THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

pubs.acs.org/JPCA

Determining In Situ Protein Conformation and Orientation from the Amide-I Sum-Frequency Generation Spectrum: Theory and Experiment

S. J. Roeters,*[,†] C. N. van Dijk,[†] A. Torres-Knoop,[†] E. H. G. Backus,[‡] R. K. Campen,[§] M. Bonn,*[,‡] and S. Woutersen*[,†]

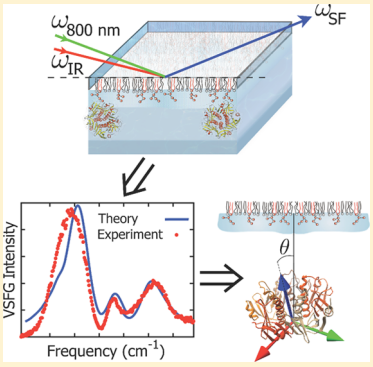
[†]Van't Hoff Institute for Molecular Sciences, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

[‡]Max-Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

[§]Department of Physical Chemistry, Fritz Haber Institute, Faradayweg 4-6, 14195 Berlin, Germany

Supporting Information

ABSTRACT: Vibrational sum-frequency generation (VSFG) spectra of the amide-I band of proteins can give detailed insight into biomolecular processes near membranes. However, interpreting these spectra in terms of the conformation and orientation of a protein can be difficult, especially in the case of complex proteins. Here we present a formalism to calculate the amide-I infrared (IR), Raman, and VSFG spectra based on the protein conformation and orientation distribution. Based on the protein conformation, we set up the amide-I exciton Hamiltonian for the backbone amide modes that generate the linear and nonlinear spectroscopic responses. In this Hamiltonian, we distinguish between nearest-neighbor and non-nearest-neighbor vibrational couplings. To determine nearest-neighbor couplings we use an ab initio 6-31G+(d) B3LYP-calculated map of the coupling as a function of the dihedral angles. The other couplings are estimated using the transition-dipole coupling model. The local-mode frequencies of hydrogen-bonded peptide bonds and of peptide bonds to proline residues are red-shifted. To obtain realistic hydrogen-bond shifts we perform a molecular dynamics simulation in which the protein is solvated by water. As a first application, we measure and calculate the amide-I IR, Raman, and VSFG spectra of cholera toxin B subunit docked to a model cell membrane. To deduce the orientation of the protein with respect to the membrane from the VSFG spectra, we compare the experimental and calculated spectral shapes of single-polarization results, rather than comparing the relative amplitudes of VSFG spectra recorded for different polarization conditions for infrared, visible, and sum-frequency light. We find that the intrinsic uncertainty in the interfacial refractive index—essential to determine the overall amplitude of the VSFG spectra—prohibits a meaningful comparison of the intensities of the different polarization combinations. In contrast, the spectral shape of most of the VSFG spectra is independent of the details of the interfacial refractive index and provides a reliable way of determining molecular interfacial orientation. Specifically, we find that the symmetry axis of the cholera toxin B subunit is oriented at an angle of 6° ± 17° relative to the surface normal of the lipid monolayer, in agreement with 5-fold binding between the toxin's five subunits and the receptor lipids in the membrane.



INTRODUCTION

The membrane is one of the most important organs of the cell. It plays a vital role in the exchange of ions, polynucleotides, and proteins and provides a barrier against pathogens. Over the length of only a few nanometers, it contains a complex and dynamic mixture of lipids and proteins that perform many crucial biological functions. Obtaining a better understanding of these processes is challenging, if only because one wants to measure exclusively a limited amount of interfacial molecules that are surrounded by an overwhelming majority of very similar bulk molecules. One solution to this problem is to label the molecules of interest, but this can perturb the molecular interactions and spatially averaged membrane properties, for example, in the case of isotope[1] or fluorescent[2] labeling. Vibrational spectroscopic techniques, in which molecules are probed by their intrinsic vibrations, provide a label-free alternative. For instance, the amide-I (predominantly C═O stretch) mode profile of a protein directly reflects its secondary structure.[3] Conventional vibrational spectroscopic techniques, for example, IR absorbance or Raman scattering, are not interface-sensitive and are thus in general not capable of distinguishing the spectral response of molecules present at an interface from the much larger number often present in bulk. Attempts to overcome this problem have been made using attenuated total reflection type sample geometries.[4] In this approach, the spectral response of moieties near the interface,

Special Issue: Prof. John C. Wright Festschrift

Received: January 31, 2013
Revised: April 8, 2013
Published: April 8, 2013

   dx.doi.org/10.1021/jp401159r | J. Phys. Chem. A 2013, 117, 6311−6322

within a layer of a thickness set by the penetration depth of the evanescent wave, is probed. This penetration depth is determined by the refractive indices of the relevant media and the wave vector of the incident infrared light. At amide-I frequencies in most practical sample geometries this penetration depth is approximately 250 nm: much larger than the few nanometer thickness of a single membrane.

The nonlinear optical technique of vibrational sum frequency generation (VSFG) spectroscopy is by its symmetry selection rules interface specific[5] and extremely sensitive.[6] By overlapping an infrared and a visible laser pulse in space and time at an interface, emission is generated at the sum frequency of the two incident fields. Like other vibrational spectroscopic techniques, VSFG allows label-free measurements with high temporal and spatial resolution suitable for most biomolecular processes. Because of its extraordinary sensitivity, short acquisition times are sufficient to obtain a high signal-to-noise ratio. Finally, in contrast to many conventional surface characterization methods, VSFG experiments can be performed under biologically relevant conditions, and in principle even in vivo, although such experiments in the amide-I region (analogous to the in vivo experiments by Koelsch et al. in the CH- and CD-region)[7,8] will be challenging to interpret due to the complex mixture of amide-group containing molecules in and near the cellular membrane. Previous authors have applied VSFG spectroscopy to investigate biomimetic interfaces in general[9−14] and to probe proteins at interfaces in particular.[15−20]

The amide-I normal modes of a protein are collective oscillations of numerous interacting amide-I oscillators in the backbone of the protein. As this interaction is dependent on the protein conformation, the amide-I region is particularly suited to study proteins.[3] The spectral lineshapes and the relative magnitudes of different polarization combinations of the VSFG amide-I spectrum are also determined by the protein orientation relative to the surface and to other modes present at the interface, for example, C═O-stretch modes in lipid head groups. It is evident that to deduce, from amide-I VSFG measurements of interfacial proteins, quantitative information on the orientation, surface coverage and structure, a theoretical framework to interpret the spectra is required. Interestingly, both for IR and Raman spectroscopy such theoretical frameworks are already available. The relation between the conformation and the linear (steady-state) amide-I IR spectrum has been explored extensively.[21−26] Gorbunov et al.,[27] Falvo et al.,[28] Ganim et al.,[29] and Hamm et al.[30] have created similar excitonic models for interpreting nonlinear amide-I 2D-IR protein spectra. Huang et al.[31] and Tsuboi et al.[32] have performed such calculations to predict protein Raman spectra. For VSFG, similar efforts have been made to develop theories of structure and orientation determination using polarization-dependent VSFG. Specifically, Zhuang et al. have presented a method to determine the orientation and conformation of relatively simple molecules in which only a small number of modes are present.[33] Nguyen et al. have provided analytical solutions for ideal secondary structures (α- and 3_{10}-helix and β-sheet), allowing orientation determination for proteins that have an isolated band mainly from one of these secondary structure motifs.[34,35] Simpson and co-workers have published a number of benchmark papers on calculating the sum frequency spectra of electronically and vibrationally resonant amide-I spectra, focusing mainly on the influence of the frequency of the Raman excitation.[36,37] Finally, Xiao et al. have performed VSFG spectrum calculations based on binning the β-sheets of human islet amyloid polypeptide aggregates into bins of six neighboring amino acids, on which they performed an ab initio normal-mode analysis.[38] From the ratio of two peaks that are dominated by β-sheet normal mode contributions the orientation of the aggregate in lipid monolayers at the air−lipid−water interface was deduced. We note, however, that a formalism to calculate spectra for proteins of arbitrary size and complexity has not yet been developed. Here we present such formalism, thus making it possible to study the conformation and orientation of proteins in and near membranes. By including nearest-neighbor coupling, non-nearest neighbor coupling, and hydrogen-bond effects we calculate the VSFG hyperpolarizability tensor. One can fit the elements of this tensor to their corresponding experimental spectra to identify proteins and obtain protein orientations. We anticipate that this will make VSFG an even more powerful tool for gaining insight into biomolecular processes.

As a first demonstration, we apply the formalism to determine the orientation of cholera toxin B subunit (CTB) docked to a biomimetic membrane. We measured the VSFG spectrum of this protein for four polarization combinations. Comparison of our calculations with the experimental results shows good agreement of the spectra and moreover enables us to determine the orientation of the CTB protein with respect to the membrane surface.

FORMALISM

We use the theoretical framework of refs 22, 31−33, and 39 to calculate the VSFG amide-I spectrum of proteins, based on their conformation and orientation. From the protein conformation, the amide-I exciton Hamiltonian of the system is determined. We obtain the delocalized vibrational eigenmodes by solving the time-independent Schrödinger equation. From these eigenmodes, the IR and Raman responses of the molecule are calculated. Their tensor product gives the molecular hyperpolarizability: the SFG response of the molecule. By averaging this molecular response over the protein orientation distribution one can determine the response in the lab frame. To determine the effective hyperpolarizability for different polarization combinations one has to account for the differences in refractive index via multiplication with the corresponding Fresnel factors.[33] Finally, in some cases, a nonresonant SFG signal needs to be taken into account. In the next sections we will discuss each of the above steps in detail and show how the VSFG spectrum is determined by the protein orientation and conformation. This dependency makes it possible to gain insight into biomolecular processes by comparing calculated with experimental VSFG spectra.

