THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

pubs.acs.org/JPCA

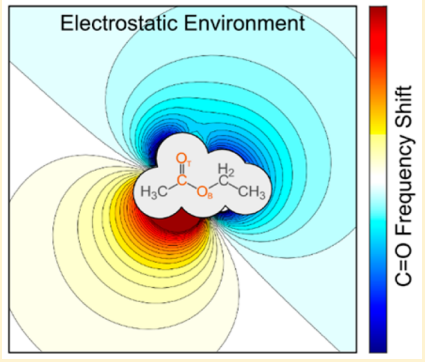
An Empirical IR Frequency Map for Ester C═O Stretching Vibrations

Sean C. Edington, Jennifer C. Flanagan, and Carlos R. Baiz*

Department of Chemistry, University of Texas at Austin, 105 E. 24th St. Stop A5300, Austin, Texas 78712-1224, United States

Supporting Information

ABSTRACT: We present an approach for parametrizing spectroscopic maps of carbonyl groups against experimental IR absorption spectra. The model correlates electric fields sampled from molecular dynamics simulations with vibrational frequencies and line shapes in different solvents. We perform an exhaustive search of parameter combinations and optimize the parameter values for the ester carbonyl stretching mode in ethyl acetate by comparing to experimental FTIR spectra of the small molecule in eight different solvents of varying polarities. Hydrogen-bonding solvents require that the peaks are fit independently for each hydrogen bond ensemble to compensate for improper sampling in molecular dynamics simulations. Spectra simulated using the optimized electrostatic map reproduce C═O IR absorption spectra of ethyl acetate with a line center RMSD error of 4.9 cm^{-1} over 12 different solvents whose measured line centers span a 45 cm^{-1} range. In combination with molecular dynamics simulations, this spectroscopic map will be useful in interpreting spectra of ester groups in heterogeneous environments such as lipid membranes.



1. BACKGROUND

Carbonyl stretching modes are useful probes of molecular dynamics due to their large oscillator strengths and sensitivity to solvent environments. The local backbone amide I modes in proteins, which consist primarily of C═O stretching and N−H bending motions, are used extensively as reporters of protein secondary structure, backbone disorder, and folding.[2−5] Carbonyls are also present in countless other biological molecules, including DNA and lipids. However, carbonyl vibrations other than the amide I modes remain underutilized as probes of structure and dynamics in biological systems. This can be attributed to multiple factors: (1) In the 1950s, a relation was discovered between protein structure and amide I absorption spectra. This work drove the early development and subsequent widespread use of IR spectroscopy for structural analysis.[6−12] (2) In turn, spectroscopic models that facilitate the interpretation of amide I spectra were extensively developed. Models for similar carbonyl stretches in ketones, acetyls, aldehydes, and esters remain comparatively undeveloped.[3,13−26] (3) With some exceptions, the protein backbone represents a uniform array of oscillators with one-to-one correspondence between oscillators and protein residues. Backbone C═O stretches are vibrationally coupled and form delocalized modes that reflect the underlying arrangement of the oscillators, namely the protein structure.[3,27−29]

Ester carbonyl stretches in lipid bilayers show substantial spectral changes with respect to composition and local lipid order, but interpretation of spectra has thus far been largely phenomenological.[30−35] Unlike proteins, lipid membranes encompass multiple environments, varying compositions, and rapidly changing molecular orientations. These attributes complicate the interpretation of experiments.[36−40] However, ester carbonyls are of special importance for understanding the cell membrane interface since they are located within the nanometer thick boundary region between solvated, polar lipid heads and aliphatic lipid tails.[41] Theoretical models are needed to aid in interpreting IR spectra of these complex environments. In this paper, we report the development of an empirical approach for parametrizing electrostatic models using experimental IR absorption spectra of ethyl acetate in a range of solvents.

Electrostatic frequency maps are useful for modeling the amide I bands of proteins and providing a connection between molecular structure and IR spectra. These tools facilitate the interpretation of linear and nonlinear IR line shapes in terms of molecular conformations.[5,42−45] The theoretical basis for this approach relies on the modulation of transition frequencies by changes in the underlying potential energy surface along the vibrational coordinate. Thus, frequency fluctuations largely depend on the forces exerted on the oscillator by the bath. Since the origin of these forces is primarily electrostatic in polar environments, frequency fluctuations are correlated with the electrostatic environment around an oscillator. More formally, the instantaneous frequency of an oscillator can be expressed as a linear combination of the electrostatic parameters at different positions. Atomic positions are usually used for simplicity:

Received: March 20, 2016
Revised: May 21, 2016

DOI: 10.1021/acs.jpca.6b02887
J. Phys. Chem. A XXXX, XXX, XXX−XXX

$$\omega = \omega_0 + \sum_{n=1}^{N\ \text{sites}} \left(C_n^{\Phi}\Phi_n + \sum_{i=x,y,z} C_{n,i}^{F} \frac{\partial \Phi_n}{\partial x_i} + \sum_{i,j=x,y,z} C_{n,i}^{G} \frac{\partial^2 \Phi_n}{\partial x_i\, \partial x_j} \right) \quad (1)$$

where Φ_n represents the electrostatic potential at a specific site and ω_0 is the zero-field frequency. We express the electric fields and gradients as first and second partial derivatives of the potential to emphasize that these terms are interdependent. Coefficients C_n^{Φ}, $C_{n,i}^{F}$, and $C_{n,i}^{G}$ represent the map parameters for the electrostatic potential Φ, electric field F, and gradients G, respectively. Each site supports up to ten parameters. In short, the above equation converts the electrostatic environment around an oscillator into a vibrational frequency shift. Frequency trajectories and their derived IR spectra are typically obtained from molecular dynamics simulations as described below.

2. EXPERIMENTAL METHODS AND THEORETICAL MODELS

2.1. Selection of Solute and Solvents. Ethyl acetate (Figure 1) was selected as our ester probe. It is soluble in a wide range of solvents and cannot interact with itself through hydrogen bonding as it does not donate hydrogen bonds.[46] Our selection of solvent was guided by the following criteria: (1) Solvents must be commonly used in chemistry and biology and must be liquid at room temperature. (2) Solvents must not contain C═O groups which would spectrally overlap with our probe molecule. (3) Ethyl acetate, the probe, must be able to sample a range of hydrogen-bonding environments. (4) Solvents must emulate the range of polarities encountered in biological samples. Alkanes are poorly modeled by the electrostatic approach we use and are excluded from the set. Table 1 lists the solvents used in this study along with their polarities. Solvents were selected to sample a wide variety of polarities, with more focus on the polar end of the spectrum.

