The Journal of Chemical Physics

AIP Publishing

Hydrogen Bonding Blues: Vibrational Spectroscopy of the TIP3P Water Model

Zeke A. Piskulich[1, a)] and Qiang Cui[2, b)]
[1)] *Rutgers University, Institute for Quantitative Biomedicine, 174 Frelinghuysen Road, Piscataway, NJ, 08854*
[2)] *Department of Chemistry, Boston University, 590 Commonwealth Avenue, Boston, MA, 02215, USA*

(Dated: 5 November 2024)

The computational spectroscopy of water has proven to be a powerful tool for probing the structure and dynamics of chemical systems and for providing atomistic insight into experimental vibrational spectroscopic results. However, such calculations have been limited for biochemical systems due to the lack of empirical vibrational frequency maps for the TIP3P water model, which is used in many popular biomolecular force fields. Here we develop an empirical map for the TIP3P model and evaluate its efficacy for reproducing the experimental vibrational spectroscopy of water. We observe that the calculated Infrared and Raman spectra are blue-shifted and narrowed compared to the experimental spectra. Further analysis finds that the blue-shift originates from a shifted distribution of frequencies, rather than other dynamical effects, suggesting that the TIP3P model forms a significantly different electrostatic environment than other three-point water models. This is explored further by examining the two-dimensional infrared spectra, which demonstrates that the blue-shift is significant for the first two vibrational transitions. Likewise, spectral diffusion timescales, evaluated through both the center-lines slope and the frequency-frequency time correlation function demonstrate that TIP3P exhibits significantly faster spectral dynamics than other three-point models. Finally, sum-frequency generation spectroscopy calculations suggest that despite these challenges, the TIP3P empirical map can provide phenomenological, qualitative, insight into the behavior of water at the air-water and lipid-water interfaces. As these interfaces are models for hydrophobic and hydrophilic environments observed in biochemical systems, the presently developed empirical map will be useful for future studies of biochemical systems.

I. INTRODUCTION

Vibrational spectroscopy is a powerful tool for understanding the structure and dynamics of biochemical systems.[1–5] Linear techniques, like infrared (IR) and Raman spectroscopy, have often been used for determinations of biochemical structures by signature contributions from specific functional groups. For example, amide vibrational modes are often used to provide insight into the secondary structure of proteins due to the appearance of vibrational bands associated with specific protein conformations. Alternatively, many spectroscopic studies have focused on water molecules in the interior or near the surface of biomolecules.[2,6–8] On the one hand, these efforts are motivated by the consideration that water molecules make crucial contributions to the structural, energetic and dynamic properties of biomolecules;[9–11] for example, water molecules in protein cavities play major roles in enzyme catalysis, ion transport and ligand binding.[12–16] On the other hand, the structural and dynamical properties of water molecules at biochemical interfaces are influenced by local surface topology and interactions with surface-exposed atoms.[17–19] Therefore, such systems are well-suited for spectroscopic characterization.[20–24]

Measurements of water molecules in biological systems are challenged by the presence of significant populations of water in the bulk region and often require the use of model systems, like reverse micelles,[25,26] to remove these "background" contributions. Along this line, non-linear optical techniques, like Sum-Frequency Generation (SFG) spectroscopy and two-dimensional Infrared (2DIR) spectroscopy, which can explore biochemical systems *in situ* without the use of model systems, have grown in popularity over recent years and have successfully provided significant structural, dynamical and mechanistic insights into the role of water in biochemical environments.[1,27–31]

To help interpret these linear and nonlinear spectra at the molecular level, it is essential to compute these vibrational spectra based on molecular simulations. From classical molecular dynamics simulations, the vibrational frequency mapping (empirical mapping) approach developed by Corcelli, Skinner, and others, allows for both linear and non-linear vibrational spectra to be obtained through the calculation of time-correlation functions (TCFs).[32–38] These TCFs, as will be described in greater detail below, depend on quantum mechanical observables (*e.g.*, the vibrational transition frequency, transition dipole moment, transition polarizability) that are not available in a purely classical simulation. In this approach, a small set of classical configurations are used to generate a map between quantum mechanical observables calculated from quantum mechanics/molecular