Excitonic Amide-I Hamiltonian. The calculation of the spectra is based on the protein conformation. We start the numerical implementation from the conformation in PDB-format, so that spectra can be calculated for any structure obtained by X-ray crystallography, nuclear magnetic resonance (NMR), or molecular dynamics (MD) simulations that has been uploaded to a protein structure database, as well as for theoretically predicted structures that might be expected for a particular experimental situation. The excitonic amide-I Hamiltonian is constructed by analyzing the local modes and the couplings between them, calculated from the atomic coordinates in the PDB file. The Hamiltonian has the following form in the local-mode basis:

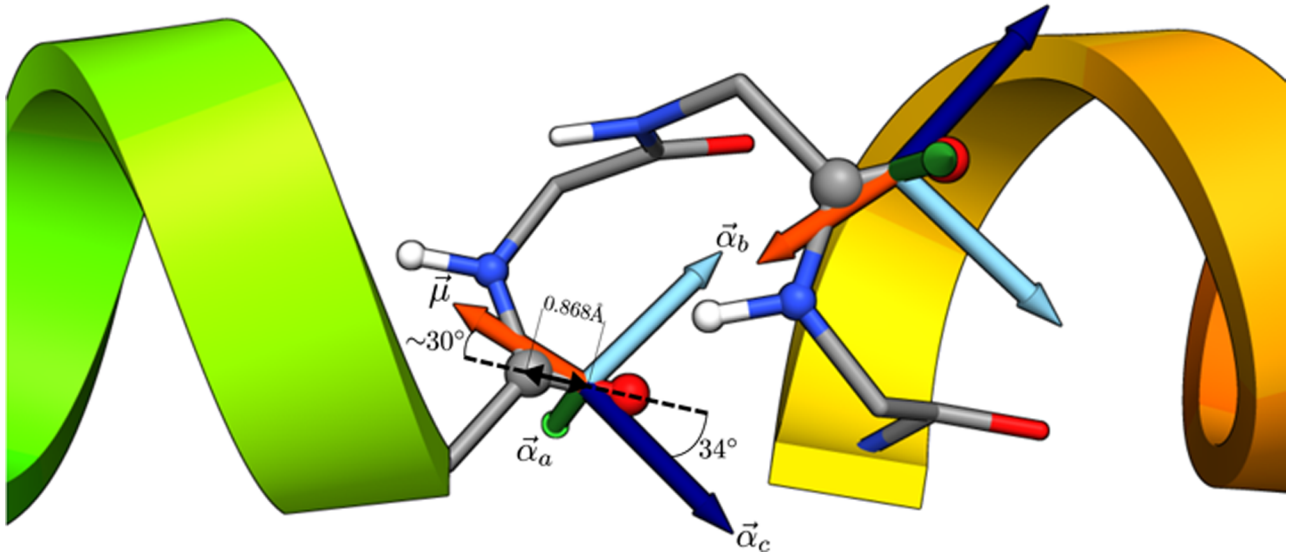


Figure 1. IR transition-dipole moment $\vec{\mu}$ and Raman-tensor principal axes of an amide group. The IR transition-dipole moment (orange) is positioned in the amide plane at a distance of 0.868 Å from the C=O dipole.[55,56] The value indicated for the angle between the transition-dipole moment $\vec{\mu}$ and the C=O bond axis is the average angle calculated using the transition charges,[30] for a typical protein. The amide plane is made up by the C, O, N, and H atoms of the backbone (depicted as balls here). The Raman tensor *c*- and *b*-axes (dark and light blue, respectively) are also positioned in the amide plane, at angles of −34° and 56°, respectively, with respect to the C═O bond. The *a*-axis (green) is perpendicular to the *c*,*b*-plane.[39,49]

$$H = \begin{pmatrix} \hbar\omega_1^0 & \kappa_{12} & \kappa_{13} & \kappa_{14} & \cdots \\ \kappa_{12} & \hbar\omega_2^0 & \beta_{23} & \kappa_{24} & \\ \kappa_{13} & \kappa_{23} & \hbar\omega_3^0 & \kappa_{34} & \\ \kappa_{14} & \kappa_{24} & \kappa_{34} & \hbar\omega_4^0 & \\ \vdots & & & & \ddots \end{pmatrix} \quad (1)$$

with ω_i^0 the frequency of local-mode i and κ_{ij} the coupling between local-modes i and j. The amide-I vibrations of proteins are mainly determined by the secondary structure and are not significantly affected by the nature of the side chains.[3,40] This approximation generally works well, but more sophisticated local-mode-frequency approximations may be applied.[41] We assume that the amide-I modes are spectrally isolated from other modes, so that the backbone modes only couple with each other and not with other modes.

Nondiagonal Elements. While the coupling between non-nearest neighbor local amide-I modes can be well-estimated by through-space transition-dipole interactions,[23,42] this model breaks down for nearest neighbor couplings, in which through-bond interactions play a significant role.[43] We therefore apply two different models to calculate the two categories of nondiagonal elements of the Hamiltonian. To determine the nearest-neighbor couplings (i.e., $\kappa_{i,i\pm1}$ and $\kappa_{i\pm1,i}$) we use values obtained from ab initio density functional theory (DFT). All other couplings are calculated using the transition-dipole coupling model (TDCM).

The nearest-neighbor DFT approach is implemented using a state-of-the-art map of the coupling as a function of the dihedral angles (ϕ,ψ), based on comprehensive quantum-chemical calculations on "glycine dipeptide" (Ac-Gly-NHCH$_3$), using the 6-31G+(d) basis set and B3LYP-functional.[27] We use a parametrized map of these calculations that can be obtained from the Web site of ref 44.

The model for amide-I moieties that are further apart approximates the interaction with a Coulomb-like interaction between their transition-dipole moments:[45]

$$\kappa_{ij} = \frac{1}{4\pi\varepsilon_0}\left(\frac{\vec{\mu}_i\cdot\vec{\mu}_j}{|\vec{r}_{ij}|^3} - 3\frac{(\vec{r}_{ij}\cdot\vec{\mu}_i)(\vec{r}_{ij}\cdot\vec{\mu}_j)}{|\vec{r}_{ij}|^5}\right) \quad (2)$$

with ε_0 the dielectric constant, $\vec{\mu}_i$ the transition-dipole of peptide bond i, and $\vec{r}_{ij}$ the vector connecting dipole i and j. We obtain the coordinates of the backbone C, O, N, and H atoms from the PDB file, and determine the transition-dipoles from these coordinates using ab initio precalculated values for the displacements, charges, and charge flows,[30,46] that can also be obtained from the Web site of ref 44.

Diagonal Elements. The local amide-I frequencies ω_i^0 are not the same for all backbone amide groups. We apply two types of local-mode frequency shifts.

Amide-I Red-Shift in Proline Residues. For the peptide bond upstream in the amino acid chain of proline residues, the amide-I frequency is red-shifted because the nitrogen atom of the peptide bond is bound to the C$_\delta$-atom of the proline ring, instead of to a hydrogen atom as in the case of other residues. We implement a $\delta\omega_{\mathrm{Pro}}$ = 19 cm^{-1} red-shift, the calculated difference in the amide-I′ frequency of *N*-deuterio-*N*-methylacetamide (NMA-D) and acetyl-*N*-pyrrolidine, on the B3PLYP+D3/aug-cc-pVDZ level (see the Supporting Information).

Amide-I Red-Shift due to Hydrogen Bonding. In general, an empirical local-mode frequency shift $\delta\omega_{\mathrm{HB},i}$ is used to account for hydrogen bonding, both to the C═O bond of the peptide groups and to the proton of the peptide groups.[25,45,47,48] Here, we use an adapted form of the empirical formula used in ref 45 to estimate this shift (see the Results and Discussion section). To obtain realistic hydrogen-bond shifts MD simulations can be performed to make an estimation of the intra- and extra-molecular hydrogen-bonding network.

dx.doi.org/10.1021/jp401159r | J. Phys. Chem. A 2013, 117, 6311−6322

We thus estimate the local-mode frequency of backbone amides upstream of proline residue i as follows:

$$\omega_i^0 = \Omega^0 - \delta\omega_{\mathrm{Pro}} - \delta\omega_{\mathrm{HB},i} \tag{3}$$

with Ω^0 the isolated local-mode frequency, which is red-shifted by $\delta\omega_{\mathrm{Pro}}$ and $\delta\omega_{\mathrm{HB},i}$ to obtain the corrected local-mode frequency ω_i^0. For the other hydrogen-bonded amides we use:

$$\omega_i^0 = \Omega^0 - \delta\omega_{\mathrm{HB},i} \tag{4}$$

VSFG Spectrum. *Vibrational Eigenmode Construction.* The delocalized vibrational eigenmodes are determined by diagonalizing the one-exciton Hamiltonian, thus obtaining the eigenvalues ω^ν and eigenvectors $c^{\sigma\nu}$ of the eigenmodes $|\nu\rangle$ of the protein. These can be expanded as:

$$|\nu\rangle = \sum_\sigma c^{\sigma\nu}|\sigma\rangle \tag{5}$$

Here $|\sigma\rangle$ is the localized $\nu = 1$ state of the amide-I mode of peptide unit σ and $c^{\sigma\nu}$ the eigenvector component that gives the contribution of each local mode σ to eigenmode $|\nu\rangle$.

The local-mode transition dipoles and Raman tensors are determined from the backbone atom coordinates. Specifically, the IR transition-dipole moments are determined using the transition charges[30] (see Figure 1). We find that for a typical protein this results in an average angle of 149.6° ± 1.5° with respect to the C═O bond. To obtain the orientation of the principal Raman axes we use refs 39 and 49. The c-axis (defined to be in the direction in which the polarizability is largest) makes an angle of 34° with the C═O bond, in the amide plane, while the a-axis is perpendicular to the amide plane, and the b-axis is perpendicular to both the a- and the c-axis. In the principal axis (a, b, c)-frame, the polarizability of local amide-I mode σ is:[32]

$$\alpha_{ij}^\sigma = \begin{pmatrix} 0.05 & & \\ & 0.20 & \\ & & 1.00 \end{pmatrix} \tag{6}$$

for i, j = a, b, c. We apply the following transformations to obtain the Raman tensors of the local modes in the molecular (x, y, z)-frame of the protein from the Raman tensors in the local-mode (a, b, c)-frame:[32]

$$\begin{aligned}
\alpha_{xx}^\sigma &= l_x^2\alpha_{aa}^\sigma + l_y^2\alpha_{bb}^\sigma + l_z^2\alpha_{cc}^\sigma \\
\alpha_{yy}^\sigma &= m_x^2\alpha_{aa}^\sigma + m_y^2\alpha_{bb}^\sigma + m_z^2\alpha_{cc}^\sigma \\
\alpha_{zz}^\sigma &= n_x^2\alpha_{aa}^\sigma + n_y^2\alpha_{bb}^\sigma + n_z^2\alpha_{cc}^\sigma \\
\alpha_{yz}^\sigma &= m_x n_x\alpha_{aa}^\sigma + m_y n_y\alpha_{bb}^\sigma + m_z n_z\alpha_{cc}^\sigma \\
\alpha_{xy}^\sigma &= l_x m_x\alpha_{aa}^\sigma + l_y m_y\alpha_{bb}^\sigma + l_z m_z\alpha_{cc}^\sigma \\
\alpha_{xz}^\sigma &= l_x n_x\alpha_{aa}^\sigma + l_y n_y\alpha_{bb}^\sigma + l_z n_z\alpha_{cc}^\sigma
\end{aligned} \tag{7}$$

with l_x, l_y, ..., m_y, m_z the direction cosines.[50]

To obtain the IR and the Raman response of eigenmode ν we apply eq 5:

$$\mu_k^\nu = \langle 0|\hat{\mu}_k|\nu\rangle = \sum_\sigma c^{\sigma\nu}\mu_k^\sigma \tag{8a}$$

$$\alpha_{ij}^\nu = \langle 0|\hat{\alpha}_{ij}|\nu\rangle = \sum_\sigma c^{\sigma\nu}\alpha_{ij}^\sigma \tag{8b}$$

for i, j, k = x, y, z, with $\hat{\mu}_k$ and $\hat{\alpha}_{ij}$ the electric dipole operator and the Raman scattering operator in the molecular frame, respectively, and μ_k^σ and α_{ij}^σ the transition-dipole moment and Raman tensor of the local amide-I mode of peptide unit σ. Assuming homogeneous line broadening for the individual normal modes, the IR and Raman intensities can now be calculated as follows:

$$I_{\mathrm{IR}} \propto \sum_\nu \left| \frac{\vec{\mu}^\nu}{\omega^\nu - \omega_{\mathrm{IR}} - i\Gamma} \right|^2 \tag{9}$$

and:

$$I_{\mathrm{Raman}} \propto \sum_\nu \left| \frac{\alpha^\nu}{\omega^\nu - (\omega_{\mathrm{laser}} - \omega_{\mathrm{Stokes}}) - i\Gamma} \right|^2 \tag{10}$$

with ω^ν the eigenvalues of the Hamiltonian (that give the normal-mode frequencies), ω_{IR} the frequency of the IR field, ω_{laser} the frequency of the excitation laser, ω_{Stokes} the frequency of the Stokes field and Γ the line width of the Lorentzian, assumed to be equal for all normal modes. The expressions for α^ν for the (unpolarized) Raman scattering intensity in terms of the tensor components α_{ij} can be found in ref 51. To test the calculation of the Raman spectra from the Hamiltonian, we have applied our formalism to the Hamiltonian of the tetraalanine-peptide presented in ref 51, to obtain exactly identical values for the isotropic Raman scattering ratios.

