \boldsymbol{\mu}(t) \rangle] \tag{1}$$

where the prefactor ω^2 restores detailed balance. The angular frequency ω is connected with the wavenumber λ by $\lambda^{-1} = \frac{\omega}{2\pi c}$, where *c* is the velocity of light. F_{ω} is the Fourier transformation (FT) written as a functional applied to the auto correlation function. $\boldsymbol{\mu}(t) \equiv \boldsymbol{\mu}_{\text{QM}}(t)$ is the total wave function dipole moment vector obtained as a sum of electronic and nuclear dipole moments of the QM region of a simulation trajectory. In the first step, the absorption spectrum A_{QM} is calculated for comparison with experiments and as a benchmark for the subsequent decomposition.

In order to enable the decomposition of the total dipole moment into single contributions of each atom to the total dipole moment we take advantage of a partial charge (PC) model by introducing an atomic moment for each atom of the QM region^{18,23,24}

$$\mathbf{s}_{k}(t) = \mathbf{q}_{k}(t) \cdot \mathbf{r}_{k}(t)$$
⁽²⁾

combining partial charges $\mathbf{q}_k(t)$ and positions $\mathbf{r}_k(t)$. We approximate the dipole moment by the sum $\mathbf{m}(t) \equiv \mathbf{m}_{PC}(t)$ over all N atoms of all atomic moments $\mathbf{s}_k(t)$ as

$$\boldsymbol{\mu}(t) \approx m(t) = \sum_{k} \mathbf{s}_{k}(t)$$
(3)

Using the Wiener-Khintchin theorem absorption is written

$$A_{PC}(\omega) \propto \omega^{2} F_{\omega} \left[\left\langle \mathbf{m}(0) \cdot \mathbf{m}(t) \right\rangle \right]$$
$$\propto \omega^{2} \mathbf{m}(\omega) \cdot \mathbf{m}^{*}(\omega) \qquad (4)$$
$$= \sum_{kl} \underbrace{\omega^{2} \mathbf{s}_{k}(\omega) \cdot \mathbf{s}^{*}_{l}(\omega)}_{S_{kl}}.$$

Here, $\mathbf{m}(\omega)$ and $\mathbf{s}_k(\omega)$ are the Fourier transforms of $\mathbf{m}(t)$ and $\mathbf{s}_{k}(t)$, respectively. $A_{PC}(\omega)$ in eq 4 provides already the total spectrum of the QM region based on the PC model. As the total spectrum depends on the quality of the used charge model the latter needs to be checked by comparison with $A_{OM}(\omega)$. The challenge is to get the decomposition of $A_{\rm PC}(\omega)$ into the contributions of each atom. The double sum in eq 4, which reveals the nonadditive, interference-based structure of absorption, needs to be converted into a simple sum of atomic contributions. We will simplify and approximately diagonalize the spectral matrix S_{kl} by the following steps. First, the imaginary part of matrix elements can be neglected since the interesting sum in eq 4 is by definition real. Second, all fluctuation of $\mathbf{s}_{k}(t)$ that do not contribute to dipole absorption, the IR-inactive modes, can be eliminated by projection onto the total dipole moment,

$$\mathbf{s}_{\mathbf{k}}(\omega) \to \hat{\mathbf{s}}_{\mathbf{k}}(\omega) = \mathbf{s}_{\mathbf{k}}(\omega) \cdot \mathbf{m}(\omega) |\mathbf{m}(\omega)|^{-2} \mathbf{m}(\omega)$$
 (5)

Replacing the $\mathbf{s}_k(\omega)$ by $\mathbf{\hat{s}}_k(\omega)$ would leave the absorption in eq 4 unchanged. However, to tackle the remaining problem of the nonuniform sign of nondiagonal elements we apply the modified projection procedure which also leaves absorption unchanged, namely

$$\mathbf{s}_{\mathbf{k}}(\omega) \to \hat{\mathbf{s}}_{\mathbf{k}}(\omega) = \alpha(\omega) |\mathbf{s}_{\mathbf{k}}(\omega) \cdot \mathbf{m}(\omega)| |\mathbf{m}(\omega)|^{-2} \mathbf{m}(\omega)$$
(6)

with a prefactor $\alpha(\omega)$ chosen such that the sum of all atomic moments $\hat{\mathbf{s}}_{\mathbf{k}}(\omega)$ is still the dipole moment $\mathbf{m}(\omega)$ like before.

The resulting spectral matrix $\hat{S}_{kl} = \text{Re}(\hat{s}_k \cdot \hat{s}_l^*) \ge 0$ has a simpler and symmetric shape with only real, non-negative elements.

Diagonal elements \hat{S}_{kk} originate from atom k alone, whereas nondiagonal \hat{S}_{kl} elements are due to interference between atoms k and l. Here we regard the nondiagonal elements approximately by distributing each $\hat{S}_{kl} = \hat{S}_{lk}$ among the two interfering atoms as follows

$$\begin{split} \tilde{S}_{kk} &= \hat{S}_{kk} + \hat{S}_{kl} \cdot \frac{\sqrt{\hat{S}_{kk}}}{\sqrt{\hat{S}_{kk}} + \sqrt{\hat{S}_{ll}}} \\ \tilde{S}_{ll} &= \hat{S}_{ll} + \hat{S}_{kl} \cdot \frac{\sqrt{\hat{S}_{ll}}}{\sqrt{\hat{S}_{kk}} + \sqrt{\hat{S}_{ll}}} \end{split}$$
(7)

and set $\tilde{S}_{kl} = 0$ for $k \neq l$.

The weights are chosen proportional to the strength of the respective atomic moment, $|\hat{s}_k| \approx \sqrt{\hat{S}_{kk}}$. The resulting absorption spectrum is

$$A_{\rm PC}(\omega) \propto \sum_{k=1}^{N} \left(\hat{S}_{kk}(\omega) + 2 \sum_{l \neq k} \hat{S}_{kl}(\omega) \cdot \frac{\sqrt{\hat{S}_{ll}(\omega)}}{\sqrt{\hat{S}_{kk}(\omega)} + \sqrt{\hat{S}_{ll}(\omega)}} \right)$$
$$= \sum_{k=1}^{N} \tilde{S}_{kk}(\omega) \tag{8}$$

The term $\omega^2 \tilde{S}_{kk}(\omega)$ is the contribution of atom k to the absorption spectrum at frequency ω . Note that the two heuristic transformations (eq 6) and (eq 7) did not change the absorption spectrum $A_{PC}(\omega)$ given by eq 4. They merely represent a procedure that yields an approximate decomposition (eq 8) into atomic contributions that is guided by our interpretation of calculated correlations and has to be validated by application. For each atom or group it indicates its contribution everywhere in the absorption spectrum (Figure 1).

Beyond the assignment of atoms to observed absorption peaks, spectroscopy aims at assigning particular modes to investigate how they correlate with function. We propose the following scheme for defining and extracting vibrational modes from the IR-active atomic moments $\tilde{\mathbf{s}}_k(\omega)$. Given a particular vibrational mode, we calculate a function $S_{\text{mode}}(\omega)$ that represents its contribution to the absorption spectrum.