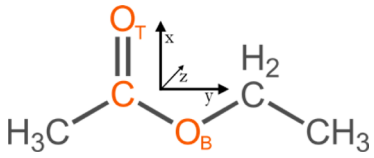


Figure 1. Molecular structure of ethyl acetate along with the molecular coordinate system used here. The z-axis points into the paper. Atoms used as sites for the electrostatic map are indicated in orange. The terminal and bridging oxygen atoms are labeled O_T and O_B, respectively.

2.2. Experimental FTIR: Data Collection and Fitting. Solutions of ethyl acetate (Sigma, 99.8%, anhydrous) at 20 mg/mL concentration were prepared in diethyl ether (Fisher, 99.9%, anhydrous), tetrahydrofuran (Fisher, 99.9%), DMSO (Fisher, 99.9%), acetonitrile (Fisher, 99.9%), methanol (Sigma, 99.8%, anhydrous), ethanol (Sigma, 99.5%), hexanol (Sigma, 98%), and D_2O (Cambridge Isotopes, 99.8%). Solvents were dried over 3-Å-pore-size molecular sieves for 12 h before use.

Spectra were collected at 2 cm^{-1} resolution on a Bruker Vertex 70 FTIR spectrometer equipped with a DTGS detector. A 50−100 μL volume of each solution was placed between two CaF_2 windows separated by a 50 μm PTFE spacer and secured within a sample cell.

Experimental IR absorption spectra (Figure 2) were fit to profiles defined by either 1, 2, or 3 Lorentzian line shapes using the curve-fitting tool provided by the MATLAB R2015b package. The multiple peaks observed in hydrogen-bonding solvents are attributed to ensembles containing different numbers of hydrogen bonds to the terminal C═O oxygen atom.[47] Experimental fit parameters including center frequencies, line widths, and hydrogen bond ensemble populations are provided in sections S1 and S2 of the Supporting Information.

Table 1. List of Solvents Used for Constructing the Electrostatic Maps[a]

solvent	polarity	solvent	polarity
diethyl ether	0.12	hexanol	0.55
tetrahydrofuran (THF)	0.21	ethanol	0.65
dimethyl sulfoxide (DMSO)	0.44	methanol	0.76
acetonitrile	0.46	deuterium oxide (D_2O)	0.99

[a]Polarity values are from Reichard's normalized solvatochromic scale.[1]

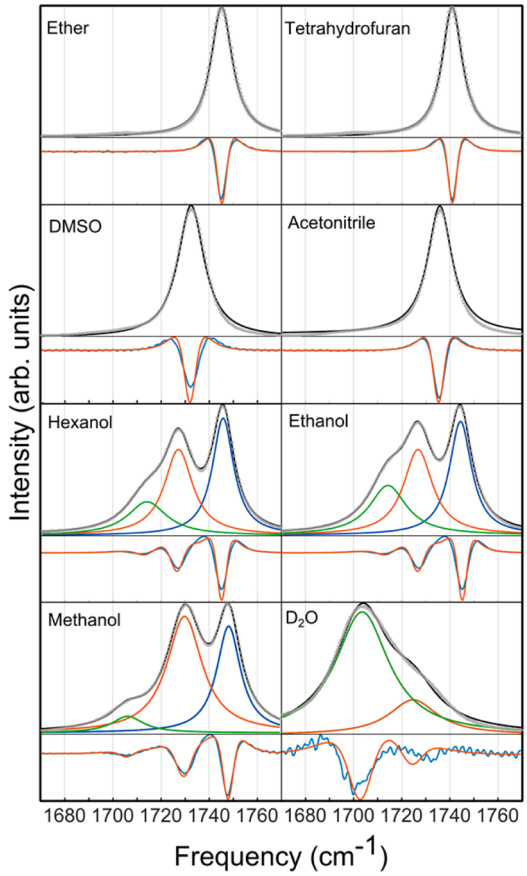


Figure 2. Experimental IR spectra (gray) of ethyl acetate in the eight solvents used to construct the map. The Lorentzian fits are shown as black curves. The component Lorentzian functions (blue, red, green) are shown for 0, 1, and 2 hydrogen bond ensembles in protic solvents. The second-derivative spectrum (blue curves) $d^2I(\omega)/d\omega^2$ is shown along with the second derivative of the fits (red curves) underneath each spectrum.

2.3. MD Simulations and Electrostatic Calculations. MD simulations were carried out using the generalized Amber force field (GAFF).[48,49] GAFF topologies for solvent and solute molecules were generated using the Amber 12 tools package. The topologies were converted to GROMACS format and used for all subsequent simulations. A 4 nm cubic solvent box was

DOI: 10.1021/acs.jpca.6b02887
J. Phys. Chem. A XXXX, XXX, XXX−XXX

first equilibrated for 200 ps in an NVT ensemble followed by a 2 ns equilibration in an NPT ensemble. The solute was then inserted into the equilibrated box and further equilibrated for 1 ns. Bonds were constrained using the LINCS algorithm in all simulations. Long-range electrostatics were treated using particle mesh Ewald summations. Production trajectories were run for 10 ns (2 fs steps) in an NPT ensemble at 300 K and 1 atm using the Berendsen thermostat.[50] Snapshots were stored every 20 fs for analysis. All MD simulations were carried out using the GROMACS 4.5.3 package.[51]

Electrostatic potentials, fields, and gradients were computed at the O_T, C, and O_B atomic positions (Figure 1) of the ester molecule using the assigned GAFF atomic charges. Contributions from all atoms in the 4 nm box are included without cutoffs, except contributions from the two carbon atoms covalently bonded to the ester group, which were excluded from the electrostatic calculations.

2.4. Spectral Simulations. Once a frequency trajectory, $\omega(t)$, was obtained, the linear absorption spectrum was computed using the following optical response function:[52]

$$A(\omega) = \mathrm{Im}\left[\int_0^\infty U(\tau) e^{-\tau/2T_1} e^{i(\omega+\omega_R)\tau}\, d\tau\right] \quad (2)$$

where

$$U(\tau) = \left\langle \exp\left(-\frac{i}{\hbar}\int_0^\tau [\omega(t) - \omega_R]\, dt\right)\right\rangle \quad (3)$$

The angled brackets indicate an average over the ensemble, and ω_R is an arbitrary rotating-frame frequency typically selected to coincide with the average frequency of the trajectory. Here, the initial 10 ns trajectories were subdivided into 500 partially overlapping 100 ps subtrajectories. The absorption spectrum $A(\omega)$ is analogous to a rotating-frame single-sided Fourier transform of $U(\tau)$. T_1 is an ad-hoc term that represents broadening due to population relaxation and was set to 1.5 ps for all simulations described here.[15,47,52]

2.5. Map Coefficient Optimization Procedure. The fit was performed by directly minimizing the least-squares differences between simulated spectra and normalized IR absorption spectra, except for the three alcohols. In the case of the alcohols, each hydrogen-bond (HB) ensemble was fit independently to the Lorentzian fits of the FTIR spectra (Figure 2 and Table S2). D_2O was fit directly to the experimental spectrum, but the MD trajectory was reweighted to reflect the correct number of 1 and 2 HB ensembles. This is due to the fact that the uncertainties in the Lorentzian line widths are larger in D_2O compared to other solvents (Table S2). The global parameter optimization requires generating a total of 12 individual spectra for each coefficient evaluation within the optimization routine.