The molecular VSFG hyperpolarizability $\beta_{ijk}^{(2)}$ of mode ν in the molecular (x,y,z)-frame is given by the tensor product of the IR and Raman response of the mode:

$$\beta_{ijk}^{(2)\nu} = \mu_k^\nu \otimes \alpha_{ij}^\nu \tag{11}$$

for i, j, k = x, y, z. As a check, we apply the formalism to a 170 amino acids long perfect polyalanine α-helix. We obtain similar polarization combination ratios for $\beta_{ijk}^{(2)\nu}$ for the E_1 and A-modes as reported previously:[34] the authors of ref 34 calculate β_{xxz}/β_{zzz} = 0.31 and β_{xzx}/β_{zzz} = 0.59, where we find 0.33 and 0.58, respectively. Our approach differs from ref 34 and ref 35 in that we do not use analytical calculations of the normal modes of perfect secondary structure motifs but instead construct the normal modes from the interacting local modes of an arbitrary protein conformation. The calculated A–E-frequency splitting also matches the splitting observed in ref 53.

The vibrational eigenmodes of the protein sum up to give the frequency-dependent hyperpolarizability of the protein:[54]

$$\tilde{\beta}_{ijk}^{(2)\mathrm{protein}} = \frac{-1}{2\hbar}\sum_\nu \frac{\beta_{ijk}^{(2)\nu}}{(\omega^\nu - \omega_{\mathrm{IR}} - i\Gamma)} \tag{12}$$

for i, j, k = x, y, z, with ω_{IR} the IR frequency (obeying conservation of energy, so $\omega_{\mathrm{IR}} = \omega_{\mathrm{SF}} - \omega_{\mathrm{VIS}}$), ω^ν the resonance frequency of normal mode ν that is given by the corresponding eigenvalue (obtained by diagonalizing the Hamiltonian) and Γ the line width of the Lorentzians, assumed to be equal for all normal modes.

Effective Nonlinear Susceptibility $\chi_{IJK}^{(2)}$. Now that the frequency-dependent protein hyperpolarizability $\tilde{\beta}_{ijk}^{(2)}$ is constructed in the molecular (x, y, z)-frame, we can determine the observed VSFG response in the lab (X, Y, Z)-frame. The orientation of a single protein can be characterized by the three Euler angles (θ, ϕ, ψ) that relate the molecular (x, y, z)- to the lab (X, Y, Z)-coordinate frame (see Figure 2). The measured

dx.doi.org/10.1021/jp401159r | J. Phys. Chem. A 2013, 117, 6311−6322

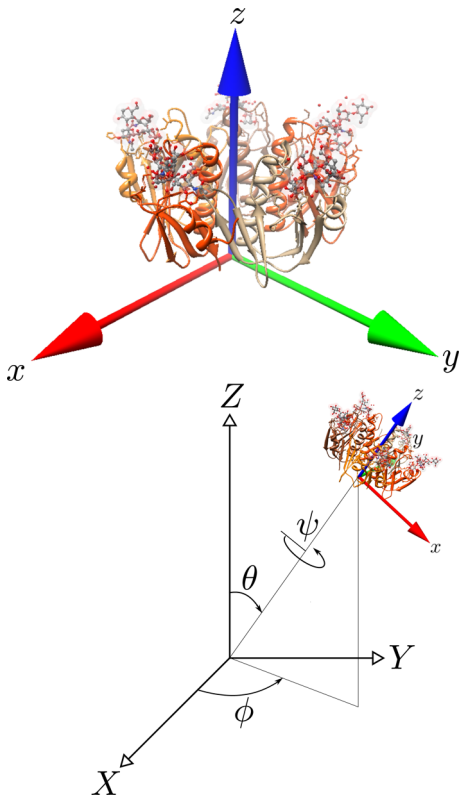


Figure 2. Definition of the molecular axes of CTB and the Euler angles (θ, ϕ, ψ) that transform the atom coordinates from the molecular to the lab frame. The z-axis of the protein is defined as the symmetry axis of the homopentamer.

signal is a collective response of the molecules in the region where the IR and the visible beams overlap and hence is an average over the distribution of protein orientations. We therefore have to average the Euler transformation over the molecular orientation distribution in order to obtain the nonlinear susceptibility in the lab frame:[33]

$$\chi_{IJK}^{(2)\text{protein}} = N \sum_{x,y,z} \langle (\hat{X}\cdot\hat{x})(\hat{Y}\cdot\hat{y})(\hat{Z}\cdot\hat{z})\rangle \tilde{\beta}_{ijk}^{(2)\text{protein}} \tag{13}$$

for $I, J, K = X, Y, Z$, for $i, j, k = x, y, z$, and where $\hat{X}$, $\hat{Y}$, $\hat{Z}$ and $\hat{x}$, $\hat{y}$, $\hat{z}$ are the unit vectors in the lab frame and molecular frame coordinates, respectively, N is the number of molecules that contribute to the response, and $\langle ... \rangle$ indicates the averaging over the molecular orientation distribution.

From the nonlinear susceptibility we obtain the measured VSFG intensity as follows:

$$I_{\text{VSFG}} \propto I_{\text{IR}} I_{\text{VIS}} \left| \chi_{IJK}^{(2)\text{protein}} \right|^2 \tag{14}$$

for $I, J, K = X, Y, Z$ or S, P (see Figure 3), and with I_{IR} the IR intensity, and I_{VIS} the visible intensity.

Sometimes a nonresonant (or, more precisely, off-resonant) contribution influences the shape of the VSFG spectrum to a certain degree. This can for instance be the OH- or the OD-stretch vibration (depending on the sample), for which the resonance frequency lies outside the spectral window of interest. However, the trailing or leading edge may be sufficiently intense to affect the VSFG response. In that case, a nonresonant term must be added to obtain the measured VSFG intensity:

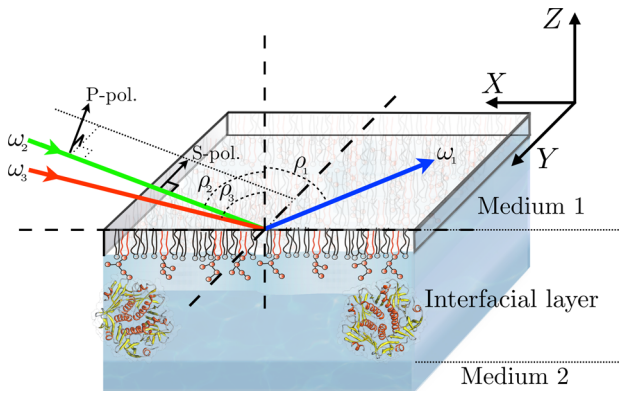


Figure 3. Schematic of our experimental setup, depicting the macroscopic- or lab-coordinate system of VSFG, medium 1 with refractive index n_1, an interfacial layer that may consist of multiple molecular layers with possibly each a different interfacial refractive index n', and the subphase (medium 2) with refractive index n_2. The X,Y-plane is the surface, and Z is the surface normal. The S- or P-polarization of the beams refer to whether they are polarized perpendicular (*senkrecht*) or parallel, respectively, to the plane of incidence.

$$I_{\text{VSFG}} \propto I_{\text{IR}} I_{\text{VIS}} \left| A_{\text{NR}} e^{i\phi_{\text{NR}}} + \chi_{IJK}^{(2)\text{protein}} \right|^2 \tag{15}$$

with A_{NR} and ϕ_{NR} the amplitude and phase of the nonresonant contribution, respectively, together determining the nonresonant nonlinear susceptibility $\chi_{\text{NR}}^{(2)}$.

The refractive indexes at either side of, and in the interfacial layer will in general all be different. These different refractive indexes cause the local fields at the interface to be different from the fields incident from the far field. To account for these local field effects originating from differences in refractive index, which influence the observed response differently for different polarization combinations, we multiply the nonlinear susceptibility with the Fresnel factors. For the four most commonly observed elements of the macroscopic SFG second-order susceptibility tensor (that probe the seven independent elements of the molecular SFG hyperpolarizability tensor of nonchiral molecules), this is done as follows:[33]

$$\chi_{\text{eff,SSP}}^{(2)} = L_{YY}(\omega_1) L_{YY}(\omega_2) L_{ZZ}(\omega_3) \times \sin\rho_3 \times \chi_{YYZ}^{(2)} \tag{16a}$$

$$\chi_{\text{eff,SPS}}^{(2)} = L_{YY}(\omega_1) L_{ZZ}(\omega_2) L_{YY}(\omega_3) \times \sin\rho_2 \times \chi_{YZY}^{(2)} \tag{16b}$$

$$\chi_{\text{eff,PSS}}^{(2)} = L_{ZZ}(\omega_1) L_{YY}(\omega_2) L_{YY}(\omega_3) \times \sin\rho_1 \times \chi_{ZYY}^{(2)} \tag{16c}$$

$$\begin{aligned} \chi_{\text{eff,PPP}}^{(2)} = {} & -L_{XX}(\omega_1) L_{XX}(\omega_2) L_{ZZ}(\omega_3) \times \cos\rho_1 \times \cos\rho_2 \\ & \times \sin\rho_3 \times \chi_{XXZ}^{(2)} - L_{XX}(\omega_1) L_{ZZ}(\omega_2) L_{XX}(\omega_3) \\ & \times \cos\rho_1 \times \sin\rho_2 \times \cos\rho_3 \times \chi_{XZX}^{(2)} \\ & + L_{ZZ}(\omega_1) L_{XX}(\omega_2) L_{XX}(\omega_3) \times \sin\rho_1 \times \cos\rho_2 \\ & \times \cos\rho_3 \times \chi_{ZXX}^{(2)} + L_{ZZ}(\omega_1) L_{ZZ}(\omega_2) L_{ZZ}(\omega_3) \\ & \times \sin\rho_1 \times \sin\rho_2 \times \sin\rho_3 \times \chi_{ZZZ}^{(2)} \end{aligned} \tag{16d}$$