For defining modes, we consider triples of subsequently bound atoms A, B, and C as shown in Figure 3 and in particular the bond vector $\mathbf{d}_{AB}(t) = \mathbf{r}_A(t) - \mathbf{r}_B(t)$ connecting atom A with B. The change of \mathbf{d}_{AB} along \mathbf{d}_{AB} itself is defined as stretching vibrational mode. Bending vibrational modes occur when a third atom C is bound to A or B, which defines a plane ABC. The bond vector \mathbf{d}_{AB} , the normal vector $\mathbf{n}_{ABC}(t)$ of the plane ABC and $\mathbf{a}_{ABC}(t) = \mathbf{d}_{AB}(t) \times \mathbf{n}_{ABC}$ span an orthogonal axis system { $\mathbf{d}_{AB}, \mathbf{n}_{ABC}, \mathbf{a}_{ABC}$ }. The in-plane bending is defined as a change of \mathbf{d}_{AB} along \mathbf{a}_{AB} , and out-of-plane bending as a change of \mathbf{d}_{AB} along \mathbf{n}_{ABC} . Note that the latter movement is torsion around the axis BC.

The definition of mode directions depends on which of the atoms contributing to the mode is defined as anchor atom. To label one atom as anchor atom we introduced a mass weighted ranking of all atoms contributing to the mode by summing up their own mass with the one from the first neighbors. The atom with the highest mass is defined as anchor atom, which is atom C in the above-described case.

Finally, partial mode spectra are obtained analogous to the decomposition of the absorption spectrum into atomic spectra

D



Figure 3. Definition of vibrational modes within LMA. Modes are here defined by motions of atoms relative to the plane (yellow) spanned by a triple of subsequently bound atoms A, B, and C. At the stretching mode (orange), atom A follows the direction of the bond AB. Considering the first bending-mode direction (cyan), atom A moves in plane perpendicular to the bond AB. Considering the second bending-mode (purple), atom A moves perpendicularly out of the plane ABC. Thus, motions of the distance vector of atoms AB are attributed to vibrational modes. The contribution of each mode to absorption is assumed to be proportional to its frequency-dependent spectral power that is calculated from velocities.

as described in eq 8. We start from the velocities $\dot{\mathbf{d}}_{AB}(t)$ of bond vectors $\mathbf{d}_{AB}(t)$ and calculate the projections onto the mode directions, $\mathbf{P}_{mode} \cdot \dot{\mathbf{d}}_{AB}(t)$, i.e., three components \mathbf{P}_{mode} along the above-defined directions \mathbf{d} , \mathbf{n} , \mathbf{a} . Their Fourier transforms yield partial mode spectra

$$\sigma_{\text{mode}}(\omega) = \sum_{\text{mode}} |F_{\omega}[\mathbf{P}_{\text{mode}} \cdot \dot{\mathbf{d}}_{AB}(t)]|^2$$
(12)

that have to be normalized appropriately. For this purpose we make the assumption that all modes of bonds involving Atom A with its contribution c_{A} , thus forming the set M_{A} , contribute equally to the spectrum of atom A,

$$\sum_{\text{modes}\in M_{A}} c_{A}(\omega)\sigma_{\text{mode}}(\omega) = \tilde{S}_{AA}(\omega)$$
(13)

For the sake of simplicity we here use the same symbol for atoms and atom numbers. Consequentially, the full spectrum of a particular mode defined by the mode type and the two bound atoms is

$$S_{\text{mode}}(\omega) = (c_{\text{A}}(\omega) + c_{\text{B}}(\omega))\sigma_{\text{mode}}(\omega)$$
(14)

Eq 13 shows that the mode spectra sum up to the same value as the atomic spectra. Using eq 8 one finds

$$A_{\rm PC}(\omega) \propto \sum_{\rm A} \tilde{S}_{\rm AA}(\omega) = \sum_{\rm modes} S_{\rm mode}(\omega)$$
 (15)

Eq 14 represents the division of the spectrum into partial modes spectra. Note that our definition of vibrational modes is clear-cut, but the decomposition remains an approximation for assigning modes to spectral features.

Band-Pass Filter. The band-pass filter addresses the reverse question as to which atoms and vibrational modes contribute to a given IR-band observed in the spectrum. It does not simply invert and recover the same data. The idea is to select the frequency band of interest, apply reverse Fourier transformation and visualize the resulting frequency-restricted trajectory, which is done best by animated molecular graphics.

Starting with the Fourier transforms of the position vectors $\mathbf{r}_{\mathbf{k}}(\omega) = F_{\omega}[\mathbf{r}_{\mathbf{k}}(t)]$ of all N atomic trajectories we want to look

Article

Table 1. Simulation System Setup for Bulk Water and t-NMA^a

system	QM region	MM _{QM} region	MM region	equilibration (MM) time/ns	production (QM/MM) time/ps	analyzed frames
bulk water	27 H_2O molecules (81 atoms)	~1350 atoms	1486 H ₂ O molecules (4458 atoms)	5	14	56 000
solvated <i>t-</i> NMA	1 <i>t</i> -NMA + 23 H_2O molecules (81 atoms)	~1506 atoms	850 H ₂ O molecules (2550 atoms)	5	14	56 000

^aGiven are the number of atoms within the differently treated simulation region of the hybrid QM/MM simulations as well as the duration of the simulation time. The number of frames that deliver data points for the calculation of the theoretical spectra are given in the last column.

only at the motion within the trajectory of the complete system that evokes in the frequency band $[\omega_1, \omega_2]$. By using a boxcar shape function

$$B(\omega_1, \omega_2) = \begin{cases} 1 \text{ for } \omega_1 < \omega < \omega_2 \\ 0 \text{ elsewhere} \end{cases}$$
(9)

as band-pass filter while applying reverse transformation F_t^{-1} , we extract two fractions from $\mathbf{r}_k(\omega)$, the slow motion up to a frequency ω_c far below the IR regime

$$\mathbf{r}_{k}^{(s)}(t) = F_{t}^{-1}[B(0,\,\omega_{s})\cdot\mathbf{r}_{k}(\omega)]$$
(10)

and the IR-relevant motion (Figure 2 middle)

$$\mathbf{r}_{\mathbf{k}}^{(\mathrm{IR})}(t) = F_t^{-1}[B(\omega_1, \omega_2) \cdot \mathbf{r}_{\mathbf{k}}(\omega)]$$
(11)

The resulting trajectory $\{\mathbf{r}_{k}^{(s)}(t) + \mathbf{r}_{k}^{(IR)}(t), k = 1 \cdots N\}$ shows fast vibrational motions of all atoms, the molecules moving slowly due to diffusion and conformational changes which are reflected by the slow-motion band.

The same trajectory is obtained no matter if the frequency band $[\omega_1, \omega_2]$ is selected in the IR spectrum - as proposed here - or in the fluctuation spectrum of spatial displacements as the band-structure of both spectral types is almost identical despite remarkable differences in intensity.