We used a genetic algorithm (GA) to sample the parameter space and find the optimum starting point for direct optimization. The initial GA population was set to 50 individuals, the elite count set to 3, and the number of generations set to 20. The GA was run 50 consecutive times starting from random initial populations. Once the initial GA optimization was complete, the best scoring parameter values were used as input for a direct, unconstrained, quasi-Newton minimization. Optimization algorithms were used as implemented in the MATLAB R2015a package. Computational costs can be reduced by performing the GA optimization using average frequencies instead of the overall spectrum. In experimental spectra, these correspond to the center of the Lorentzian bands. Finally, the least-squares minimization was carried out on full IR spectra, as outlined above, in order to take into account the peak centers and line widths. This hybrid approach reduces the computational cost and enabled us to quickly screen the full parameter space, thus increasing the likelihood that the unconstrained optimization converges to a global minimum.

2.6. Exhaustive Search of Parameter Combinations. We discarded any potentials and gradients and restricted our optimizations to include only electric fields. Three electric field parameters per atom gives a total of nine electrostatic parameters for the molecule (three atoms of interest, see Figure 1). Out of the nine parameters, we selected a subset of six parameters. The reduced parameter dimensionality avoids overfitting the experimental data. Overfitting can result in an unwanted dependence of the frequency upon the spatial configuration of the solvent atoms, which is analogous to fitting noisy data to high-order polynomials. Instead, the map must produce frequencies that vary smoothly with respect to the solvent coordinates. Figure 9 shows that the six-parameter map produces a smooth spatial dependence of the frequency even with the higher number of parameters (see section 3.3).

Our goal is to find parameter combinations that best reproduce the experimental data. The total number of six-parameter combinations is 84 (9 choose 6). We performed an exhaustive search of all combinations using the fitting procedure outlined above. A full GA search plus direct minimization routine was performed for each combination. We only took into account combinations excluding electrostatic potentials since these parameters are interdependent. For instance, the difference between the electrostatic potentials at the C and O_T atoms is highly correlated with the field along the C═O bond direction since it is similar to computing a numerical derivative of the field along the bond coordinate. To assess the interdependence of the parameters, we computed the Pearson cross-correlation values for the electrostatic parameters using values sampled from the MD trajectories (Figure 3). The figure shows that the potentials at the three atomic positions are well-correlated, with coefficients of approximately 0.6−0.8. Given charge-neutrality considerations, potential coefficients must sum to zero and therefore must be included in pairs. Since

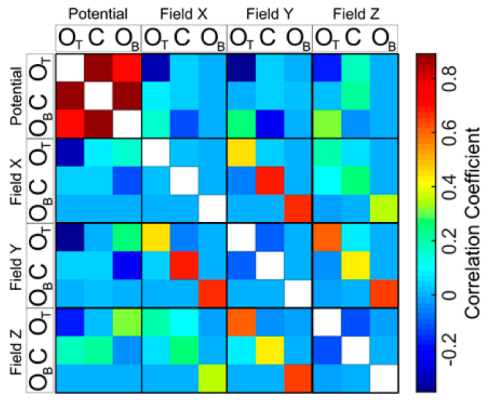


Figure 3. Correlation coefficient matrix between electrostatic parameters for different atoms in ethyl acetate. These parameters are computed using the electrostatic environments sampled from the MD simulations described above. The coefficient values are averaged over all eight solvents.

DOI: 10.1021/acs.jpca.6b02887
J. Phys. Chem. A XXXX, XXX, XXX−XXX

the correlation between these terms is high, including two potentials on different atoms with opposite signs is equivalent to including one additional field parameter. We thus excluded potentials from the map. Correlations between electric field components range from approximately −0.4 to +0.4, indicating that the parameters are sufficiently independent.

The best-fit parameter combination from the exhaustive search was found to be

$$\omega = \omega_0 + C^F_{O_T,x}\frac{\partial \Phi_{O_T}}{\partial x} + C^F_{O_T,y}\frac{\partial \Phi_{O_T}}{\partial y} + C^F_{C,x}\frac{\partial \Phi_C}{\partial x} + C^F_{C,y}\frac{\partial \Phi_C}{\partial y} + C^F_{C,z}\frac{\partial \Phi_C}{\partial z} + C^F_{O_B,y}\frac{\partial \Phi_{O_B}}{\partial y} \quad (4)$$

where subscripts and superscripts indicate the atomic position at which the potential or field is computed and the projection coordinates, respectively. The center frequency, ω_0, is also included as a parameter in the fitting routine. The best-fit values are given in Table 2.

Table 2. Optimized Electrostatic Parameters for Ethyl Acetate[a]

ω_0	$C^F_{O_T,x}$	$C^F_{O_T,y}$	$C^F_{C,x}$	$C^F_{C,y}$	$C^F_{C,z}$	$C^F_{O_B,y}$
1745.0	1154.6	−1964.2	1967.6	−640.4	−835.4	−2776.0

[a]The center frequency, ω_0, is in units of cm^{-1} and field in $cm^{-1}/(E_h/A_0e)$, where E_h denotes hartrees, e is the elementary charge, and A_0 is the Bohr radius.

3. RESULTS AND DISCUSSION

Equation 4 shows the final set of coefficients from the optimized map. Out of the six electric field coefficients, five lie in the plane of the molecule. One component ($C^F_{C,z}$) lies perpendicular to the plane. The five in-plane components reflect the fact that the transition dipole moment also lies in the plane of the molecule at approximately 11.5° with respect to the C═O bond axis.[53] We observe that five of the parameters involve fields at the C and O_T sites, reflecting the fact that the frequency is most sensitive to the electrostatics near the π-conjugated C═O bond and less sensitive to parameters near O_B. Since the molecule is symmetric, we expect the component perpendicular to the plane of the molecule to be zero. A charge above or below the plane of the molecule should induce the same shift, but electric fields are inherently directional. While this term is smaller than four of the six parameters, it is not zero, which could be due to an artifact in the fit. To account for symmetric shifts induced by charges above or below the plane of the molecule, electric field squared parameters should be included. However, this further increases the combination of possible parameters and may introduce additional artifacts. Finally, it is important to note that ω_0 the "zero-field" frequency is 1745 cm^{-1} (Table 2), approximately 5 cm^{-1} lower than the frequency of ethyl acetate in *n*-hexane a nonpolar solvent and lower than the 1764 cm^{-1} frequency in the gas phase. This shift is likely due to van der Waals and polarizability effects that are not captured by the electrostatic map. Thus, the "zero-field" frequency should not be considered a vacuum frequency.