where S and P refer to the polarization of the incoming and outgoing beams (see Figure 3). For nonchiral molecules, these are the only nonzero elements. For chiral molecules, also $\chi_{\text{eff,PSP}}^{(2)}$, $\chi_{\text{eff,SPP}}^{(2)}$, and $\chi_{\text{eff,PPS}}^{(2)}$ are nonzero.[57] The Fresnel factors L_{ii} for beam j are:

dx.doi.org/10.1021/jp401159r | J. Phys. Chem. A 2013, 117, 6311−6322

$$L_{xx}(\omega_j) = \frac{2n_1(\omega_j)\cos\gamma_j}{n_1(\omega_j)\cos\gamma_j + n_2(\omega_j)\cos\rho_j} \quad (17a)$$

$$L_{yy}(\omega_j) = \frac{2n_1(\omega_j)\cos\rho_j}{n_1(\omega_j)\cos\rho_j + n_2(\omega_j)\cos\gamma_j} \quad (17b)$$

$$L_{zz}(\omega_j) = \frac{2n_2(\omega_j)\cos\rho_j}{n_1(\omega_j)\cos\gamma_j + n_2(\omega_j)\cos\rho_j} \times \left(\frac{n_1(\omega_j)}{n'(\omega_j)}\right)^2 \quad (17c)$$

with ω_1, ω_2, and ω_3 the sum, visible, and IR frequencies, respectively, ρ_2 and ρ_3 the angle of incidence of the visible and IR beam, respectively, ρ_1 the angle at which the sum frequency beam is generated, given by $\sin\rho_1 = (k_2 \sin\rho_2 + k_3 \sin\rho_3)/(k_2 + k_3)$ (with k_2 and k_3 the wave vectors of the visible and IR field, respectively), n_1, n_2, and n' the indexes of refraction of the dispersive media as defined in Figure 3, and γ_j the refracted angle, given by $\sin\gamma_j = (n_1(\omega_j)/n_2(\omega_j))\sin\rho_j$.

Now that we have calculated the effective nonlinear hyperpolarizability as a function of $\beta_{ijk}^{(2)\nu}$ (the normal modes of the protein), one can relate observed VSFG signals to protein conformation and orientation. Equations 1 to 11 show the dependency of the VSFG response on the structure and eq 13 shows the dependency on the protein orientation.

RESULTS AND DISCUSSION

As a first application of the above formalism, we apply it to the membrane docking of the cholera toxin B subunit (CTB). This protein binds the pentasaccharide headgroup of specific ganglioside lipids in a 5-fold fashion: one GM_1-lipid per subunit.[58,59] This results in a parallel orientation of the z-axis of the protein with the normal of the membrane[58,60] (see Figure 2). Reproducing the spectra and the protein orientation can validate the formalism for protein identification and orientation determination purposes.

We measure (see Figure 4 and the Experimental Details section) and calculate (see Figure 5) the IR, Raman, and VSFG spectra.

Calculating the IR, Raman, and VSFG Spectra of CTB. To calculate the IR, Raman, and VSFG spectra of CTB, we apply the formalism to PDB-entry 2CHB, the cholera toxin B subunit complexed with GM_1 pentasaccharide. We first optimize and solvate the structure in a 10 ps MD simulation to obtain realistic hydrogen-bond lengths, including bonds from solvating water molecules (see Experimental Details section for details on the MD simulation). We construct the Hamiltonian, in which we estimate the hydrogen-bond induced local-mode amide-I frequency shifts on the diagonal of the Hamiltonian (see eqs 3 and 4) using an adapted form of the empirical algorithm used in ref 45:

$$\delta\omega_{\mathrm{HB},i} = \sum_{j=1}^{n_{\mathrm{HB}}} \Delta\omega_{\mathrm{HB},j}(3.5\,\text{Å} - r_{\mathrm{O}\cdots\mathrm{H},j}) \quad (18)$$

with i the local-mode index, n_{HB} the number of hydrogen bonds that amide group i participates in, $\Delta\omega_{\mathrm{HB},j}$ the empirical proportionality constant of hydrogen bond j, and $r_{\mathrm{O}\cdots\mathrm{H},j}$ the distance between the accepting oxygen atom and the donating hydrogen atom.[45]

There are five possible values that $\Delta\omega_{\mathrm{HB},j}$ can take, for the five different types of hydrogen bonds that a peptide bond can participate in. We chose the values of $\Delta\omega_{\mathrm{HB},j}$ such that eq 18 is a linear approximation of the distance dependency of the red-shift obtained with ab initio calculations on the RHF/6-311++G** level by previous authors.[47] For hydrogen-bond

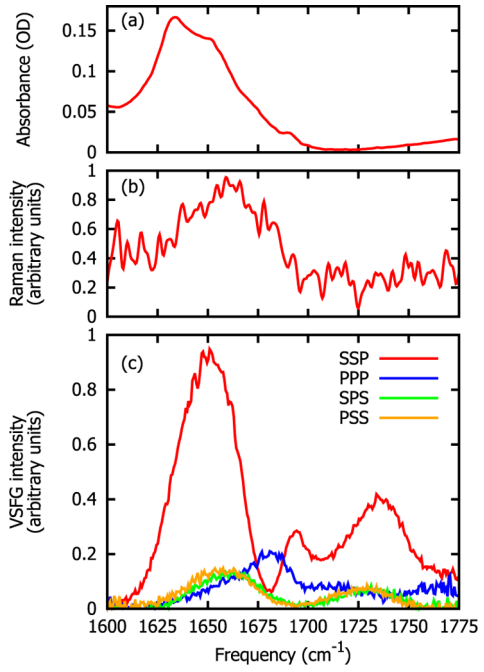


Figure 4. Experimental (a) IR spectrum of 20 μM CTB, (b) the Raman spectrum of a 0.2 mM solution of CTB, and (c) the VSFG spectrum of 43 nM CTB under 1:9 GM_1/d_{75}-DPPC monolayer with a surface pressure of 30 mN/m. In all experiments CTB was dissolved in a D_2O buffer solution of pH = 7.4 to create a biologically relevant environment.

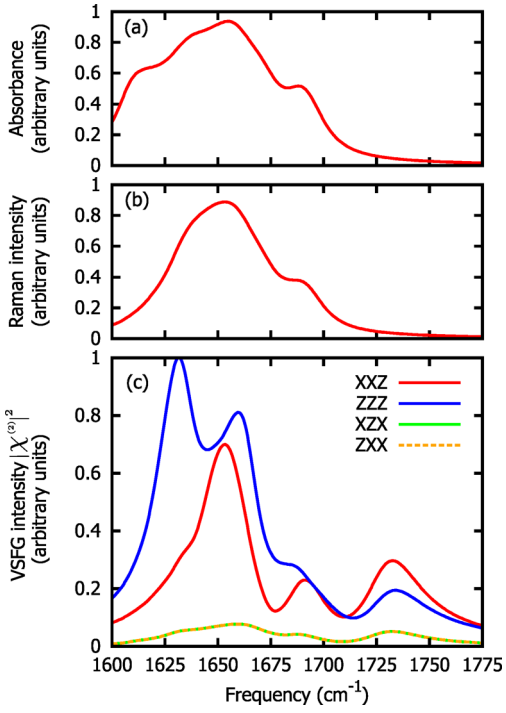


Figure 5. Calculated (a) IR and (b) Raman spectrum of CTB, and (c) lab frame second-order susceptibility of CTB summed with the C=O stretch peak of a monolayer of DPPC lipids. All VSFG spectra are scaled such that the *ZZZ* spectrum is normalized to unity.

dx.doi.org/10.1021/jp401159r | J. Phys. Chem. A 2013, 117, 6311−6322

accepting secondary amides this approach results in $\Delta\omega_{HB,acc.}$ = 20 cm^{-1}. For tertiary amides (i.e., the peptide bonds located upstream of a proline residue) $\Delta\omega_{HB,proline\ acc.}$ is downscaled to 16 cm^{-1} according to the ratio reported in ref 61. If we apply all proportionality constants as derived from ref 47, the calculated splitting between the two protein peaks in the experimental SSP spectrum is too small, and only one protein peak is calculated. Hence, to match the spectrum of CTB we set $\Delta\omega_{HB,donor}$ to 13 cm^{-1}, approximately two times the value reported in ref 47, and $\Delta\omega_{HB,2nd\ acc.}$, the hydrogen bond shift for second accepted hydrogen bonds, to 10 cm^{-1}, approximately half the value reported in ref 47. These adjustment are probably required because Hartree−Fock methods, like the one used in ref 47, generally do not predict hydrogen-bond shifts very well.[62] We assume that a hydrogen bond does not significantly influence the local-mode frequency if it does not meet the requirements described in ref 63. The software package Chimera[64] is used to perform this assessment.

As discussed in the Formalism section, the off-diagonal elements of the Hamiltonian are determined using a parametrized map of quantum-chemical calculations[27] for the nearest neighbors and the TDC model[45] for non-nearest neighbors. We find that using the transition-point-charge model[30] for non-nearest neighbor couplings results in significantly worse agreement with the experimental data (see Figure S1 of the Supporting Information). Subsequently, the Hamiltonian is diagonalized to attain its eigenvectors and eigenvalues. We then construct the transition-dipole moments and the principal Raman axes in the molecular frame for the conformation obtained from a time slice of the equilibrated MD simulation and combine them with the eigenvectors to obtain the IR and Raman responses of the normal modes. From these, using eqs 9−17, we calculate the IR and Raman spectra and the orientation-dependent VSFG spectra. As the measured VSFG spectrum is a sum of the protein response and the C═O-stretch response of the ester in the lipid head groups, we perform the calculation both for the amide-I normal modes of the protein and for the lipid C═O-stretch mode.

The formalism is thus based on a number of parameters that determine the shape of the calculated spectra: the orientational distribution functions $P(\theta_{protein}, \phi_{protein}, \psi_{protein})$ and $P(\theta_{lipid\text{-}C=O\ dipole}, \phi_{lipid\text{-}C=O\ dipole}, \psi_{lipid\text{-}C=O\ dipole})$ (which enter into eq 13, see also Figure 2), $\Gamma_{protein}$, Γ_{lipid}, an overall frequency offset, and the intensity ratio of the lipid response with respect to the protein response.