Simulation Setup. QM/MM²⁵ simulations were carried out for bulk water and solvated *t*-NMA as summarized in Table 1. For the QM part of the simulations we employed Gaussian 03^{26} at the level of B3LYP $(6-31G^*)^{27}$ for the MM part we used Gromacs-4.0.7²⁸ with the OPLS-AA force field²⁹ and the TIP4P water model. The bulk water simulation system contains 27 water molecules within the QM region surrounded by 1486 MM treated water molecules of which all MM atoms that are in radius of 10 Å within a QM atom are considered as $\mathrm{MM}_{\mathrm{OM}}$ region and contribute as partial charge points to the QM calculation. The whole simulation system is a cubic box of 3.7 nm edge length. The solvated t-NMA simulation system comprises t-NMA surrounded by 23 water molecules of the first solvation sphere within the QM region and 850 MM treated water molecules of which approximately 502 water molecules are part of the MM_{QM}-region. The whole simulation system is a cubic box of 3 nm edge length. In order to equilibrate the two systems a 5 ns MM simulation with 1 fs stepsize was performed followed by a QM/MM production run of 14 ps with 0.25 fs stepsize. The simulations were performed using NpT at 298.15 K and 1.0 bar using thermal coupling to the Nose-Hoover thermostat^{30,31} and isotopic pressure coupling by a Parrinello-Rahman barostat.^{32,33} The cutoff was set to 10 Å and periodic boundary conditions were applied. The point charges of the QM region, which were used to calculate the atomic moments are Mulliken charges, which are calculated for every step of the simulation.

Experimental Spectra. *t*-NMA was purchased from Sigma-Aldrich (Germany). A 100 mM solution compared to pure water as the reference was measured by a modified IFS88 spectrometer (Bruker, Germany) equipped with a LN₂-cooled MCT detector and a diamond μ -ATR (attenuated total reflection) cell. The wavenumber resolution was 2 cm⁻¹, zero-filling factor 4, Blackmann-Harris 3-term apodization, a Ge-filter, and a folding limit of 7899 cm⁻¹ was used. For the water spectrum, a thin film of deionized water (thickness approximately 1–2 μ m) in between two CaF₂ windows was measured in transmission using a Bruker Vertex 80V Spectrometer equipped with a LN₂-cooled MCT detector. A wavenumber resolution of 4 cm⁻¹, a zero filling factor of 2, and folding limit of 3949 cm⁻¹ were used.

The experimental carbon dioxide (CO_2) spectrum in the supplement is taken from Dierenfeldt's publication.³⁴

Data Processing. FT is performed using the Fast Fourier Transformation (FFT) algorithm.³⁵ Differently to the common approach in IR spectroscopy we do not use window functions prior to the FT.

In order to remove translations, trajectories of atomic coordinates were fitted by superimposing the center of mass of all atoms. No fitting was applied to rotation.

Wavenumbers in all theoretical spectra were scaled by 0.96 to correct the systematic error of the B3LYP (6-31G*) functional. 27,36

Normal mode spectra of the QM optimized CO_2 molecule was obtained with the QM program suite $ORCA^{37}$ employing the B3LYP functional and the 6-31G* basis set.

RESULTS

In this section we discuss and validate our here developed method called local mode analyses (LMA) and give application examples. In order to prove the economical applicability and reveal the benefits of the LMA for experimentalists as well as theoreticians we applied it to localize the vibrational modes and spectral features for two case examples: bulk water and solvated trans-N-methylacetamide (t-NMA). Employing partial charges and coordinates to obtain absorption spectra is validated for bulk water and its LMD is shown. LMD enables to assign to vibrational modes or even single atom movements out of the trajectory their respective contribution to the total absorption spectrum by calculating partial spectra for subgroups of the QM region. Besides the assignment of vibrational modes for the t-NMA absorption spectra we also take advantage of the LMA to obtain, out of only one QM/MM simulation trajectory, the absorption spectrum of solvated *t*-NMA with subtracted bulk water, which is a good approximation of the experimental spectrum of solvated *t*-NMA, as in the experiment bulk water is used as reference. The application of the band-pass filter, which gives the vibrational modes of the trajectory for a certain wavenumber region within the absorption spectrum, the IRband, is also demonstrated at the example of t-NMA. All discussed spectra in this section are absorption spectra and from here on referred to only as spectra.



Figure 4. Absorption spectra of bulk water. Comparison of the theoretical bulk water spectrum $A_{QM}(\omega)$ (blue) based on the total wave function dipole moment $\mu_{QM}(t)$, the theoretical bulk water spectrum $A_{PC}(\omega)$ (pink) based on the dipole moment $m_{PC}(t)$ derived from the quantum mechanically calculated partial charges and coordinates, and the experimental spectrum (black). All three spectra reveal the same pattern and IRband positions for $A_{QM}(\omega)$ are exactly the same as for $A_{PC}(\omega)$. The accordance of all three spectra proves that spectra calculations based on QM calculated Mulliken charges and coordinates are a valid approach and from here on all shown theoretical spectra are $A_{PC}(\omega)$ spectra. In this figure and all the following ones the intensities of the theoretical spectra are scaled to the highest peak of the experimental absorption spectrum.

Method Validation. The basic prerequisite for the LMA that allows for the spectral decomposition is to calculate the spectrum $A_{\rm PC}(\omega)$ based on the dipole moment $m_{\rm PC}(t)$ derived from the quantum mechanically calculated partial charges and coordinates according to eq 8. We first proved that employing Mulliken charges to derive $m_{\rm PC}(t)$ is an appropriate charge concept to calculate spectra by comparing in Figure 4 the theoretical spectrum $A_{\rm PC}(\omega)$ of bulk water with the $A_{\rm OM}(\omega)$ spectrum which is derived from the total wave function dipole moment $\mu_{QM}(t)$ according to eq 1. The experimental spectrum of bulk water is used as reference. Both theoretical absorption spectra reveal the same IR-band positions in accordance with the experiment. Thus, the approach to calculate spectra based on Mulliken charges and coordinates is validated. Note that other charge models than Mulliken can be used within LMA as well, e.g., GAPT charges³⁸ that have proved to be reliable for static charge distributions. More complex charge models might be necessary for investigation of chemical reactions. From here on every spectrum referred to as theoretical spectrum is an $A_{PC}(\omega)$ spectrum.

Besides the two experimentally and theoretically well-studied test cases of bulk water^{19,39–43} and *t*-NMA,^{20,44–47} which are described in the following paragraphs, we also employed LMA to CO_2 in vacuum to validate and prove the advances of LMA. The results for CO_2 can be found in Supporting Text ST1, Supporting Figure S1, and Supporting Movie SM1.

Local Mode Decomposition. In order to prove the advantages of LMD we separately calculated the partial spectra for the stretching $\nu(H_2O)$ and bending vibrational mode $\delta(H_2O)$ of the 27 water molecules within the quantum region of the bulk water trajectory. Figure 5 reveals that the IR-band assignment is in accordance with the experimental assignment and furthermore shows that both partial spectra sum up to the total theoretical spectrum.