[a)] piskulichz@gmail.com; Corresponding Author
[b)] qiangcui@bu.edu; Corresponding Author

mechanics (QM/MM) calculations and the local electrostatic environment described by classical point charges. This technique has proven successful for a range of water force-fields, including SPC/E,[34] TIP4P,[36,39] and E3B3,[36] as well as other modes of interest to biochemical spectroscopies such as the amide I stretch in proteins,[37,38] the carbonyl stretch in nucleobases,[40] and many others.[41]

While many empirical mapping studies have focused on biochemical systems, relatively few have looked at the vibrational spectroscopy of water in biochemical environments due to the expense involved in calculating the electric field for thousands of water molecules. Roy and co-workers simulated the SFG spectra of SPC/E water with POPC and POPE lipids described by the Berger force-field[42] and were able to provide atomistic insight into how varied water orientations around the lipid head groups resulted in differing SFG intensities.[23,24] Often, however, biochemical simulations do not use the SPC/E, TIP4P, and E3B3 water models for which the accuracy has been well-established. In part, this is because the CHARMM36 force-field, one of the most commonly used force-fields for biochemical simulation, has been parameterized with the TIP3P water model.

While CHARMM36 provides generally good agreement between experimental and simulated structures of biochemical systems, the same cannot be said for the ability of the TIP3P water model to accurately describe the structures and dynamics of real water.[43,44] The TIP3P radial distribution function shows less structuring past the second solvation shell than is observed experimentally. The TIP3P water model also exhibits diffusion, reorientation, and hydrogen-bond exchange dynamics that are nearly a factor of three times too fast.[45] While there have been a number of water models, such as the TIP3P-PME model, that have aimed to fix these issues, they have not been widely adopted in biomolecular simulations as they could negatively impact the force-field's ability to accurately describe the properties of biomolecules.[46] More sophisticated approaches, such as the charge-response kernel model of water,[47–49] have been used to model the spectroscopy of configurations generated using TIP3P. However, to the best of our knowledge, there has not yet been an attempt to use the empirical mapping approach to describe the spectroscopy of TIP3P.

In the present work, we have developed an empirical vibrational frequency map for the TIP3P water model for use in biochemical simulations. We first present the new empirical map and thoroughly review the mapping approach, then we discuss its ability to describe dynamics and structures within bulk water, and compare results with those from an improved empirical map developed for another three-point model for water, the SPC/E water model. Such a comparison helps establish the sensitivity of specific spectral features to the description of certain populations of water structures by different force field models.

A key challenge for the spectroscopy of biochemical systems is the multitude of environments that water can find itself in, both hydrophilic (as is the case for water molecules near lipid membrane head-groups) and hydrophobic (water molecules in certain protein cavities, water molecules deep in lipid membranes). Thus, it is important to understand how the TIP3P model performs in such environments. In the present work, we use the air-water interface as an example of a hydrophobic environment and the lipid-water interface as an example of a hydrophilic environment. Using SFG spectroscopy, we then demonstrate that the model can provide qualitative information about the nature of these interfacial regions. Therefore, we expect that the frequency map developed for the TIP3P model, when used with care, can provide valuable insights into the structural and dynamical properties of water in biological systems.

II. COMPUTATIONAL METHODS

A. Frequency Mapping Simulations

A simulation cell of 855 TIP3P water molecules was generated using the CHARMM-GUI Solution builder,[50,51] and was resized to match an experimental density of 0.9991 grams/cubic centimeter using `gmx editconf`. The Gromacs Simulation Package was used for all simulations in the present work.[52] Simulations were equilibrated in the canonical (NVT) ensemble using a Nosé-Hoover thermostat with a coupling time of 1.0 ps.[53,54] A 2.0 fs simulation timestep was used for all simulations. The equilibration for the water box happened over the duration of a 2.0 ns simulation. Following equilibration, a production run was conducted for 10 ns, from which configurations were output every 2.0 ps. This resulted in a total of 5,000 configurations used to build the empirical map. The procedure for the SPC/E water model simulations were performed in an identical manner, with just the water model changed. The `LINCS` algorithm was used to hold the geometry of the water model rigid.[55]

B. Frequency Mapping Electronic Structure Calculations

For each configuration generated from the aforementioned molecular dynamics simulation, we selected a single water molecule randomly and then carved out a droplet around the hydrogen atom. For each droplet, a set of sixteen electronic structure calculations were run where the OH bond was stretched while maintaining a constant center of mass. These points were evenly sampled between an OH bond distance of 0.65 and 2.25 angstroms. The B3LYP-D3/6-311G(d,p) level of theory was used for all calculations.[56–58]

For each calculation, an inner sphere of 4.0 Å around the central hydrogen was treated quantum mechanically, while the remainder of the droplet was treated as classical

point charges. Here, the distance cutoff was performed with respect to the O$\cdots$ H distance, and molecules straddling the boundary were made whole. In the present work, for the new TIP3P map a slightly longer cutoff (8.31 Å) than the 7.831 Å utilized by Skinner and co-workers was utilized based on the radial distribution function as described below.[34] The new SPC/E map retained the Skinner value.