In order to obtain orientational information from the VSFG data, one has to assume a certain functional form of the orientational distribution in terms of the Euler angles, for example, a delta function or a Gaussian, for θ, ϕ, and/or ψ. Depending on the symmetry of the protein, the distribution can be dependent on one, two, or three of the Euler angles. CTB, like many membrane proteins, is expected to be distributed independent of ψ (due to approximate cylindrical symmetry of the protein) and ϕ, but dependent on θ (see also Figure 2). If a protein does not have cylindrical symmetry, the orientational probability distribution in general depends on all three Euler angles. Hence, the number of parameters that describe the orientational distribution is larger, which makes it more difficult to uniquely determine these parameters. However, if one assumes a not too complicated form for the distributions, it may still be possible to derive the average angles (and, if applicable, Gaussian widths) characterizing the orientation distribution from the VSFG spectra, as we expect the total set of band structures of multiple polarization combinations to be unique for a protein.

If we assume a delta-function distribution $P(\theta_{protein}) = \delta(\theta - \theta_{protein})$ for the θ-distribution of the protein and average over all possible values of ψ and ϕ in eq 13, we obtain the dependency of the VSFG-signals of CTB on $\theta_{protein}$ (by solving the equations in Appendix II of ref 54):

$$\chi^{(2)}_{XXZ} = \frac{1}{2}\Big((\tilde{\beta}^{(2)}_{xxz} + \tilde{\beta}^{(2)}_{yyz}) \times \cos\theta_{protein} \times \sin\theta_{protein} - \frac{1}{2}(\tilde{\beta}^{(2)}_{xxz} + \tilde{\beta}^{(2)}_{xzx} + \tilde{\beta}^{(2)}_{zxx} + \tilde{\beta}^{(2)}_{yyz} + \tilde{\beta}^{(2)}_{yzy} + \tilde{\beta}^{(2)}_{zyy} + 2\tilde{\beta}^{(2)}_{zzz}) \times \cos\theta_{protein} \times \sin^3\theta_{protein}\Big) \quad (19a)$$

$$\chi^{(2)}_{XZX} = \frac{1}{2}\Big((\tilde{\beta}^{(2)}_{xzx} + \tilde{\beta}^{(2)}_{yzy}) \times \cos\theta_{protein} \times \sin\theta_{protein} - \frac{1}{2}(\tilde{\beta}^{(2)}_{xxz} + \tilde{\beta}^{(2)}_{xzx} + \tilde{\beta}^{(2)}_{zxx} + \tilde{\beta}^{(2)}_{yyz} + \tilde{\beta}^{(2)}_{yzy} + \tilde{\beta}^{(2)}_{zyy} + 2\tilde{\beta}^{(2)}_{zzz}) \times \cos\theta_{protein} \times \sin^3\theta_{protein}\Big) \quad (19b)$$

$$\chi^{(2)}_{ZXX} = \frac{1}{2}\Big((\tilde{\beta}^{(2)}_{zxx} + \tilde{\beta}^{(2)}_{zyy}) \times \cos\theta_{protein} \times \sin\theta_{protein} - \frac{1}{2}(\tilde{\beta}^{(2)}_{xxz} + \tilde{\beta}^{(2)}_{xzx} + \tilde{\beta}^{(2)}_{zxx} + \tilde{\beta}^{(2)}_{yyz} + \tilde{\beta}^{(2)}_{yzy} + \tilde{\beta}^{(2)}_{zyy} + 2\tilde{\beta}^{(2)}_{zzz}) \times \cos\theta_{protein} \times \sin^3\theta_{protein}\Big) \quad (19c)$$

$$\chi^{(2)}_{ZZZ} = \frac{1}{2}(\tilde{\beta}^{(2)}_{xxz} + \tilde{\beta}^{(2)}_{xzx} + \tilde{\beta}^{(2)}_{zxx} + \tilde{\beta}^{(2)}_{yyz} + \tilde{\beta}^{(2)}_{yzy} + \tilde{\beta}^{(2)}_{zyy}) \times \cos\theta_{protein} \times \sin^3\theta_{protein} + \tilde{\beta}^{(2)}_{zzz} \times \cos^3\theta_{protein} \times \sin\theta_{protein} \quad (19d)$$

For CTB, assuming a delta function distribution for $\theta_{protein}$ is valid due to the very strong 5-fold binding between CTB and GM_1.[58,60] Furthermore, if we fit the spectra assuming that $P(\theta_{protein})$ is a Gaussian distribution, we obtain a very narrow width of this distribution, indicating that it is valid to assume a delta distribution. We also assume a delta function distribution for the lipid C═O dipoles, because applying a Gaussian distribution instead of a delta distribution does not improve the fit significantly. We assume $\theta_{lipid\text{-}C=O\ dipole}$ = 64°, the average angle between the dipole moment and the surface normal, as reported in ref 65. We assume that the nonresonant amplitude (see eq 15) is negligible compared to the resonant amplitude, as an experimentally measured concentration series shows that the spectral line shape does not significantly change for different concentrations.[66]

Deriving the Protein Orientation Distribution. *Challenges in Extracting Protein Orientation from Polarization-Dependent VSFG Measurements.* Ideally, it should be possible to use the relative magnitude of different polarization combinations to determine the protein orientation, similar to the orientation determination of uncoupled modes by previous authors.[33] However, for CTB, with the assumptions regarding the interfacial refractive indices used in ref 33, there is no orientation or orientational distribution that is consistent with either the observed relative magnitudes of the different polarization combinations or the line shape of the PPP-

dx.doi.org/10.1021/jp401159r | J. Phys. Chem. A 2013, 117, 6311−6322

polarization combination. This is probably both due to the uncertainty in the value of the interfacial refractive indices n' and due to the uncertainty of how to correctly model the surface when applying the Fresnel factors.

In previous VSFG papers all three n' values are assumed to be similar[33,57] and/or assumed to be similar to one of the bulk phases.[67] Zhuang et al. have derived a lower limit for n' from a simple estimate of the local-field correction at the interface using a modified Lorentz model.[33] However, the n' values can vary strongly as a function of molecular density and orientation, especially for interfaces consisting of large chromophores,[57,68] for example, from 1.4 to 1.9 for a 10° change in orientation of the interfacial liquid crystal molecule 48-*n*-octyl-4-cyanobiphenyl.

Because each of the polarization combinations is dependent on one or more interfacial indexes ($n'(\omega_{IR})$, $n'(\omega_{VIS})$, and $(n'(\omega_{SF}))$, incorrect values of n' lead to wrong predictions for the ratios of the SF signal of different polarization combinations for a given orientation (distribution). Inversely, incorrect n' values lead to false orientation distributions as derived from the measurements. Furthermore, oversimplifying the interface as only consisting of one boundary between air and water, where maybe multiple boundaries should be taken into account, also leads to incorrect orientation distributions.

Deriving the Protein Orientation Distribution Based on the Spectral VSFG-Line Shape. As an alternative for using the intensity ratio between different elements of the effective second-order susceptibility tensor (i.e., different polarization combinations), which requires accurate handling of the reflections at the interface (via the Fresnel factors), the line shape of the molecular second-order susceptibility of some elements can be fitted to the experimental data. This only works for polarization combinations that originate from a single molecular second-order susceptibility lab frame element, as the Fresnel factors constitute only an overall scaling factor in this case (for example, for the SSP-, SPS-, and PSS-elements). This approach cannot be applied to the PPP-element due to the more complex dependency on the Fresnel factors (see eq 16d). In the case of CTB, the SSP-response is well-suited to this approach, as the protein spectrum comprises two peaks: one high-frequency peak at 1690 cm^{-1} that originates from non-hydrogen-bonded amide moieties, and a low-frequency peak at 1645 cm^{-1} that originates from hydrogen-bonded moieties. As the corresponding parts of the protein have different orientations with respect to the membrane surface, the peak ratio is a function of the protein orientation (see Figure 6). Hence the 1645 cm^{-1}/1690 cm^{-1}-peak ratio can be used to derive the protein orientation, without any assumptions regarding n', nor on the implementation of Fresnel factors.

Calculated Spectra. The calculated spectra reproduce the experimental data reasonably well, as is evident from a comparison between Figures 4 and 5. While the IR and Raman spectrum are, of course, independent of orientation of the protein, for the calculation of the lab frame SFG susceptibilities, knowledge of specifically the angle θ is required to relate the molecular (x, y, z)- to the lab (X, Y, Z)-coordinate frame. By least-squares fitting the experimental SSP spectrum to the calculated *XXZ*-spectrum, we find that $\theta_{CTB} = 6° \pm 17°$. The residual sum of squares (RSS) increases by 10% if θ_{CTB} is fixed to a value larger than 30° and increases monotonously for increasing values of θ_{CTB} (see Figure 6, and also the Supporting Information for a graph of the RSS as a function of θ_{CTB}). Fits to the SPS and PSS susceptibility and a global fit to all three

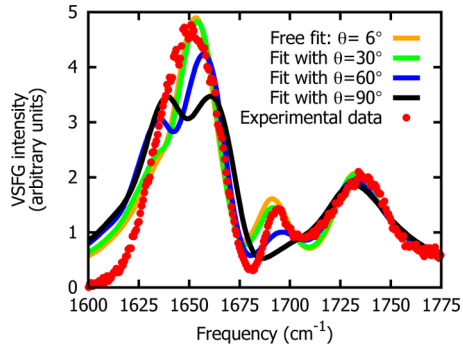


Figure 6. Calculated spectra as a function of protein orientation, assuming $\theta_{lipidC=Odipole} = 64°$.[65] This assumption is derived from an IRRAS measurement, so a value of −116° is equally likely. Spectra in which $\theta_{protein}$ and $\theta_{lipidC=Odipole}$ do not point in the same direction resemble the experimental spectra worse (see Supporting Information Figure S2). As homodyne SFG is not sensitive to an overall inversion of the orientation distributions over $\theta = 90°$ (i.e., it cannot distinguish a single dipole pointing up or down; only the relative orientation of different dipoles determines the signal), only spectra for $\theta_{protein} < 90°$ are shown.

(SSP, SPS, and PSS) hyperpolarizabilities (see Supporting Information Figure S3) lead to similarly small angles of θ_{CTB}. The RSS values increase in the same manner for the SPS and PSS fits as for the SSP fit for increasing values of θ_{CTB} (see Figure S4 of the Supporting Information). A θ_{CTB} close to 0°, as determined by previous authors,[58,60] can be expected from the 5-fold binding between the toxin's five subunits and the receptor lipids in the membrane.

The spectra in Figure 5 are calculated with the same isolated local-mode frequency Ω^0 and line width Γ for the IR, Raman, and VSFG spectra that we obtained by fitting the VSFG-SSP spectrum. The central frequencies of the three peaks in the IR spectrum (at 1630, 1655, and 1690 cm^{-1}) are reproduced correctly by the formalism. However, the relative magnitudes of the peaks at 1630 and 1655 cm^{-1} are not very well reproduced. This may be a result of the comparatively simple estimation of the hydrogen bond-induced amide-I red-shift that we apply here. Due to the very small Raman cross section of the protein, it is difficult to obtain a good signal-to-noise ratio in the measurement, but the line shape of the experimental Raman spectrum appears to be well-predicted by the formalism.