3.1. Electrostatic Map Performance: Comparison of Simulations with Experiments. Figure 4 shows experimental and simulated spectra of ethyl acetate in different solvents. Spectra in alcohols and D_2O exhibit multiple peaks. In alcohols, the individual peaks, from high to low frequency, correspond to

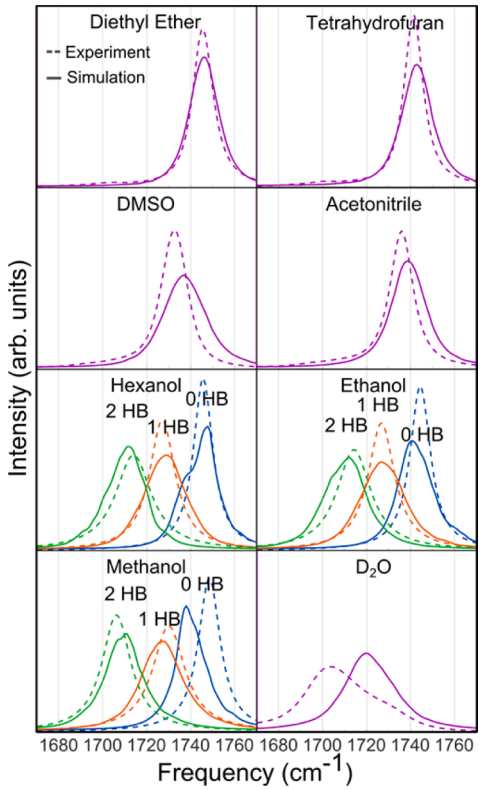


Figure 4. Experimental (dashed) and simulated (solid) linear absorption spectra of ethyl acetate carbonyl stretches in the eight reference solvents. Hydrogen-bonding solvents were split into 0, 1, and 2 hydrogen bond ensembles. All spectra are normalized to their area.

ensembles with 0, 1, and 2 HB between the O_T and solvent O−H. In D_2O, the spectrum is fit directly to experiment due to uncertainties in the Lorentzian fit parameters (see Table S2).[47] Alcohol spectra are fit independently by splitting the MD trajectories into multiple subtrajectories, each corresponding to a different HB ensemble. This approach sidesteps an important shortcoming of the GAFF force field: its failure to capture the underlying thermodynamics of hydrogen bonding (see section S1 in the Supporting Information).

Overall agreement between experiment and simulations is very good, particularly given the simplicity and computational efficiency of the electrostatic models. The models qualitatively capture the peak frequencies in all eight solvents as well as the effects of hydrogen bonding in alcohols. Spectra in low-polarity solvents such as diethyl ether are particularly well reproduced by the map.

Experimental and computed spectra in low-polarity solvents are in excellent agreement, but a blue-shift of approximately 3−5 cm^{-1} is observed in the two polar aprotic solvents, DMSO and acetonitrile. The comparison suggests that the GAFF force fields may underestimate the atomic charges for these solvents, which are characterized by multiple bonds between polar atoms (e.g., S═O in DMSO or C≡N in acetonitrile). Such double bonds are not present in ethyl ether or tetrahydrofuran. These effects could be further explored with ab initio electronic structure calculations to test how the electrostatic potential computed from GAFF charges differs from the full QM potential for polar solvent molecules containing double or triple bonds.

Experimental and simulated spectra for the three alcohols are in good agreement. The experimental shift of approximately 17

DOI: 10.1021/acs.jpca.6b02887
J. Phys. Chem. A XXXX, XXX, XXX−XXX

cm^{-1} per HB is semiquantitatively reproduced by our model. We observe that agreement is best for the 1 and 2 HB cases. This can be attributed to the fact that electric fields dominate in the 1 and 2 HB cases, whereas van der Waals interactions—which are not captured by the model—can contribute significantly to line broadening in the 0 HB case. In D_2O, the simulated spectrum remains significantly blue-shifted with respect to experiment. This could be due to strong hydrogen-bonding effects that are not captured by the simulations. It is important to note that simulations were carried out in H_2O whereas experimental spectra were measured in D_2O since the C═O stretch spectrally overlaps with water's O−H bend. We performed simulations in D_2O and found that both spectra and hydrogen bond populations show virtually no difference compared to H_2O, confirming that the slight change in mass has negligible effects on the fast dynamics of water. Finally, part of the line narrowing observed in the simulations can be attributed to a shorter vibrational lifetime compared to nonpolar solvents.[47] In our simulation, the lifetime is fixed at 1.5 ps for all solvents. Thus, solvent-dependent lifetime effects are not captured.

Figure 5 shows a comparison between simulated and experimental absorption spectra for the three alcohols. Spectra

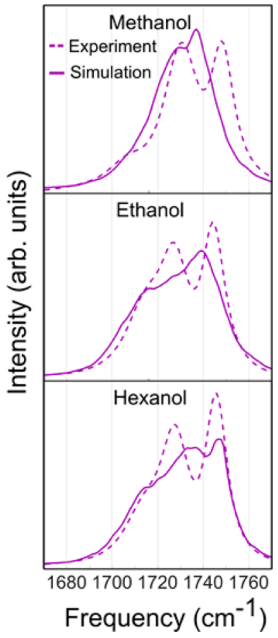


Figure 5. Experimental (dashed) and simulated (solid) linear absorption spectra of ethyl acetate in the three different alcohols. To make a direct comparison with experiment, simulated spectra for different hydrogen bond ensembles (Figure 4) are weighted by the hydrogen bond populations extracted from experiment (Table S1).

are computed by weighting each individual HB ensemble spectrum (Figure 4) by its experimentally determined population (Table S1). Qualitatively, the three peaks in the spectrum are reproduced by the simulation. However, the three HB peaks are not as clearly defined as they are in the experiment. This is a result of the simulated line shapes being 3−6 cm^{-1} (full width at half-maximum) broader than their experimental counterparts (see Table S3).