Next, we applied LMD to calculate partial spectra for the vibrational modes of the QM region comprising *t*-NMA and the 23 water molecules of the first solvation shell out of the solvated *t*-NMA QM/MM trajectory. Summarized in Figure 6 is the comparison of the partial spectra of the local vibrational modes of *t*-NMA with the experimental IR-band assignment through isotopic labeling.

The partial spectrum of the vibrational modes of the CO group reveals an IR-band at 1612 cm^{-1} , which is in accordance



Article

Figure 5. LMD of the absorption spectrum of bulk water. LMD was employed to calculate separately the partial spectra of the stretching $\nu({\rm H_2O})$ (blue) and bending vibrational mode $\delta({\rm H_2O})$ (orange) of the 27 water molecules of the quantum region of a QM/MM trajectory of bulk water. These partial spectra of the two vibrational modes can be clearly assigned to the two main IR-bands of the total theoretical bulk water spectrum (pink) and are in accordance to the IR-band assignment in the experimental spectrum (black). The attribution of the structural motion corresponding to the two IR-bands extracted from the trajectory by the band-pass filter are provided in Supporting Movie SM2.

with the experimental assignment of ν (CO) (amide I) at 1618 cm⁻¹. However, the partial spectrum of the bending vibrational mode δ (NH) of the NH group also has its IR-band at the same wavenumber as ν (CO). The experimentally separate assigned IR-band for δ (NH) at 1580 cm⁻¹ (amide II) is in the theoretical spectra shoulder to which both, ν (CO) and δ (NH), contribute in the same manner. A large fraction of this IR-band originates from the couplings of the NH group with the neighboring water molecules.⁴⁸ It is experimentally shown that the IR-band at 1580 cm⁻¹ almost vanishes in other solvents than water,⁴⁹ which is a proof that this IR-band is mainly evoked by water molecule interaction.

Figure 6B shows the partial spectra of the bending vibrational modes $\delta(C1H_3)$ and $\delta(C2H_3)$ of the two methyl groups $C1H_3$ bound to the nitrogen atom of *t*-NMA and $C2H_3$ bound to the carboxyl carbon atom. The IR-bands at 1417 cm⁻¹, 1379 cm⁻¹, and 1316 cm⁻¹ are experimentally assigned as bending methyl vibrations. In accordance with the experiment LMD assigns and further distinguishes $\delta(C1H_3)$ to 1419 cm⁻¹ and $\delta(C2H_3)$ to



Figure 6. LMD of the absorption spectrum of *t*-NMA. Shown in A is the partial spectrum of the stretching ν (CO) and bending vibrational mode δ (CO) of the CO group as well as the bending vibrational mode δ (NH) of the NH group. Shown in B is the partial spectrum of the bending vibrational mode δ (C1H₃) of the first methyl group (C1, light green) and the bending vibrational mode δ (C2H₃) of the second methyl group (C2, dark green). The contribution of all other *t*-NMA atoms to the partial *t*-NMA spectrum (magenta) is shown in gray. The experimental spectrum (black) with IR-band assignments through isotopic labeling is given as reference. The contributions of the partial spectra of the vibrational modes to the *t*-NMA spectrum agree with the experimental IR-band assignment.



Figure 7. Band-pass filter. Highlighted are the five main IR-bands of the theoretical partial spectrum of *t*-NMA (magenta). The *t*-NMA structures in the boxes highlight the atoms of only the main vibrational mode that contributes most with their motion within the QM/MM trajectory to the IR-band with the color of the box frame. Supporting Movies SM4 to SM8 display the motion of each local vibrational mode extracted from the QM/MM trajectory by the band-pass filter for each of the five IR-bands. The experimental spectrum of *t*-NMA (black) is given as reference.

1384 cm⁻¹ as well as 1334 cm⁻¹. Furthermore, δ (C1H₃) contributes to the IR-band at 1612 cm⁻¹, which is also in accordance with the experiment, as the isotopic labeling of the methyl groups effects a shift of this IR-band.⁴⁸ In addition, LMD reveals a contribution of both δ (C1H₃) and δ (NH) to the IR-band at 1153 cm⁻¹, which is in accordance with the experimentally assigned rocking mode r(NCH₃) at 1165 cm⁻¹.

All in all, LMD applied to bulk water and *t*-NMA provides partial spectra for local vibrational modes, which are in accordance with the experimental IR-band assignment. Thus, LMD is proven to allow for an easy assignment of the contributions of each atom to an IR-band and is capable of giving suggestions for targeted isotopic labeling, which can replace expensive human work time for tedious trial and error experiments by cheaper computational time.

Solvation Effects. Besides the assignment of vibrational modes we also take advantage of the LMD to consider solvation effects at the quantum level. LMD enables inclusion of the solvent surrounding the substrate of interest at the quantum level while allowing to obtain the partial spectrum only for the substrate. Interactions with the solvent are taken into account when calculating the substrate spectrum, however, the vibrational modes of the substrate and the solvent can be separated. Including the influence of the solvents at the quantum level leads to a better accordance between the calculated and the experimental spectra as shown for the

example of *t*-NMA in Supporting Figure S4. The resulting substrate spectrum is comparable to the experimental spectrum as the solvent is usually used as reference and thus subtracted from the total spectrum. The agreement between the theoretical and experimental substrate spectrum is shown in Figures 6 and 7 for the example of *t*-NMA.

Band-Pass Filter. Another benefit of LMA is that it can not only assign partial spectra to local modes through LMD but also vice versa give the local vibrational modes or even rare events within the QM/MM trajectory for specific IR-bands through the band-pass filter. We applied the band-pass filter to extract the local vibrational modes shown in Supporting Movies SM4 to SM8 for the IR-bands of *t*-NMA, which are highlighted in Figure 7. These movies illustrate how the band-pass filter visualizes the structure and dynamics of the IR-band features obtained from LMD, as described above. For example Supporting Movie SM8 shows the overlay of the stretching vibrational mode $\nu(CO)$ and the bending vibrational mode $\delta(NH)$, which have the main contribution to the IR-band from 1650 to 1550 cm⁻¹ as indicated by the LMD shown in Figure 6A. Another example is the rocking of the NC1H₃ group for the IR-band at 1180 to 1130 cm⁻¹ which can be observed in Supporting Movie SM4 and which was indicated by the LMD of Figure 6A,B. All in all the extracted local vibrational modes prove that LMA is capable of visualizing the underlying structural dynamics and decode information hidden in complex experimental spectra and making them accessible also to nonspectroscopic experts.

CONCLUSIONS

We have developed and successfully applied LMA in an effort to further bridge the two powerful techniques of hybrid QM/ MM simulation and experimental time-resolved FTIR (trFTIR) spectra, which enables investigation of chemical reactions, in particular enzyme catalysis and related processes in proteins with the highest temporal and spatial resolution possible.

Calculation of IR spectra from simulation trajectories is theoretically well-established and has been demonstrated in many applications.^{7,18,21–24,47} LMA adds a combination of two methods for the bidirectional analysis of vibrational modes in space and spectra, which allows for interpretation of IR-bands and their localization in a molecule in order to gain detailed insights into the structural and dynamic processes underlying protein function. Hypothetical modes are assigned to spectral features by spectral decomposition and observed IR-bands are assigned to molecular vibrations by visualization after band-pass filtering. Through this bidirectional approach LMA is suited to answer the questions about protein function that arise when measuring IR spectra of molecular systems or trFTIR spectra of chemical reactions.