The formalism does not reproduce the polarization combination dependency of the VSFG spectra very well, probably as a result of the uncertainty in the precise values of the interfacial refractive indices and about how to model the interface correctly when applying the Fresnel factors. However, the *XXZ*-, *XZX*-, and *ZXX*-molecular hyperpolarizabilities in the lab frame (Figure 5) show a remarkable similarity in spectral shape with the experimental SSP-, SPS-, and PSS-VSFG spectra, respectively. The amplitudes of the calculated signals are determined by the refractive index at the interface, but the spectral line shape is not dependent on assumptions regarding the values of n' or the functional form of the Fresnel factors (see previous sections). The fact that the calculated spectral shapes correspond so well to the measured ones illustrates that the model correctly captures the vibrational sum frequency response of the interfacial protein. A better comparison between theory and model can be made by scaling the calculated to the experimental spectra, as shown in Figure 6 for the SSP spectrum, for several calculated spectra with a varying

dx.doi.org/10.1021/jp401159r | J. Phys. Chem. A 2013, 117, 6311−6322

angle θ_{CTB}. The deviation on the red side of the SSP spectrum is probably again largely due to the comparatively simple algorithm we apply for the hydrogen-bond induced red-shifts. Employing a more sophisticated hydrogen-bond shift algorithm[25,48] or an approach in which intermolecular interactions with the C═O and amide groups in the lipid monolayer are also taken into account may render the agreement even better.

The formalism allows for easy decomposition of the spectra into the contributions from different types of secondary structure motifs, as the different contributions can simply be turned on and off in the Hamiltonian. By performing such an analysis for CTB, we find that the most intense modes are the α-helical ones. This can partly explain the relatively large calculated *ZZZ*-molecular second-order susceptibility of CTB for small values of θ_{CTB} (see Figure 5). For such an orientation, there is a much larger contribution of the molecular-frame hyperpolarizability $\tilde{\beta}^{(2)}_{ZZZ}$ to the lab-frame susceptibility $\chi^{(2)}_{ZZZ}$ as compared to the other lab-frame susceptibilities (see eq 19). This reflects the fact that the polarization of the incoming and outgoing optical fields is parallel to the $\tilde{\beta}^{(2)}_{ZZZ}$ modes if $\theta_{CTB} \approx 0$. The $\tilde{\beta}^{(2)}_{ZZZ}$ modes are particularly strong because the A-modes of the α-helices are oriented parallel to the helical axis,[34] which in turn are parallel with the *z*-axis of the protein (see Figure 2).

The $\chi^{(2)}_{ZZZ}$ spectrum contains three major peaks at 1632, 1660, and 1690 cm^{-1}. We believe that the high-frequency peak is a result of α-helical residues that are near the end of the α-helices and are thus less strongly hydrogen bonded. The peak at 1660 cm^{-1} may be a result of α-helical residues that are more strongly hydrogen bonded. These hydrogen bonds can be both intramolecular to α-helical residue i+4 (i being the residue index) and intermolecular to water. Finally, there are also a few very exposed peptide bonds (and thus possibly hydrogen bonded to multiple water molecules) that, together with the perpendicular β-sheet modes, give rise to the peak at 1632 cm^{-1} (see Supporting Information for a graphical representation of the local mode contributions to these different normal modes). Interestingly, the intensity of the PPP-polarization combination is not much larger than that of the other polarization combinations that, for $\theta_{CTB} \approx 0°$, hardly contain any molecular *zzz*-contribution (see eqs 16 and 19). This is probably because the magnitude of the Fresnel factor for the *XXZ*-contribution is comparable to the *ZZZ*-contribution, and the two elements interfere destructively. Numerical evaluation of eq 16d shows that for a large range of values of n' this is indeed the case.

If the hydrogen bonds are not taken into account, the formalism predicts a single broad peak for the SSP spectrum of the protein (see Figure S5 of the Supporting Information for the influence of hydrogen bonding and solvation on the calculated spectrum). Running an MD simulation in which the protein is solvated results in a red-shifted and broadened low-frequency peak, as the hydrogen bonds between the water molecules and the protein backbone create a distribution in the red-shifts and a high-frequency peak that is less hydrogen bonded. This shows the strong dependency of the amide-I VSFG spectrum on the protein hydration, which can be used as a (time-dependent) probe of protein hydration near a membrane.

Experimental and Theoretical Considerations in Performing Protein VSFG Experiments. Apparently, for CTB the SSP-line shape is not very sensitive to the protein orientation for small values of $\theta_{protein}$ (see Figure S4 of the Supporting Information). If more elaborate theoretical and experimental studies improve our understanding of the local field effects that occur at a biomolecular interface (that determine the effective interfacial refractive index), also the line shape of the PPP spectrum and the relative magnitudes of the different polarization combinations can be used in combination with the presented formalism to determine protein distributions more accurately.

One way to obtain a more accurate description of these local field effects is using a CaF_2 crystal as the medium above the interface to decrease the contributions of non-*ZZZ* elements to the measured PPP signal.[15] Although this renders the shape of the PPP spectrum only dependent on the *ZZZ* susceptibility, the relative magnitude with respect to the other polarization combinations is still a function of n'. Alternatively, using the polarization null angle method[57] to derive the ratios between different lab frame hyperpolarizabilities from the measured SFG signal, is another way of reducing the influence of uncertainties in n′. Also, the information content of the measurement can be increased by performing heterodyne SFG experiments. In such experiments, the information loss that takes place in the observation of homodyne SFG signals (proportional to the modulus-square of the hyperpolarizability, see eq 14) is eliminated by measuring the interference between the homodyne signal and the signal from a local oscillator instead.[69−73] The resulting signal can be deconstructed into the imaginary part of $\chi^{(2)}$ that is most directly sensitive to the molecular orientation, which can also be predicted using our formalism. Finally, with the advent of 2D-SFG[74] a new experimental method is available to enhance the certainty of $\theta_{protein}$ determination. This technique, analogue to 2D-IR, directly reveals structural information through off-diagonal peaks that result from couplings between modes. Also, additional polarization combinations increase the (orientational) information content of the measurement.[75] Calculation of these spectra based on the protein conformation is a straightforward extension of the formalism presented here.

As can be seen by comparing Figures 4(a,b) with Figure 4(c), VSFG spectra have more informative line shapes than IR absorption or Raman spectra. This is due to the fact that it is a coherence-based technique; that is, the response is determined by a coherence between the $\nu = 0$ and the $\nu = 1$ state, induced by the IR field. This is in contrast to the IR absorption and Raman scattering processes. These are population transfer based techniques, in which the signal intensity is proportional to the sum of the modulus-squared normal mode transition-dipole moments and Raman polarizabilities, respectively (see eqs 9 and 10). However, in the case of SFG one is sensitive to the modulus-squared sum of the tensor product of the transition-dipole moment μ^{ν}_{k} and the Raman polarizability α^{ν}_{ij} of the normal modes (see eq 14). Hence, interference between different normal modes can strongly influence the observed spectra. This can for example be noted in the spectrum of CTB, where the relative orientation of the protein with respect to the C═O dipoles in the lipid head groups determines the shape of the spectrum in between their peaks: two modes pointing either in the same or in opposite directions give rise to a different spectrum.[76] By fitting the spectra we find that the C═O dipoles in the DPPC head groups are oriented in opposite direction of the protein's central *z*-axis (as defined in Figure 2, see also Figure S2 and S4 of the Supporting Information).

■ CONCLUSIONS

We have developed a formalism to predict the infrared, Raman, and VSFG amide-I spectra of proteins in and near interfaces,

dx.doi.org/10.1021/jp401159r | J. Phys. Chem. A 2013, 117, 6311−6322

based on their conformation. This formalism, which can be applied to any protein of which the conformation is known, makes it possible to identify surface-bound proteins and to determine their orientation with respect to the membrane. As a first application of the formalism we measure and calculate the VSFG spectrum of membrane-bound cholera toxin B subunit. To determine the orientation of this protein with respect to the surface, we use the dependence of the predicted VSFG spectral shape on the orientation. The strength of such an approach is that it does not rely on knowledge or estimates of the interfacial refractive indices at the three wavelengths relevant for the SFG process. Such knowledge is required when comparing signal intensities for SFG signals using different polarization combinations. Rather, our approach relies on optimizing the agreement between the calculated and measured spectral response at one specific polarization combination. The VSFG spectra can be well reproduced, and the orientation of the membrane-bound cholera toxin B subunit was determined to have its main axis oriented perpendicular to the surface normal of the membrane. Since the formalism presented here makes it possible to directly interpret amide-I VSFG spectra in terms of the protein conformation and orientation, we expect that it will significantly increase the potential of VSFG as a tool to study interfacial biomolecular processes.

EXPERIMENTAL DETAILS

Cholera toxin B subunit (CTB) is purchased as a lyophilized powder at Sigma-Aldrich, as well as the D_2O (99.96% purity). The GM_1 and d_{75}-DPPC lipids are purchased from Avanti Polar Lipids. The bulk IR spectrum is measured on a Bruker Vertex 70 FTIR spectrometer, in a CaF_2 sample cell with a 50 μm Teflon spacer, at a concentration of 20 μM in D_2O at pH 7.4, containing 0.5 M Tris buffer, 2 M NaCl, 30 mM NaN_3, and 10 mM sodium EDTA (the buffer salt constituents of the lyophilized powder). The Raman spectrum of a 200 μM solution of CTB in D_2O-PBS buffer is measured with a Bruker SENTERRA Raman microscope (λ_{exc} = 785 nm). The VSFG spectra are obtained after incubation of a planar ~30 mN/m binary lipid monolayer (consisting of 1:9 GM_1:d_{75}-DPPC) with a 12.5 nM toxin solution of pH = 7.4, for 30 min. We assume the docking process is completed after this period as the signal equilibrated after 15 min of incubation, and remained largely the same for 24 h of incubation. The sample was measured in a 20 mL coated Teflon trough. The VSFG spectroscopy setup we have used has been described in detail in a previous publication.[77] In short, we employ a regeneratively amplified Ti:sapphire system (Legend, Coherent, Inc.) that produces ~100 fs pulses at a frequency of 1 kHz, with a spectrum centered at 800 nm wavelength and a bandwidth of 12 nm. A part of the light (0.5 mJ/pulse) is split off and spectrally narrowed using an etalon, while 1 mJ/pulse is converted to tunable mid-IR pulses with an optical parametric amplifier and difference frequency generation unit. Each beam is passed through a half wave plate and a polarizer and is focused so as to impinge on the sample with angles of 62° (IR) and 58° (800 nm), respectively. After the sample the emitted sum frequency light is collimated, the remaining 800 nm light filtered out and the beam dispersed in a spectrograph, via a grating, and imaged using an an electron multiplied Charge Coupled Device camera. All measurements are performed at 22 °C.