To assess the quality of the map and test its transferability, Figure 6 shows a comparison of spectra in solvents that were not part of the initial map development set: butanol,

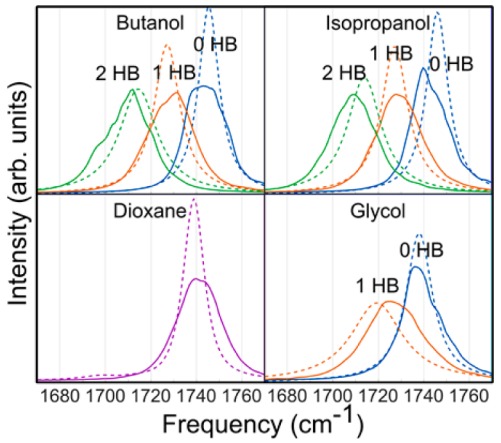


Figure 6. Experimental (dashed) and simulated (solid) linear absorption spectra of ethyl acetate in four solvents that were not part of the map development set. The performance of the map in these cases is comparable to its performance for members of map development set.

isopropanol, dioxane, and diethylene glycol. Simulated spectra are in good agreement with experimental spectra in these solvents, which is comparable to the agreement between simulation and experiment for the eight solvents in the initial map development set (Figure 4). The relative frequencies of each hydrogen-bond population are well reproduced for the alcohols. The broadness of the 1 HB peak in diethylene glycol, which is not observed in the other alcohols, is also well reproduced by the map. Spectra in dioxane are also in good agreement, indicating that the map is transferable between polar solvents and should be able to qualitatively reproduce spectra in most polar protic and aprotic solvents.

Figure 7 shows a comparison of the center frequencies for the experimental and calculated spectra used in map development and assessment. The plot shows an excellent correlation

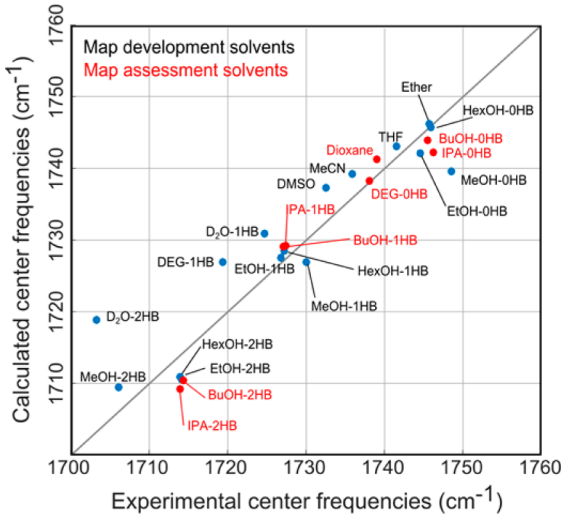


Figure 7. Comparison between computed and experimental C═O center frequencies and line widths in all solvents. Solvents used to parametrize the map are indicated in blue and solvents used to evaluate the map performance are shown in red. The solvents are labeled Ether = diethyl ether, THF = tetrahydrofuran, MeCN = acetonitrile, HexOH = hexanol, EtOH = ethanol, MeOH = methanol, BuOH = butanol, IPA = isopropanol, and DEG = diethylene glycol.

DOI: 10.1021/acs.jpca.6b02887
J. Phys. Chem. A XXXX, XXX, XXX−XXX

between calculated and measured line centers for solvents that were used to parametrize the map as well as separate solvents used to assess the quality of the map. The root-mean-squared deviations between calculated and experimental center frequencies is 4.9 cm^{-1}, suggesting this is the approximate error that should be expected when using this map. It is worth noting that this error is approximately a third of the ~17 cm^{-1} per hydrogen bond shift that we observe for the alcohols (Table S2). In addition to enabling interpretation of experimental spectra in terms of hydrogen-bonding environments, we expect the map to distinguish frequency shifts induced by changes in environment polarity.

3.2. Frequency−Frequency Correlation Functions (FFCFs). Figure 8 shows the frequency−frequency time correlation functions (FFCFs) plotted for individual HB populations. A fast (<200) fs initial decay and a slower relaxation component are observed. Interestingly, the frequency−frequency correlation functions appear to change only slightly with respect to the number of hydrogen bonds for the three alcohols and D_2O, indicating that the time scales of fast fluctuations are comparable across HB ensembles.

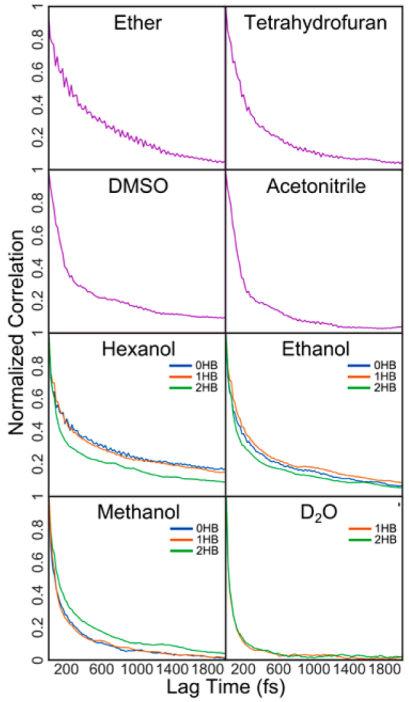


Figure 8. Frequency−frequency correlation functions of the C═O stretch in ethyl acetate generated using the electrostatic map described in the text.

3.3. Frequency Shifts and Electrostatic Environments. Figure 9 shows frequency shifts from the zero-field frequency with respect to the spatial position of a single positive +0.5 au test charge scanned across the plane of the molecule (X,Y plane). Shifts are computed using the electrostatic parameters shown in Table 2. A +0.5 au test charge was selected to represent the typical polar hydrogen atom in current MD force fields. Along each contour, the frequency shift is uniform with respect to the position of the charge in the plane. The blue contours indicate a shift to lower frequency whereas red contours are higher frequencies. As expected, a positive H atom near O_T from a H-bonded solvent molecule shifts the oscillator

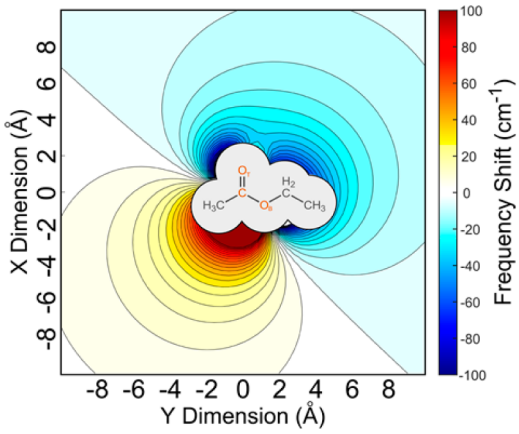


Figure 9. Frequency shifts resulting from a single +0.5 au positive test charge scanned across the (x, y) plane of the molecule. The circles indicate the van der Waals radii of the ethyl acetate atoms.

by as much as −100 cm^{-1}. It is important to consider that the oxygen atom in O−H will compensate by contributing a commensurate blue-shift, thus attenuating the overall shift to approximately 17 cm^{-1} per HB. The sum of contributions from all atoms gives rise to the final frequency shift extracted from the map.