Like all existing theoretical IR spectroscopic techniques, our LMD is an approximation, but the applications to carbon dioxide, bulk water, and *trans-N*-methylacetamide have shown that the assignments are in accordance with previous ones determined both theoretically and experimentally. Moreover, our band-pass filtering is an independent, parameter-free method to verify or falsify IR-band assignments, and has confirmed the assignments for all the cases presented here. The application to these simple but commonly used cases to validate theoretical IR spectroscopic approaches serves as proof of principle for LMA.

It is recommended to use LMA only if the experimental and the theoretical IR spectra A_{OM} in eq 1 are in agreement, which

was shown for all systems studied here. More complex systems than the systems studied here likely need longer QM/MM trajectories in order to allow for sampling of the less frequently populated conformations to reproduce the experimental data correctly. However, all information about the system is determined from one simulation trajectory by application of straightforward numerical processing, which makes LMA an economical way to evaluate spectra and to initiate further experiments to test hypotheses about underlying protein function. In case IR-band assignment is ambiguous LMA helps to perform targeted experiments to minimize experimental efforts, such as repeated isotopic replacement or random mutagenesis studies.

Band-pass filtering opens a new, intuitive access to the dynamics of molecules up to chemical reactions as it visualizes vibrational modes belonging to observed IR-bands. At application of trFTIR even time-windows of chemical reactions can be analyzed by LMA. The spirit of visualizing spectral information through structural models follows the same philosophy as X-ray crystallography that uses computational models to assign structure to diffraction data, which was the origin of modern structural biology and biochemistry. However, LMA goes one step further and reveals also the dynamics of the structure. Thus, LMA assists not only experts in the field of spectroscopy to better interpret their spectral data but also makes IR-spectroscopic results readily available to structural chemistry and biology.

ASSOCIATED CONTENT

S Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.6b09343.

Eight movies visualize the local vibrational modes and the local mode decomposition of CO_2 in vacuum, H_2O in bulk water, and solvated t-NMA; description of the application of LMA for CO_2 in vacuum and an explanation how spectral matrices are used to decompose the total spectrum into the contributions of each single atom; Four figures show the comparison between experimental and theoretical IR spectra for CO_2 , the spectral matrices of CO_2 , and the impact on the theoretical IR spectrum for t-NMA by the QM treatment of the first solvation sphere (PDF)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: juergen.schlitter@rub.de; Phone: +49 234 3225753. *E-mail: gerwert@bph.rub.de; Phone: +49 234 3224461.

ORCID

- T. Rudack: 0000-0003-2693-9561
- C. Kötting: 0000-0002-3599-9657

Present Address

^{II}(T.R.) Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign, United States

Author Contributions

[⊥]M.M. and T.R. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Sven Brucker for providing spectral data from his FTIR spectroscopic measurements for solvated *t*-NMA; Ryan McGreevy for critical reading; and Rafael Bernardi for advice for figure design. We acknowledge the Deutsche Forschungsgemeinschaft (SFB642).

REFERENCES

(1) Perilla, J. R.; Goh, B. C.; Cassidy, C. K.; Liu, B.; Bernardi, R. C.; Rudack, T.; Yu, H.; Wu, Z.; Schulten, K. Molecular Dynamics Simulations of Large Macromolecular Complexes. *Curr. Opin. Struct. Biol.* **2015**, *31*, 64–74.

(2) Rudack, T.; Xia, F.; Schlitter, J.; Kötting, C.; Gerwert, K. Ras and GTPase-Activating Protein (GAP) Drive GTP into a Precatalytic State as Revealed by Combining FTIR and Biomolecular Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (38), 15295–300.

(3) Deng, H.; Wang, J. H.; Callender, R.; Ray, W. J. Relationship between Bond Stretching Frequencies and Internal Bonding for [O-16(4)]- and [O-18(4)]phosphates in Aqueous Solution. *J. Phys. Chem. B* **1998**, *102* (18), 3617–3623.

(4) Lifson, S.; Warshel, A. Consistent Force Field for Calculations of Conformations, Vibrational Spectra, and Enthalpies of Cycloalkane and N-Alkane Molecules. *J. Chem. Phys.* **1968**, *49* (11), 5116.

(5) Grobjean, M. F.; Tavan, P.; Schulten, K. Can Normal Mode Analysis Reveal the Geometry of the L 50 Chromophore of Bacteriorhodopsin? *Eur. Biophys. J.* **1989**, *16*, 341–349.

(6) Gerwert, K.; Siebert, F. Evidence for Light-Induced 13-Cis, 14-S-Cis, Isomerization in Bacteriorhodopsin Obtained by FTIR Difference Spectroscopy Using Isotopically Labelled Retinals. *EMBO J.* **1986**, *5*, 805–811.

(7) te Heesen, H.; Gerwert, K.; Schlitter, J. Role of the Arginine Finger in Ras-RasGAP Revealed by QM/MM Calculations. *FEBS Lett.* **2007**, *581* (29), 5677–5684.

(8) Klähn, M.; Schlitter, J.; Gerwert, K. Theoretical IR Spectroscopy Based on QM/MM Calculations Provides Changes in Charge Distribution, Bond Lengths, and Bond Angles of the GTP Ligand Induced by the Ras-Protein. *Biophys. J.* **2005**, *88* (6), 3829–3844.

(9) Nonella, M.; Mathias, G.; Tavan, P. Infrared Spectrum of P-Benzoquinone in Water Obtained from a QM/MM Hybrid Molecular Dynamics Simulation. J. Phys. Chem. A **2003**, 107 (41), 8638–8647.

(10) Rudack, T.; Xia, F.; Schlitter, J.; Kötting, C.; Gerwert, K. The Role of Magnesium for Geometry and Charge in GTP Hydrolysis, Revealed by Quantum mechanics/Molecular Mechanics Simulations. *Biophys. J.* **2012**, *103* (2), 293–302.

(11) Gerwert, K.; Freier, E.; Wolf, S. The Role of Protein-Bound Water Molecules in Microbial Rhodopsins. *Biochim. Biophys. Acta, Bioenerg.* **2014**, *1837* (5), 606–613.

(12) Kötting, C.; Gerwert, K. What Vibrations Tell Us about GTPases. *Biol. Chem.* **2015**, *396* (2), 131–144.

(13) Wilson, E. B.; Decius, J. C.; Cross, P. C. Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra; Dover Publications: New York NY, USA, 1980.

(14) Mann, D.; Teuber, C.; Tennigkeit, S. A.; Schröter, G.; Gerwert, K.; Kötting, C. Mechanism of the Intrinsic Arginine Finger in Heterotrimeric G Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (50), E8041–E8050.

(15) McQuarrie, D. A. *Statistical Mechanics;* University Science Books: Sausalito CA, USA, 2000.