The lengths of the intramolecular hydrogen bonds and the hydrogen bonds to surrounding water molecules are obtained by performing a 10 ps MD simulation in GROMACS in which the 2CHB PDB-entry was solvated in a 90 × 90 × 90 Å box (with 10 Å water around the protein), using the TIP3 force field for the water molecules and the CHARMM27 force field for the protein.

ASSOCIATED CONTENT

Supporting Information

See Supporting Information for additional information on: the calculations of the local-mode frequency shift of amides upstream of proline residues; a comparison of the calculated *XXZ* spectrum using the transition-charge coupling model and the transition-dipole coupling model; the dependency of the calculated *XXZ* spectrum on the protein orientation for all possible relative angles of the protein and the lipid C═O peak; the spectra obtained by performing a global fit of the calculated *XXZ*, *XZX*, and *ZXX* hyperpolarizability to the experimentally observed SSP, SPS, and PSS susceptibility; the residual sum of squares (RSS) of the fit of the *XXZ*, *XZX*, and *ZXX* susceptibility to the experimentally measured SSP, SPS, and PSS susceptibility, respectively, as a function of protein orientation; the influence of the hydrogen bonds on the calculated *XXZ* spectrum; the influence of the interfacial refractive index n' on the shape of the PPP spectrum; visualization of the eigenmodes of the three major peaks in the *ZZZ* spectrum. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: S.Woutersen@uva.nl; S.J.Roeters@uva.nl; Bonn@mpip-mainz.mpg.de.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Huib Bakker for critically reading the manuscript, Saeed Amirjalayer for performing the amide-I′ frequency calculations on the acetyl-derivatives, and Alex de Beer for contributing to the Mathematica script. S.J.R., C.N.vD. and S.W. acknowledge the European Research Council (ERC) for funding through grant 210999. A.T.-K. acknowledges The Netherlands Organisation for Scientific Research (NWO) for financial support. Finally, we also thank H. Menges for measuring the Raman spectrum.

REFERENCES

(1) Katsaras, J.; Epand, R. F.; Epand, R. M. Absence of Chiral Domains in Mixtures of Dipalmitoylphosphatidylcholine Molecules of Opposite Chirality. *Phys. Rev. E* **1997**, *55*, 3751−3753.

(2) Clarke, R. J.; Kane, D. J. Optical Detection of Membrane Dipole Potential: Avoidance of Fluidity and Dye-Induced Effects. *Biochim. Biophys. Acta, Biomembr.* **1997**, *1323*, 223−239.

(3) Barth, A. Infrared Spectroscopy of Proteins. *Biochim. Biophys. Acta, Bioenerg.* **2007**, *1767*, 1073−1101.

(4) Harrick, N. J. *Internal Reflection Spectroscopy*; Wiley: New York, 1967.

(5) Zhu, X. D.; Suhr, H.; Shen, Y. R. Surface Vibrational Spectroscopy by Infrared-Visible Sum Frequency Generation. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1987**, *35*, 3047−3050.

(6) Hess, C.; Bonn, M.; Funk, S.; Wolf, M. Hot-Band Excitation of CO Chemisorbed on Ru(001) Studied with Broadband-IR Sum-Frequency Generation. *Chem. Phys. Lett.* **2000**, *327*, 448−448.

(7) Diesner, M.-O.; Howell, C.; Kurz, V.; Verreault, D.; Koelsch, P. In Vitro Characterization of Surface Properties Through Living Cells. *J. Phys. Chem. Lett.* **2010**, *1*, 2339−2342.
(8) Barth, C.; Jakubczyk, D.; Kubas, A.; Anastassacos, F.; Brenner-Weiss, G.; Fink, K.; Schepers, U.; Bräse, S.; Koelsch, P. Interkingdom Signaling: Integration, Conformation, and Orientation of N-Acyl- L-Homoserine Lactones in Supported Lipid Bilayers. *Langmuir* **2012**, *28*, 8456−8462.
(9) Miranda, P. B.; Shen, Y. R. Liquid Interfaces: a Study by Sum-Frequency Vibrational Spectroscopy. *J. Phys. Chem. B* **1999**, *103*, 3292−3307.
(10) Kim, J.; Cremer, P. S. IR-Visible SFG Investigations of Interfacial Water Structure Upon Polyelectrolyte Adsorption at the Liquid/Solid Interface. *J. Am. Chem. Soc.* **2000**, *122*, 12371−12372.
(11) Kim, G.; Gurau, M. C.; Lim, S.-M.; Cremer, P. S. Investigations of the Orientation of a Membrane Peptide by Sum Frequency Spectroscopy. *J. Phys. Chem. B* **2003**, *107*, 1403−1409.
(12) Liu, J.; Conboy, J. C. Direct Measurement of the Transbilayer Movement of Phospholipids by Sum-Frequency Vibrational Spectroscopy. *J. Am. Chem. Soc.* **2004**, *126*, 8376−8377.
(13) Roke, S.; Schins, J.; Müller, M.; Bonn, M. Vibrational Spectroscopic Investigation of the Phase Diagram of a Biomimetic Lipid Monolayer. *Phys. Rev. Lett.* **2003**, *90*, 128101−4.
(14) Ma, G.; Allen, H. DPPC Langmuir Monolayer at the Air-Water Interface: Probing the Tail and Head Groups by Vibrational Sum Frequency Generation Spectroscopy. *Langmuir* **2006**, *22*, 5341−5349.
(15) Wang, J.; Even, M. A.; Chen, X.; Schmaier, A. H.; Waite, J. H.; Chen, Z. Detection of Amide I Signals of Interfacial Proteins in Situ Using SFG. *J. Am. Chem. Soc.* **2003**, *125*, 9914−9915.
(16) Rocha-Mendoza, I.; Yankelevich, D. R.; Wang, M.; Reiser, K. M.; Frank, C. W.; Knoesen, A. Sum Frequency Vibrational Spectroscopy: the Molecular Origins of the Optical Second-Order Nonlinearity of Collagen. *Biophys. J.* **2007**, *93*, 4433−4444.
(17) Phillips, D. C.; York, R. L.; Mermut, O.; Mccrea, K. R.; Ward, R. S.; Somorjai, G. A. Side Chain, Chain Length, and Sequence Effects on Amphiphilic Peptide Adsorption at Hydrophobic and Hydrophilic Surfaces Studied by Sum-Frequency Generation Vibrational Spectroscopy and Quartz Crystal Microbalance. *J. Phys. Chem. C* **2007**, *111*, 255−261.
(18) Fu, L.; Ma, G.; Yan, E. C. Y. In Situ Misfolding of Human Islet Amyloid Polypeptide at Interfaces Probed by Vibrational Sum Frequency Generation. *J. Am. Chem. Soc.* **2010**, *132*, 5405−5412.
(19) vandenAkker, C. C.; Engel, M. F. M.; Velikov, K. P.; Bonn, M.; Koenderink, G. H. The Polyphenol EGCG Inhibits Amyloid Formation Less Efficiently at Phospholipid Interfaces Than in Bulk Solution. *J. Am. Chem. Soc.* **2011**, *133*, 18030−18033.
(20) Liu, Y.; Jasensky, J.; Chen, Z. Molecular Interactions of Proteins and Peptides at Interfaces Studied by Sum Frequency Generation Vibrational Spectroscopy. *Langmuir* **2012**, *28*, 2113−2121.
(21) Lee, S.; Krimm, S. General Treatment of Vibrations of Helical Molecules and Application to Transition Dipole Coupling in Amide I and Amide II Modes of α-Helical Poly (-Alanine). *Chem. Phys.* **1998**, *230*, 277−295.
(22) Torii, H.; Tasumi, M. Model Calculations on the Amide-I Infrared Bands of Globular Proteins. *J. Chem. Phys.* **1992**, *96*, 1−9.
(23) Torii, H.; Tasumi, M. Ab Initio Molecular Orbital Study of the Amide I Vibrational Interactions Between the Peptide Groups in Di- and Tripeptides and Considerations on the Conformation of the Extended Helix. *J. Raman Spectrosc.* **1998**, *29*, 81−86.
(24) Choi, J.-H.; Cho, M. Calculations of Intermode Coupling Constants and Simulations of Amide I, II, and III Vibrational Spectra of Dipeptides. *Chem. Phys.* **2009**, *361*, 168−175.
(25) Wang, L.; Middleton, C. T.; Zanni, M. T.; Skinner, J. L. Development and Validation of Transferable Amide I Vibrational Frequency Maps for Peptides. *J. Phys. Chem. B* **2011**, *115*, 3713−3724.
(26) Schweitzer-Stenner, R. Simulated IR, Isotropic and Anisotropic Raman, and Vibrational Circular Dichroism Amide I Band Profiles of Stacked β-Sheets. *J. Phys. Chem. B* **2012**, *116*, 4141−4153.
(27) Gorbunov, R. D.; Kosov, D. S.; Stock, G. Ab Initio-Based Exciton Model of Amide I Vibrations in Peptides: Definition, Conformational Dependence, and Transferability. *J. Chem. Phys.* **2005**, *122*, 224904.
(28) Falvo, C.; Zhuang, W.; Kim, Y. S.; Axelsen, P. H.; Hochstrasser, R. M.; Mukamel, S. Frequency Distribution of the Amide-I Vibration Sorted by Residues in Amyloid Fibrils Revealed by 2D-IR Measurements and Simulations. *J. Phys. Chem. B* **2012**, *116*, 3322−3330.
(29) Ganim, Z.; Chung, H. S.; Smith, A. W.; Deflores, L. P.; Jones, K. C.; Tokmakoff, A. Amide I Two-Dimensional Infrared Spectroscopy of Proteins. *Acc. Chem. Res.* **2008**, *41*, 432−441.
(30) Hamm, P.; Lim, M. H.; Degrado, W.; Hochstrasser, R. M. The Two-Dimensional IR Nonlinear Spectroscopy of a Cyclic Penta-Peptide in Relation to Its Three-Dimensional Structure. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2036−2041.
(31) Huang, Q.; Schweitzer-Stenner, R. Conformational Analysis of Tetrapeptides by Exploiting the Excitonic Coupling Between Amide I Modes. *J. Raman Spectrosc.* **2004**, *35*, 586−591.
(32) Tsuboi, M.; Benevides, J.; Thomas, G. J., Jr. Raman Tensors and Their Application in Structural Studies of Biological Systems. *Proc. Jpn. Acad., B* **2009**, *85*, 83−97.
(33) Zhuang, X.; Miranda, P. B.; Kim, D.; Shen, Y. R. Mapping Molecular Orientation and Conformation at Interfaces by Surface Nonlinear Optics. *Phys. Rev. B* **1999**, *59*, 12632−12640.
(34) Nguyen, K. T.; Le Clair, S. V.; Ye, S.; Chen, Z. Orientation Determination of Protein Helical Secondary Structures Using Linear and Nonlinear Vibrational Spectroscopy. *J. Phys. Chem. B* **2009**, *113*, 12169−12180.
(35) Nguyen, K. T.; King, J. T.; Chen, Z. Orientation Determination of Interfacial β-Sheet Structures in Situ. *J. Phys. Chem. B* **2010**, 12169−12180.
(36) Perry, J. M.; Moad, A. J.; Begue, N. J.; Wampler, R. D.; Simpson, G. J. Electronic and Vibrational Second-Order Nonlinear Optical Properties of Protein Secondary Structural Motifs. *J. Phys. Chem. B* **2005**, *109*, 20009−20026.
(37) Moad, A. J.; Moad, C. W.; Perry, J. M.; Wampler, R. D.; Goeken, G. S.; Begue, N. J.; Shen, T.; Heiland, R.; Simpson, G. J. Nlopredict: Visualization and Data Analysis Software for Nonlinear Optics. *J. Comput. Chem.* **2007**, *28*, 1996−2002.
(38) Xiao, D.; Fu, L.; Liu, J.; Batista, V. S.; Yan, E. C. Y. Amphiphilic Adsorption of Human Islet Amyloid Polypeptide Aggregates to Lipid/Aqueous Interfaces. *J. Mol. Biol.* **2012**, *421*, 537−547.
(39) Tsuboi, M.; Ikeda, T. Raman Microscopy of a Small Uniaxial Crystal: Tetragonal Aspartame. *J. Raman Spectrosc.* **1991**, *22*, 619−626.
(40) Barth, A.; Zscherp, C. What Vibrations Tell About Proteins. *Q. Rev. Biophys.* **2002**, *35*, 369−430.
(41) Measey, T.; Hagarman, A.; Eker, F.; Griebenow, K.; Schweitzer-Stenner, R. Side Chain Dependence of Intensity and Wavenumber Position of Amide I′ in IR and Visible Raman Spectra of XA and AX Dipeptides. *J. Phys. Chem. B* **2005**, *109*, 8195−8205.
(42) Krimm, S.; Abe, Y. Intermolecular Interaction Effects in the Amide I Vibrations of β Polypeptides. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 2788.
(43) Lee, C.; Cho, M. Local Amide I Mode Frequencies and Coupling Constants in Multiple-Stranded Antiparallel β-Sheet Polypeptides. *J. Phys. Chem. B* **2004**, *108*, 20397−20407.
(44) Hamm, P.; Zanni, M. *Concepts and Methods of 2D Infrared Spectroscopy*; Cambridge University Press: Cambridge, U.K., 2011.
(45) Hamm, P.; Lim, M.; Hochstrasser, R. M. Structure of the Amide I Band of Peptides Measured by Femtosecond Nonlinear-Infrared Spectroscopy. *J. Phys. Chem. B* **1998**, *102*, 6123−6138.
(46) Hamm, P.; Woutersen, S. Coupling of the Amide I Modes of the Glycine Di-Peptide. *Bull. Chem. Soc. Jpn.* **2002**, *75*, 985−988.
(47) Ham, S.; Kim, J.-H.; Lee, H.; Cho, M. Correlation Between Electronic and Molecular Structure Distortions and Vibrational Properties. II. Amide I Modes of NMA-nD_2O Complexes. *J. Chem. Phys.* **2003**, *118*, 3491−3498.
(48) Jansen, L. C. T.; Knoester, J. A Transferable Electrostatic Map for Solvation Effects on Amide I Vibrations and Its Application to