While the actual frequency shift depends on the 3-dimensional coordinates of the atoms, this figure—plotted only for charges in the plane of the molecule—provides some insight into the behavior of the map and illustrates how the map captures frequency shifts induced by polar solvent atoms. Positive and negative regions are separated by single nodal line approximately aligned with the C−O_B bond; overall, this representation of the map qualitatively resembles the contour lines of a point dipole. The nodal line angle could be reflective of the fact that the C═O transition dipole lies at approximately 11.5° with respect to the C═O bond in the plane of the molecule as reported for methyl acetate, a similar ester molecule.[53] This frequency-shift distribution is very similar to those seen in amide units using the maps of Skinner and Zanni.[42,54] It is worth noting that the plot shows a smooth spatial dependence of the shift along these two coordinates and only displays a single node. This suggests that the map does not overfit the data, as such a case can introduce additional nodes in the two-dimensional plot.

Figure 10 shows the average solvent charge distribution in the ester plane. The plots are equivalent to 2D charge histograms generated using the force-field charges and averaged over the 10 ns trajectories. The charge distribution in nonpolar solvents is relatively uniform and the distribution is similar across non-hydrogen-bonding solvents. The plot shows preference for positive charges surrounding the negative oxygen atom in nonpolar solvents. In DMSO and acetonitrile, the plot shows preferential orientation of negative charges near the CH_3 groups, indicating weak electrostatic interactions that lead to preferential solvent orientation at this position that are not as pronounced in diethyl ether and tetrahydrofuran.

In the case of alcohols and D_2O, a strong positive charge is observed near the oxygen atom as a result of hydrogen bonding. The broader negative area behind the first positive shell corresponds to the O−H oxygen atom, which exhibits a larger range of motion than the H atom. D_2O, and to a small extent methanol, also hydrogen bond to the bridging oxygen atom, O_B. The absence of hydrogen bonding at O_B in ethanol

DOI: 10.1021/acs.jpca.6b02887
J. Phys. Chem. A XXXX, XXX, XXX−XXX

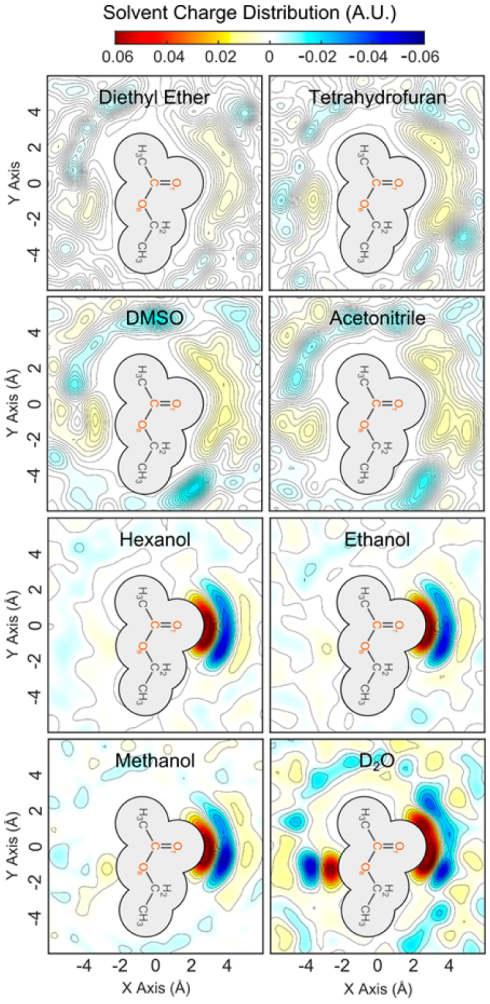


Figure 10. Spatial maps indicating the average solvent atomic charges surrounding ethyl acetate molecule in our molecular dynamics simulations. Blue contours represent areas of negative charge, and red contours represent positive charge.

and hexanol can be attributed to steric effects since the GAFF force-field charges on the O−H hydrogen atoms are virtually identical for D_2O and the alcohol series, suggesting that the electrostatic interactions should be similar. Analysis of the data did not reveal any strong correlations between carbonyl frequency and hydrogen bonding at this position, implying that the C═O stretching frequency is insensitive to the electrostatic environment near O_B.

3.4. Effect of Hydrogen Bond Switching on Dynamic Line Shapes. Since individual hydrogen bond ensembles are fit independently, the simulated line shapes could be artificially narrowed due to exclusion of the large frequency jumps that result from HB switching. To evaluate these effects, we compared the line shapes calculated from full MD trajectories to the sums of individual ensemble spectra weighted by their respective HB populations. Figure 11 shows the comparison for all four hydrogen-bonding solvents. The spectra are indistinguishable in all three alcohols. We observe $a < 1$ cm^{-1} line narrowing in D_2O. The analysis thus demonstrates the validity of this approach as a means to overcome sampling issues in molecular dynamics simulations. A similar method was used by Cho and co-workers to model HB exchange dynamics and 2D IR spectra of *N*-methylacetamide in methanol.[55,56] An

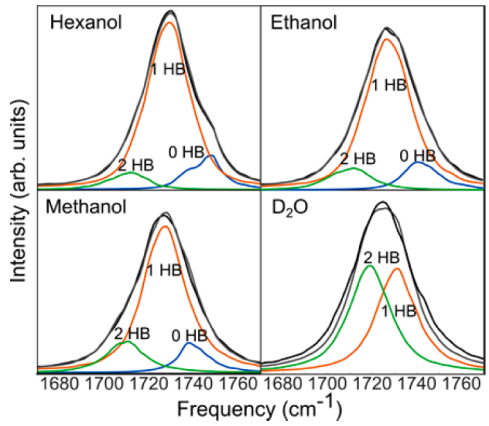


Figure 11. Simulated absorption spectra of ethyl acetate in four hydrogen-bonding solvents. The light lines show population-weighted spectra of individual hydrogen-bonding populations, and the dark curves represent the sum of the individual ensemble spectra. The black lines represent spectra calculated from the full MD trajectories.

important challenge with this approach is related to MD sampling: long MD trajectories must be carried out in order to ensure proper sampling. In addition, it is unlikely that simulations capture the correct HB switching kinetics (further discussion is provided in the Supporting Information).

4. CONCLUSIONS AND OUTLOOK

We parametrized an ester carbonyl electrostatic map using IR absorption spectra of ethyl acetate in different solvents to select and optimize parameter combinations. The GAFF parametrization was used for compatibility with the widely used Amber force field.[49] Spectra simulated with the map are in excellent agreement with experimental absorption spectra. The general procedure developed here can be easily applied to a variety of carbonyls and other vibrational probes.