(16) Babitzki, G.; Denschlag, R.; Tavan, P. Polarization Effects Stabilize Bacteriorhodopsin's Chromophore Binding Pocket: A Molecular Dynamics Study. *J. Phys. Chem. B* **2009**, *113* (30), 10483–10495.

(17) Steinbrecher, T.; Elstner, M. QM and QM/MM Simulations of Proteins. In *Biomolecular Simulations*; Monticelli, L., Salonen, E., Eds.; Humana Press: Totowa, NJ, 2013; Vol. 924, pp 91–124.

(18) Thomas, M.; Brehm, M.; Fligg, R.; Vöhringer, P.; Kirchner, B. Computing Vibrational Spectra from Ab Initio Molecular Dynamics. *Phys. Chem. Chem. Phys.* **2013**, *15* (18), 6608–22.

(19) Martinez, M.; Gaigeot, M.-P.; Borgis, D.; Vuilleumier, R. Extracting Effective Normal Modes from Equilibrium Dynamics at Finite Temperature. *J. Chem. Phys.* **2006**, *125* (14), 144106.

(20) Gaigeot, M. P.; Vuilleumier, R.; Sprik, M.; Borgis, D. Infrared Spectroscopy of N-Methylacetamide Revisited by Ab Initio Molecular Dynamics Simulations. J. Chem. Theory Comput. **2005**, 1, 772–789.

(21) Mathias, G.; Baer, M. D. Generalized Normal Coordinates for the Vibrational Analysis of Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2011**, 7 (7), 2028–2039.

(22) Sun, J.; Bousquet, D.; Forbert, H.; Marx, D. Glycine in Aqueous Solution: Solvation Shells, Interfacial Water, and Vibrational Spectroscopy from Ab Initio Molecular Dynamics. *J. Chem. Phys.* **2010**, 133, 114508.

(23) Wolf, S.; Freier, E.; Gerwert, K. A Delocalized Proton-Binding Site within a Membrane Protein. *Biophys. J.* **2014**, *107* (1), 174–184. (24) Wolf, S.; Freier, E.; Cui, Q.; Gerwert, K. Infrared Spectral Marker Bands Characterizing a Transient Water Wire inside a Hydrophobic Membrane Protein. *J. Chem. Phys.* **2014**, *141* (22), 22D524.

(25) Groenhof, G.; Bouxin-Cademartory, M.; Hess, B.; De Visser, S. P.; Berendsen, H. J. C.; Olivucci, M.; Mark, A. E.; Robb, M. A. Photoactivation of the Photoactive Yellow Protein: Why Photon Absorption Triggers a Trans-to-Cis Isomerization of the Chromophore in the Protein. J. Am. Chem. Soc. 2004, 126 (13), 4228–4233. (26) Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., J, A.; Montgomery, J., Vreven, T., Kudin, K. N., Burant, J. C., et al. Gaussian 03; Gaussian Inc.: Pittsburgh PA, 2003.

(27) Hariharan, P. C.; Pople, J. A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theor. Chim. Acta* **1973**, *28* (3), 213–222.

(28) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4* (3), 435–447.

(29) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OLPS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. J. Am. Chem. Soc. **1996**, 118 (15), 11225–11236.

(30) Nosé, S. A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys.* **1984**, *52* (2), 255–268.

(31) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, 31 (3), 1695–1697.

(32) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, 52 (12), 7182–7190.

(33) Nosé, S.; Klein, M. L. Constant Pressure Molecular Dynamics for Molecular Systems. *Mol. Phys.* **1983**, *50* (5), 1055–1076.

(34) Dierenfeldt, K. E. Isotope Ratio, Oscillator Strength, and Band Positions from CO2 IR Spectra A Physical Chemistry Experiment. *J. Chem. Educ.* **1995**, 72 (3), 281–283.

(35) Cooley, J. W.; Tukey, J. W. An Algorithm for the Machine Calculation of Complex Fourier Series. *Mathematics of Computation* **1965**, *19*, 297–301.

(36) Johnson, R. D. NIST Computational Chemistry Comparison and Benchmark Database NIST Standard Reference Database Number 101 Release 17b, September 2015.

(37) Neese, F. The ORCA Program System. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2012, 2 (1), 73–78.

(38) Richter, W. E.; Duarte, L. J.; Silva, A. F.; Bruns, R. E. Review of Experimental GAPT and Infrared Atomic Charges in Molecules. *J. Braz. Chem. Soc.* **2016**, 27 (6), 979–991.

(39) Coudert, F.-X.; Vuilleumier, R.; Boutin, A. Dipole Moment, Hydrogen Bonding and IR Spectrum of Confined Water. *Chem-PhysChem* **2006**, 7 (12), 2464–2467.

(40) Zhang, C.; Khaliullin, R. Z.; Bovi, D.; Guidoni, L.; Kühne, T. D. Vibrational Signature of Water Molecules in Asymmetric Hydrogen Bonding Environments. *J. Phys. Chem. Lett.* **2013**, *4* (19), 3245–3250.

I

(41) Goyal, P.; Qian, H.; Irle, S.; Lu, X.; Roston, D.; Mori, T.; Elstner, M.; Cui, Q. Molecular Simulation of Water and Hydration Effects in Different Environments: Challenges and Developments for DFTB Based Models. J. Phys. Chem. B 2014, 118 (38), 11007–11027.

(42) Jorgensen, W. L. Quantum and Statistical Mechanical Studies of Liquids. 10. Transferable Intermolecular Potential Functions for Water, Alcohols, and Ethers. Application to Liquid Water. J. Am. Chem. Soc. **1981**, 103 (2), 335–340.

(43) Wu, X.; Thiel, W.; Pezeshki, S.; Lin, H. Specific Reaction Path Hamiltonian for Proton Transfer in Water: Reparameterized Semiempirical Models. *J. Chem. Theory Comput.* **2013**, *9* (6), 2672–2686. (44) Mirkin, N. G.; Krimm, S. Conformers of Ck-IV-Methylaceta-

mide. J. Mol. Struct.: THEOCHEM **1991**, 236, 97–111.

(45) Chen, X. G.; Schweitzer-Stenner, R.; Krimm, S.; Mirkin, N. G.; Asher, S. A. N-Methylacetamide and Its Hydrogen-Bonded Water Molecules Are Vibrationally Coupled. J. Am. Chem. Soc. **1994**, 116 (24), 11141–11142.

(46) Kubelka, J.; Keiderling, T. A. Ab Initio Calculation of Amide Carbonyl Stretch Vibrational Frequencies in Solution with Modified Basis Sets. 1. N-Methyl Acetamide. *J. Phys. Chem. A* **2001**, *105* (48), 10922–10928.

(47) Schwörer, M.; Wichmann, C.; Tavan, P. A Polarizable QM/MM Approach to the Molecular Dynamics of Amide Groups Solvated in Water. J. Chem. Phys. **2016**, 144, 114504.

(48) Chen, X. G.; Schweitzer-Stenner, R.; Asher, S. A.; Mirkin, N. G.; Krimm, S. Vibrational Assignments of Trans-N-Methylacetamide and Some of Its Deuterated Isotopomers from Band Decomposition of IR, Visible, and Resonance Raman Spectra. *J. Phys. Chem.* **1995**, *99* (10), 3074–3083.