Linear and Two-Dimensional Spectroscopy. *J. Chem. Phys.* **2006**, *124*, 044502.
(49) Pajcini, V.; Chen, X. G.; Bormett, R. W.; Geib, S. J.; Li, P.; Asher, S. A.; Lidiak, E. G. Glycylglycine $\pi \rightarrow \pi^*$ and Charge Transfer Transition Moment Orientations: Near-Resonance Raman Single-Crystal Measurements. *J. Am. Chem. Soc.* **1996**, *118*, 9716−9726.
(50) Arfken, G. B.; Weber, H. J.; Harris, F. E. *Mathematical Methods for Physicists*; Academic Press: New York, 2012.
(51) Schweitzer-Stenner, R.; Eker, F.; Griebenow, K.; Cao, X.; Nafie, L. A. The Conformation of Tetraalanine in Water Determined by Polarized Raman, FT-IR, and VCD Spectroscopy. *J. Am. Chem. Soc.* **2004**, *126*, 2768−2776.
(52) Schweitzer-Stenner, R. Dihedral Angles of Tripeptides in Solution Directly Determined by Polarized Raman and FTIR Spectroscopy. *Biophys. J.* **2002**, *83*, 523−532.
(53) Miyazawa, T. Perturbation Treatment of the Characteristic Vibrations of Polypeptide Chains in Various Configurations. *J. Chem. Phys.* **1960**, *32*, 1647.
(54) Moad, A.; Simpson, G. A Unified Treatment of Selection Rules and Symmetry Relations for Sum-Frequency and Second Harmonic Spectroscopies. *J. Phys. Chem. B* **2004**, *108*, 3548−3562.
(55) Sandeman, I. Amide Bands in Infra-Red Spectra: the Direction of the Transition Moments of Bands in N, N′-Diacetylhexamethylenediamine. *Proc. R. Soc. London, A* **1955**, *232*, 105−113.
(56) Moore, W. H.; Krimm, S. Transition Dipole Coupling in Amide I Modes of β-Polypeptides. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 4933.
(57) Wang, H. F.; Gan, W.; Lu, R.; Rao, Y.; Wu, B. H. Quantitative Spectral and Orientational Analysis in Surface Sum Frequency Generation Vibrational Spectroscopy (SFG-VS). *Int. Rev. Phys. Chem.* **2005**, *24*, 191−256.
(58) Haan, L.; Hirst, T. R. Cholera Toxin: a Paradigm for Multi-Functional Engagement of Cellular Mechanisms (Review). *Mol. Membr. Biol.* **2004**, *21*, 77−92.
(59) Shi, J.; Yang, T.; Kataoka, S.; Zhang, Y.; Diaz, A. J.; Cremer, P. S. GM1 Clustering Inhibits Cholera Toxin Binding in Supported Phospholipid Membranes. *J. Am. Chem. Soc.* **2007**, *129*, 5954−5961.
(60) Miller, C. E.; Majewski, J.; Faller, R.; Satija, S.; Kuhl, T. L. Cholera Toxin Assault on Lipid Monolayers Containing Ganglioside GM1. *Biophys. J.* **2004**, *86*, 3700−3708.
(61) Lessing, J.; Roy, S.; Reppert, M.; Baer, M.; Marx, D.; Jansen, T. L. C.; Knoester, J.; Tokmakoff, A. Identifying Residual Structure in Intrinsically Disordered Systems: a 2D IR Spectroscopic Study of the GVGXPGVG Peptide. *J. Am. Chem. Soc.* **2012**, *134*, 5032−5035.
(62) Campen, R. K.; Kubicki, J. D. Interaction Energy and the Shift in OH Stretch Frequency on Hydrogen Bonding for the $H_2O \rightarrow H_2O$, $CH_3OH \rightarrow H_2O$, and $H_2O \rightarrow CH_3OH$ Dimers. *J. Comput. Chem.* **2009**, *31*, 963−972.
(63) Mills, J. E. J.; Dean, P. M. Three-Dimensional Hydrogen-Bond Geometry and Probability Information from a Crystal Survey. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 607−622.
(64) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E.; Ferrin, T. E. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605−12.
(65) Gericke, A.; Flach, C. R.; Mendelsohn, R. Structure and Orientation of Lung Surfactant SP-C and L-Alpha-Dipalmitoylphosphatidylcholine in Aqueous Monolayers. *Biophys. J.* **1997**, *73*, 492−499.
(66) Roeters, S. J.; Campen, R. K.; Bonn, M. VSFG Study on the Interaction of Cholera Toxin B-Subunit with a GM1/DPPC Binary Monolayer, in preparation.
(67) Tyrode, E.; Johnson, C.; Baldelli, S.; Leygraf, C. A Vibrational Sum Frequency Spectroscopy Study of the Liquid-Gas Interface of Acetic Acid-Water Mixtures: 2. Orientation Analysis. *J. Phys. Chem. B* **2005**, *109*, 329−341.
(68) Ye, P.; Shen, Y. R. Local-Field Effect on Linear and Nonlinear Optical Properties of Adsorbed Molecules. *Phys. Rev. B* **1983**, *28*, 4288−4294.
(69) Ostroverkhov, V.; Waychunas, G. A.; Shen, Y. R. New Information on Water Interfacial Structure Revealed by Phase-Sensitive Surface Spectroscopy. *Phys. Rev. Lett.* **2005**, *94*, 046102.
(70) Stiopkin, I. V.; Jayathilake, H. D.; Bordenyuk, A. N.; Benderskii, A. V. Heterodyne-Detected Vibrational Sum Frequency Generation Spectroscopy. *J. Am. Chem. Soc.* **2008**, *130*, 2271.
(71) Nihonyanagi, S.; Yamaguchi, S.; Tahara, T. Direct Evidence for Orientational Flip-Flop of Water Molecules at Charged Interfaces: a Heterodyne-Detected Vibrational Sum Frequency Generation Study. *J. Chem. Phys.* **2009**, *130*, 204704.
(72) Chen, X.; Hua, W.; Huang, Z.; Allen, H. C. Interfacial Water Structure Associated with Phospholipid Membranes Studied by Phase-Sensitive Vibrational Sum Frequency Generation Spectroscopy. *J. Am. Chem. Soc.* **2010**, *132*, 11336.
(73) Pool, R. E.; Versluis, J.; Backus, E. H. G.; Bonn, M. Comparative Study of Direct and Phase-Specific Vibrational Sum-Frequency Generation Spectroscopy: Advantages and Limitations. *J. Phys. Chem. B* **2011**, *115*, 15362−15369.
(74) Bredenbeck, J.; Ghosh, A.; Smits, M.; Bonn, M. Ultrafast Two Dimensional-Infrared Spectroscopy of a Molecular Monolayer. *J. Am. Chem. Soc.* **2008**, *130*, 2152−2153.
(75) Laaser, J. E.; Zanni, M. T. Extracting Structural Information from the Polarization Dependence of One- and Two-Dimensional Sum Frequency Generation Spectra. *J. Phys. Chem. A* **2012**, DOI: 10.1021/jp307721y.
(76) Sovago, M.; Vartiainen, E.; Bonn, M. Determining Absolute Molecular Orientation at Interfaces: a Phase Retrieval Approach for Sum Frequency Generation Spectroscopy. *J. Phys. Chem. C* **2009**, *113*, 6100−6106.
(77) Smits, M.; Sovago, M.; Wurpel, G. W. H.; Kim, D.; Muller, M.; Bonn, M. Polarization-Resolved Broad-Bandwidth Sum-Frequency Generation Spectroscopy of Monolayer Relaxation. *J. Phys. Chem. C* **2007**, *111*, 8878−8883.

■ NOTE ADDED IN PROOF

A computational study closely related to the formalism presented in this article has recently been published by Liang, C. et al., in *J. Phys. Chem. Lett.* **2013**, *4*, 448−452.