Electrostatic maps are particularly useful for interpreting the IR spectra of amide groups in proteins. We expect that this new ester map will further extend the IR toolbox to carbonyl groups in lipids and similar biological molecules. Ultrafast methods, such as two-dimensional IR spectroscopy, can provide additional experimental inputs including vibrational lifetimes and hydrogen bond exchange rates. Such data will help further refine the electrostatic maps and extend their utility for exploring the dynamics of carbonyl molecules in diverse molecular environments.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.6b02887.

Hydrogen bond analysis of molecular dynamics trajectories and hydrogen bond populations: hydrogen bond populations extracted from molecular dynamics trajectories (Table S1); analysis of experimental and computed spectra; Lorentzian fit parameters extracted from experimental IR absorption spectra (Table S2); Lorentzian fit parameters extracted from computed IR absorption spectra (Table S3) (PDF)

DOI: 10.1021/acs.jpca.6b02887
J. Phys. Chem. A XXXX, XXX, XXX−XXX

AUTHOR INFORMATION

Corresponding Author

*Phone 512-471-6488; e-mail cbaiz@cm.utexas.edu (C.R.B.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge the College of Natural Sciences at the University of Texas at Austin for generous support through a startup grant. C.B. acknowledges the Welch Foundation (Grant F-1891). Molecular dynamics simulations were performed at the Texas Advanced Computing Center (TACC) at the University of Texas at Austin. We thank Thomas la Cour Jansen (University of Groningen) and Dimitrii Makarov (University of Texas at Austin) for the insightful comments on the manuscript.

REFERENCES

(1) Reichardt, C. Solvatochromic Dyes as Solvent Polarity Indicators. *Chem. Rev.* **1994**, *94* (8), 2319−2358.

(2) Barth, A. Infrared Spectroscopy of Proteins. *Biochim. Biophys. Acta, Bioenerg.* **2007**, *1767* (9), 1073−1101.

(3) Bandekar, J. Amide Modes and Protein Conformation. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1992**, *1120* (2), 123−143.

(4) Ganim, Z.; Hoi, S. C.; 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.

(5) Baiz, C. R.; Tokmakoff, A. Structural Disorder of Folded Proteins: Isotope-Edited 2D IR Spectroscopy and Markov State Modeling. *Biophys. J.* **2015**, *108* (7), 1747−1757.

(6) Elliott, A.; Ambrose, E. J. Structure of Synthetic Polypeptides. *Nature* **1950**, *165* (4206), 921−922.

(7) Miyazawa, T.; Shimanouchi, T.; Mizushima, S. I. Characteristic Infrared Bands of Monosubstituted Amides. *J. Chem. Phys.* **1956**, *24* (2), 408−418.

(8) Byler, D. M.; Susi, H. Examination of the Secondary Structure of Proteins by Deconvolved FTIR Spectra. *Biopolymers* **1986**, *25* (3), 469−487.

(9) Miyazawa, T. Perturbation Treatment of the Characteristic Vibrations of Polypeptide Chains in Various Configurations. *J. Chem. Phys.* **1960**, *32* (6), 1647−1652.

(10) Krimm, S.; Bandekar, J. Vibrational Spectroscopy and Conformation of Peptides, Polypeptides, and Proteins. *Adv. Protein Chem.* **1986**, *38*, 181−364.

(11) Moore, W. H.; Krimm, S. Transition Dipole Coupling in Amide I Modes of Beta Polypeptides. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72* (12), 4933−4935.

(12) Krimm, S.; Abe, Y. Intermolecular Interaction Effects in Amide I Vibrations of Beta Polypeptides. *Proc. Natl. Acad. Sci. U. S. A.* **1972**, *69* (10), 2788−2792.

(13) Cai, K.; Du, F.; Zheng, X.; Liu, J.; Zheng, R.; Zhao, J.; Wang, J. General Applicable Frequency Map for the Amide-I Mode in beta-Peptides. *J. Phys. Chem. B* **2016**, *120* (6), 1069−79.

(14) Malolepsza, E.; Straub, J. E. Empirical Maps for the Calculation of Amide I Vibrational Spectra of Proteins from Classical Molecular Dynamics Simulations. *J. Phys. Chem. B* **2014**, *118* (28), 7848−55.

(15) Reppert, M.; Tokmakoff, A. Electrostatic Frequency Shifts in Amide I Vibrational Spectra: Direct Parameterization Against Experiment. *J. Chem. Phys.* **2013**, *138* (13), 134116.

(16) Karjalainen, E. L.; Ravi, H. K.; Barth, A. Simulation of the Amide I Absorption of Stacked beta-Sheets. *J. Phys. Chem. B* **2011**, *115* (4), 749−757.

(17) Maekawa, H.; Ge, N. H. Comparative Study of Electrostatic Models for the Amide-I and -II Modes: Linear and Two-Dimensional Infrared Spectra. *J. Phys. Chem. B* **2010**, *114* (3), 1434−1446.

(18) Lin, Y. S.; Shorb, J. M.; Mukherjee, P.; Zanni, M. T.; Skinner, J. L. Empirical Amide I Vibrational Frequency Map: Application to 2D-IR Line Shapes for Isotope-Edited Membrane Peptide Bundles. *J. Phys. Chem. B* **2009**, *113* (3), 592−602.

(19) Ganim, Z.; Tokmakoff, A. Spectral Signatures of Heterogeneous Protein Ensembles Revealed by MD Simulations of 2DIR Spectra. *Biophys. J.* **2006**, *91* (7), 2636−2646.

(20) Jansen, T. L.; Dijkstra, A. G.; Watson, T. M.; Hirst, J. D.; Knoester, J. Modeling the Amide I Bands of Small Peptides. *J. Chem. Phys.* **2006**, *125* (4), 044312.

(21) Choi, J. H.; Kim, J. S.; Cho, M. H. Amide I Vibrational Circular Dichroism of Polypeptides: Generalized Fragmentation Approximation Method. *J. Chem. Phys.* **2005**, *122* (17), 174903.

(22) Lee, C.; Cho, M. H.; Local Amide, I. Mode Frequencies and Coupling Constants in Multiple-stranded Antiparallel beta-sheet Polypeptides. *J. Phys. Chem. B* **2004**, *108* (52), 20397−20407.

(23) Bour, P.; Keiderling, T. A. Empirical Modeling of the Peptide Amide I Band IR Intensity in Water Solution. *J. Chem. Phys.* **2003**, *119* (21), 11253−11262.

(24) Hamm, P.; Woutersen, S. Coupling of the Amide I Modes of the Glycine Dipeptide. *Bull. Chem. Soc. Jpn.* **2002**, *75* (5), 985−988.

(25) 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* (1), 81−86.

(26) Torii, H.; Tasumi, M. Model-Calculations on the Amide-I Infrared Bands of Globular-Proteins. *J. Chem. Phys.* **1992**, *96* (5), 3379−3387.