(49) DeCamp, M. F.; DeFlores, L.; McCracken, J. M.; Tokmakoff, A.; Kwac, K.; Cho, M. Amide I Vibrational Dynamics of N-Methylacetamide in Polar Solvents: The Role of Electrostatic Interactions. *J. Phys. Chem. B* **2005**, *109* (21), 11016–11026.

Supplementary Information

for

Local Mode Analysis: Decoding IR-Spectra by Visualizing Molecular Details

M. Massarczyk^{#1}, T. Rudack^{#1,2}, J. Schlitter^{1*}, J. Kuhne¹, C. Kötting¹ and K. Gerwert^{1,3*}

¹ Department of Biophysics, Ruhr-University, Bochum, Germany

² Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana Champaign, U.S.A.

³ Chinese Academy of Sciences–Max-Planck Partner Institute for Computational Biology (PICB), Shanghai Institutes for Biological Sciences (SIBS), Shanghai, China

* corresponding authors

these authors contributed equally

Supporting Movies

Supporting Movie SM1: Visualization of CO₂ local vibrational modes extracted from the simulation trajectory of CO₂ in vacuum. The shown structural motion that evokes the theoretical IR spectrum within the regions 2300 to 2400 cm⁻¹ (blue), 1300 to 1500 cm⁻¹ (gray), and 600 to 650 cm⁻¹ (orange) were extracted from the QM trajectory through the band–pass filter. The filtered QM trajectory for the regions 2300 to 2400 cm⁻¹ reveals the asymmetrical CO₂ stretching vibrational mode $v_{as}(CO_2)$ (blue box) and the one for the regions 600 to 650 cm⁻¹ reveals the asymmetrical CO₂ bending vibrational mode $\delta(CO_2)$ (orange box). Both vibrational modes are IR active. The filtered QM trajectory for the regions 1300 to 1500 cm⁻¹ reveals the symmetrical CO₂ stretching vibrational mode $v_s(CO_2)$ (gray box), which is IR inactive and therefore does not evoke a signal in the IR absorption spectrum. For a better visualization the vibrational motions of the vibrational modes are amplified.

Watch the animation here: <u>Supportin_Movie_SM1</u>

Supporting Movie SM2: Visualization of H₂O local vibrational modes extracted from the simulation trajectory of bulk water. The shown structural motion that evokes the theoretical IR spectrum within the regions 3100 to 3600 cm⁻¹ (blue) and 1600 to 1800 cm⁻¹ (orange) were extracted from the QM trajectory through the band–pass filter. The structural motion of the unfiltered QM region of the trajectory is shown in the black box. The filtered QM trajectory for the regions 3100 to 3600 cm⁻¹ reveals the H₂O stretching vibrational mode v(H₂O) (blue box) and the one for the regions 1600 to 1800 cm⁻¹ reveals the H₂O bending vibrational mode δ (H₂O) (orange box). For a better visualization the vibrational motions of the vibrational modes are amplified.

Watch the animation here: <u>Supporting_Movie_SM2</u>

Supporting Movie SM3: Visualization of solvation effects. Shown in the black box is the theoretical IR spectrum of solvated *t*-NMA (red) with its decomposition in the partial spectra for *t*-NMA (magenta) and its 1st solvation sphere (blue). The structural motion of the unfiltered QM region of the trajectory evoking the IR spectrum of solvated *t*-NMA is shown in the red box. The shown structural motions of only the *t*-NMA (magenta box) and its 1st solvation sphere (blue box) that evoke the theoretical IR spectrum within the region 1650 to 1800 cm⁻¹ (yellow) were extracted from the QM trajectory through the band–pass filter. The filtered QM trajectory of the 1st solvation sphere reveals the H₂O bending vibrational mode δ (H₂O) whereas the one for t-NMA reflects almost no motion. The motion of these two filtered trajectories visualize that water motions solely evoke the IR spectrum for the region 1650 to 1800 cm⁻¹, which is in agreement with the result from the Local Mode Decomposition. For a better visualization the vibrational modes are amplified.

Watch the animation here: <u>Supporting_Movie_SM3</u>

Supporting Movie SM4: Visualization of the bidirectional assignment in structure and spectrum of the local rocking mode r(NC1H3) of solvated *t*-NMA for the IR-band at 1130 to 1180 cm⁻¹. The left panel illustrates the assignment of structural motion to spectral feature. Partial spectra from the motion of the NH group, highlighted in blue, and the CH3 group, highlighted in green, are calculated through Local Mode Decomposition. The right panel illustrates the assignment of spectral IR-bands to structural motion. The structural motion of only the *t*-NMA that evokes the IR-band from 1130 to 1180 cm⁻¹ were extracted from the QM trajectory through the band–pass. The visualization of this IR-band reveals the local rocking mode r(NC1H₃) as main motion. Shown is a representative sequence for the average motion of the filtered trajectory. For a better visualization the vibrational motions of the vibrational modes are amplified.

Watch the animation here: <u>Supporting_Movie_SM4</u>

Supporting Movie SM5: Visualization of the bidirectional assignment in structure and spectrum of the local bending vibrational mode $\delta(C2H_3)$ of solvated *t*-NMA for the IR-band at 1290 to 1350 cm⁻¹. The left panel illustrates the assignment of structural motion to spectral feature. The partial spectrum from the motion of the second methyl group C2H3 group, highlighted in dark green, is calculated through Local Mode Decomposition. The right panel illustrates the assignment of spectral IR-bands to structural motion. The structural motions of only the *t*-NMA that evoke the IR-band from 1290 to 1350 cm⁻¹ were extracted from the QM trajectory through the band–pass. The visualization of this IR-band reveals the local bending vibrational mode $\delta(C2H_3)$ as main motion. Shown is a representative sequence for the average motion of the filtered trajectory. For a better visualization the vibrational motions of the vibrational modes are amplified.

Watch the animation here: Supporting Movie SM5

Supporting Movie SM6: Visualization of the bidirectional assignment in structure and spectrum of the local bending vibrational mode $\delta(C2H_3)$ of solvated *t*-NMA for the IR-band at 1350 to 1400 cm⁻¹. The left panel illustrates the assignment of structural motion to spectral feature. The partial spectrum from the motion of the second methyl group C2H3 group, highlighted in dark green, is calculated through Local Mode Decomposition. The right panel illustrates the assignment of spectral IR-bands to structural motion. The structural motion of only the *t*-NMA that evokes the IR-band from 1350 to 1400 cm⁻¹ were extracted from the QM trajectory through the band–pass. The visualization of this IR-band reveals the local bending vibrational mode $\delta(C2H_3)$ as main motion. Shown is a representative sequence for the average motion of the filtered trajectory. For a better visualization the vibrational motions of the vibrational modes are amplified.

Watch the animation here: <u>Supporting_Movie_SM6</u>

Supporting Movie SM7: Visualization of the bidirectional assignment in structure and spectrum of the local bending vibrational mode δ (C2H₃) of solvated *t*-NMA for the IR-band at 1400 to 1450 cm⁻¹. The left panel illustrates the assignment of structural motion to spectral

feature. The partial spectrum from the motion of the first methyl group C1H3 group, highlighted in dark green, is calculated through Local Mode Decomposition. The right panel illustrates the assignment of spectral IR-bands to structural motion. The structural motion of only the *t*-NMA that evokes the IR-band from 1400 to 1450 cm⁻¹ were extracted from the QM trajectory through the band–pass. The visualization of this IR-band reveals the local bending vibrational mode δ (C1H₃) as main motion. Shown is a representative sequence for the average motion of the filtered trajectory. For a better visualization the vibrational motions of the vibrational modes are amplified.

Watch the animation here: <u>Supporting_Movie_SM7</u>

Supporting Movie SM8: Visualization of the bidirectional assignment in structure and spectrum of the overlay of the local stretching vibrational mode v(CO) and the local bending vibrational mode $\delta(NH)$ of solvated *t*-NMA for the IR-band at 1550 to 1650 cm⁻¹. The left panel illustrates the assignment of structural motion to spectral feature. Partial spectra from the motion of the CO group, highlighted in orange, and the NH group, highlighted in blue, are calculated through Local Mode Decomposition. The right panel illustrates the assignment of spectral IR-bands to structural motion. The structural motions of only the *t*-NMA that evoke the IR-band from 1550 to 1650 cm⁻¹ were extracted from the QM trajectory through the band–pass. The visualization of this IR-band reveals an overlay of the local stretching vibrational mode v(CO) and the local bending vibrational mode $\delta(NH)$ as main motion. Shown is a representative sequence for the average motion of the filtered trajectory. For a better visualization the vibrational modes are amplified.

Watch the animation here: <u>Supporting_Movie_SM8</u>

All supporting movies can be downloaded here:

http://www.ks.uiuc.edu/~trudack/lma-animations/

Calculation of the CO₂ spectrum

A single CO_2 molecule was simulated in vacuum for 30 ps using the Gromacs/ORCA QM/MM interface to drive the QM simulation of CO_2 employing B3LYP (6-31G*). CO_2 was equilibrated for 100 ps by a classical MM simulation with Gromacs-4.5.5 with the OPLS-AA force field.

A comparison between the theoretical IR spectra of CO₂ obtained through Local Mode Analysis and normal mode analysis with the experimental IR spectra is shown in Figure S1. In the regions between 600 cm⁻¹ and 800 cm⁻¹ and between 2200 cm⁻¹ and 2400 cm⁻¹ the CO₂ spectrum has two well studied IR-bands (KE Dierenfeldt, *J. Chem. Educ.* 1995, 72:281). The IR-band around 650 cm⁻¹ represents the CO₂ bending vibrational mode. LMA gives an IR spectrum that clearly reveals a double peak for the IR-band from 2200 cm⁻¹ to 2400 cm⁻¹, which represents the CO₂ asymmetric stretching vibrational mode. This double peak cannot be observed through normal mode analysis. Furthermore, CO₂ has an IR inactive vibrational mode for the region from 1300 cm⁻¹ and 1500 cm⁻¹, which is the symmetric stretching vibrational mode of the CO₂. This mode is observed in the trajectory after applying the band-pass filter for the region between 1300 cm⁻¹ and 1500 cm⁻¹, which can be seen in Supporting Movie SM1.



Figure S1: Comparison of IR spectra for CO₂. The experimental IR spectrum of CO₂ (black) is compared to the theoretical IR spectra observed through LMA (pink) and by normal mode analysis (green). The contribution of the asymmetrical stretching vibrational mode (blue) and the bending vibrational mode (orange) are obtained through local mode decomposition. The arrows indicating the direction of atom movement are obtained by analyzing the trajectory after applying the bandpass filter for each of the IR-bands. The symmetrical stretching mode of CO₂, shown in gray the gray box for the region from 1300 cm⁻¹ to 1500 cm⁻¹ is IR-inactive and therefore not seen as IR-band in the IR spectrum but still observed as motion through the band-bass filter. The double peak of the asymmetrical stretching vibrational mode which is observed in the experimental spectrum is also observed by LMA but not by normal mode analysis.

Spectral matrices. Spectral matrices are a way to show the contribution of each atom to the overall spectrum. By the processing described in this paper these spectral matrices are simplified by the processing of the elements \mathbf{s}_{kl} . It is important to note that the sum of all elements of each spectral matrix results in the same total spectrum of the given system. It is just the representation that is changed. The residual spectral matrices in Figure S2 and S3 show exemplarily on CO₂ how the decomposition is done. Basically all bands which are not IR active are removed. They are compensated by the correlation spectra with negative intensity which do not appear any more as well. The modification of the spectral matrices from S2 to S3 are done by equation (5), (6), (7) and (8). The raw spectral matrix (Figure S2 A) is the unmodified spectral matrix of the standard DAF spectrum. The diagonal elements S_k with k = l are the elements in correlation with them self and the nondiagonal elements S_k with $k \neq 1$ are the elements in correlation with each other. In the example, both diagonal elements of the two oxygen atoms have vibrational modes which are not IR active, since their negative correlations remove this mode in the sum. By equation (5) and (6) this cancellation is already taken into account which makes interpretation of the spectra easier. In the second step by equation (7) and (8) the nondiagonal elements are removed completely by decomposing all elements \tilde{S}_{kl} to get all \tilde{S}_{kk} .



Figure S2: Spectral matrices of CO₂ S_{kl} elements, (A) is the raw spectral matrix and (B) is the residual spectral matrix containing the IR inactive bands. (A) The raw residual spectral matrix shows all elements S_{kl} in an unmodified state, like they are shown in equation (4). (B) This residual spectral matrix shows all bands which are eliminated through projection onto the overall moment $\mathbf{m}(t)$ which is the sum of all S_{kl} . The bands at 1350 cm⁻¹ in both oxygen elements (marked by green circles) and their correlations are the ones which are filtered out by the LMD algorithm.



Figure S3: Modified Spectral matrices of CO₂ \tilde{S}_{kl} elements, (A) is the spectral matrix which is projected onto the overall moment $\mathbf{m}(t)$ and (B) is the completely decomposed spectral matrix. (A) The modified Matrix contains all elements \tilde{S}_{kl} that were generated by the projection procedure (equation (6)). (B) The decomposed Matrix shows the result after the LMD procedure according to equation (7). The band that appears in the Figure S2 in the labeled region around 1350 cm⁻¹ in both oxygen elements and their correlations are removed here by the LMD algorithm.



Figure S4: Impact on theoretical spectra of QM treatment of 1st **solvation sphere.** LMA enables to calculate only the spectrum of the protein while including the influence of the solvent at the quantum level. Shown is the comparison of the theoretical t-NMA spectrum obtained from the QM/MM simulation treating *t*-NMA and the 1st solvation sphere at the quantum level with the theoretical spectrum including only *t*-NMA in the QM-region and treating the complete solvents by classical mechanics. Including the influence of the solvents at the quantum level leads to a red shift of the spectrum in direction of the experimental spectrum.