(27) Barth, A.; Zscherp, C. What Vibrations Tell us About Proteins. *Q. Rev. Biophys.* **2002**, *35* (4), 369−430.

(28) Jackson, M.; Mantsch, H. H. The Use and Misuse of FTIR Spectroscopy in the Determination of Protein-Structure. *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30* (2), 95−120.

(29) Torii, H.; Tasumi, M. 3-Dimensional Doorway-State Theory for Analyses of Absorption-Bands of Many-Oscillator Systems. *J. Chem. Phys.* **1992**, *97* (1), 86−91.

(30) Blume, A.; Hubner, W.; Messner, G. Fourier Transform Infrared Spectroscopy of 13C═O-labeled Phospholipids Hydrogen Bonding to Carbonyl Groups. *Biochemistry* **1988**, *27* (21), 8239−49.

(31) Cameron, D. G.; Mantsch, H. H. Metastability and Polymorphism in the Gel Phase of 1,2-dipalmitoyl-3-SN-phosphatidylcholine. A Fourier Transform Infrared Study of the Subtransition. *Biophys. J.* **1982**, *38* (2), 175−84.

(32) Arsov, Z.; Quaroni, L. Direct Interaction Between Cholesterol and Phosphatidylcholines in Hydrated Membranes Revealed by ATR-FTIR Spectroscopy. *Chem. Phys. Lipids* **2007**, *150* (1), 35−48.

(33) Ueda, I.; Chiou, J.-S.; Krishna, P. R.; Kamaya, H. Local Anesthetics Destabilize Lipid Membranes by Breaking Hydration Shell: Infrared and Calorimetry Studies. *Biochim. Biophys. Acta, Biomembr.* **1994**, *1190* (2), 421−429.

(34) Lewis, R. N.; McElhaney, R. N. Structures of the Subgel Phases of n-saturated diacyl phosphatidylcholine Bilayers: FTIR Spectroscopic Studies of 13C═O and 2H Labeled Lipids. *Biophys. J.* **1992**, *61* (1), 63−77.

(35) Hübner, W.; Blume, A. Interactions at the Lipid−Water Interface. *Chem. Phys. Lipids* **1998**, *96* (1−2), 99−123.

(36) Kaiser, H. J.; Lingwood, D.; Levental, I.; Sampaio, J. L.; Kalvodova, L.; Rajendran, L.; Simons, K. Order of Lipid Phases in Model and Plasma Membranes. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (39), 16645−50.

(37) Engelman, D. M. Membranes are more mosaic than fluid. *Nature* **2005**, *438* (7068), 578−80.

(38) Jacobson, K.; Mouritsen, O. G.; Anderson, R. G. Lipid Rafts: At a Crossroad Between Cell Biology and Physics. *Nat. Cell Biol.* **2007**, *9* (1), 7−14.

(39) Vigh, L.; Escriba, P. V.; Sonnleitner, A.; Sonnleitner, M.; Piotto, S.; Maresca, B.; Horvath, I.; Harwood, J. L. The Significance of Lipid Composition for Membrane Activity: New Concepts and Ways of Assessing Function. *Prog. Lipid Res.* **2005**, *44* (5), 303−44.

DOI: 10.1021/acs.jpca.6b02887
J. Phys. Chem. A XXXX, XXX, XXX−XXX

(40) Phillips, R.; Ursell, T.; Wiggins, P.; Sens, P. Emerging Roles for Lipids in Shaping Membrane-protein Function. *Nature* **2009**, *459* (7245), 379−85.
(41) MacCallum, J. L.; Tieleman, D. P. Hydrophobicity Scales: A Thermodynamic Looking Glass into Lipid-Protein Interactions. *Trends Biochem. Sci.* **2011**, *36* (12), 653−62.
(42) 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.
(43) Mukherjee, P.; Kass, I.; Arkin, I. T.; Zanni, M. T. Structural Disorder of the CD3 xi Transmembrane Domain Studied with 2D IR Spectroscopy and Molecular Dynamics Simulations. *J. Phys. Chem. B* **2006**, *110* (48), 24740−24749.
(44) Lessing, J.; Roy, S.; Reppert, M.; Baer, M.; Marx, D.; Jansen, T. L.; 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* (11), 5032−5035.
(45) Smith, A. W.; Lessing, J.; Ganim, Z.; Peng, C. S.; Tokmakoff, A.; Roy, S.; Jansen, T. L. C.; Knoester, J. Melting of beta-hairpin peptide using isotope-edited 2D IR spectroscopy and simulations. *J. Phys. Chem. B* **2010**, *114*, 10913−10924.
(46) De Marco, L.; Thamer, M.; Reppert, M.; Tokmakoff, A. Direct Observation of Intermolecular Interactions Mediated by Hydrogen Bonding. *J. Chem. Phys.* **2014**, *141* (3), 034502.
(47) Banno, M.; Ohta, K.; Tominaga, K. Ultrafast Vibrational Dynamics and Solvation Complexes of Methyl acetate in Methanol Studied by Sub-picosecond Infrared Spectroscopy. *J. Raman Spectrosc.* **2008**, *39* (11), 1531−1537.
(48) Wang, J. M.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graphics Modell.* **2006**, *25* (2), 247−260.
(49) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157−1174.
(50) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular-Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684−3690.
(51) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations Through Multi-level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.
(52) Jansen, T. L.; Knoester, J.; Transferable, A. Electrostatic map for Solvation Effects on Amide I Vibrations and its Application to Linear and Two-dimensional Spectroscopy. *J. Chem. Phys.* **2006**, *124* (4), 044502.
(53) Fang, B.; Wang, T.; Chen, X.; Jin, T.; Zhang, R.; Zhuang, W. Modeling Vibrational Spectra of Ester Carbonyl Stretch in Water and DMSO Based on Molecular Dynamics Simulation. *J. Phys. Chem. B* **2015**, *119*, 12390.
(54) Fayer, M. D. *Ultrafast Infrared Vibrational Spectroscopy*; Taylor & Francis: Boca Raton, FL, 2013; p *xii*, 475 pp.
(55) Kwac, K.; Lee, H.; Cho, M. H. Non-Gaussian Statistics of Amide I Mode Frequency Fluctuation of N-methylacetamide in Methanol Solution: Linear and Nonlinear Vibrational Spectra. *J. Chem. Phys.* **2004**, *120* (3), 1477−1490.
(56) Woutersen, S.; Mu, Y.; Stock, G.; Hamm, P. Hydrogen-bond Lifetime Measured by Time-resolved 2D-IR Spectroscopy: N-methylacetamide in Methanol. *Chem. Phys.* **2001**, *266* (2−3), 137−147.

DOI: 10.1021/acs.jpca.6b02887
J. Phys. Chem. A XXXX, XXX, XXX−XXX