C. Spectroscopic Calculations

1. HOD/D2O Spectra

A room temperature molecular dynamics simulation of TIP3P water was run in the canonical ensemble (NVT) for 10 nanoseconds using the same parameters described above. From this simulation, configurations were output every 10 picoseconds, resulting in 1,000 total configurations which acted as starting points for a series of short (10 ps) simulations in the microcanonical ensemble (NVE). This multi-trajectory approach allowed for sampling of the canonnical ensemble, while also allowing for the calculation of spectra in a trivially parallelized way.

For each NVE trajectory, the trajectory was written every 4.0 fs, providing a total of 2,500 configurations from which the spectra were calculated. For each configuration in these trajectories, we used our `Electric Field Cell List` package[59] to calculate the electric field projected onto the OH bond vector, using the cutoff value from the empirical map. This program calculates the electric field using a cell-list decomposition approach as has been described elsewhere.[60] While this approach provides relatively minimal advantage for this small water system, it provides significant scaling advantages for the membrane system described below.

Finally, we used our `Generalized Water Spectroscopy` package[61] to calculate the Infrared, Raman, and 2D-IR spectra of the water system using the time-correlation approach described below. For these time correlation calculations, time origins were selected every 100 configurations (400 fs), and the correlation time was 500 fs. For the `Fast Fourier Transform`, the FFTW3 Library was used with a frequency resolution of 4.0 cm^{-1}.

2. Air-Water Interface

A slab system of 2,258 water molecules was initialized into a 6 nm $\times$ 6 nm $\times$ 2 nm box, and the latter dimension was then expanded to 6 nm to produce an interface. The system was equilibrated for 10.0 ns in the NVT ensemble, again using a Nosé-Hoover thermostat with a coupling time of 1.0 ps and a simulation timestep of 2.0 fs. Configurations were output every 10.0 ps, for a total of 1,000 configurations used as initial conditions for spectroscopic calculations which were conducted in the same manner as described above. Sum-frequency generation spectra were then calculated with the center of mass of the air-water system translated such that the water slab spanned the lower periodic boundary.

3. Water-Lipid Interface

A similar approach was utilized for the POPC and POPE membrane systems. Initial configurations for these systems were built using the `CHARMM-GUI Membrane Builder`.[62,63] For the purposes of these calculations, the system was built to solely contain water and lipids, no ions were added. For each simulation, 200 lipids were added to the box. The POPC and POPE simulations included 12,305 and 10,594 water molecules, respectively.The long simulation in this case was 20.0 nanoseconds and performed in the NVT ensemble. As before, 1,000 configurations were used as starting points for short constant energy simulations.

III. RESULTS

A. Vibrational Frequency Mapping

In the vibrational frequency mapping approach, a mixed quantum-classical approach is used to generate empirical correlations between quantities that are inherently quantum mechanical (*e.g.*, transition dipoles) and quantities that describe the local solvation environment that can be calculated from a classical molecular dynamics simulation. In this method, the local electric field generated by the nearby classical point charges is projected onto the oscillator; though, it is worth noting that more complicated approaches involving machine learning have been proposed and developed.[64,65] For the most part, however, such maps that have been proposed for HOD take the form,

$$\tilde{\Upsilon} = v_1 + v_2\mathcal{E} + v_3\mathcal{E}^2, \tag{1}$$

where $\mathcal{E}$ is electric field projected onto the OH bond vector, v_1, v_2, and v_3 are the parameters of the fit, and $\tilde{\Upsilon}$ is the value of the quantum mechanical observable, Υ, that is reconstructed using the empirical map.

In the following section, we develop an empirical vibrational frequency map for the TIP3P water model. The parameters for this empirical map may be found in Table I. An updated SPC/E water empirical map has also been generated with the new approaches developed over the course of this present work, specific fitting details of that map are included in the Supporting Information in Figure S1 and Table S1.

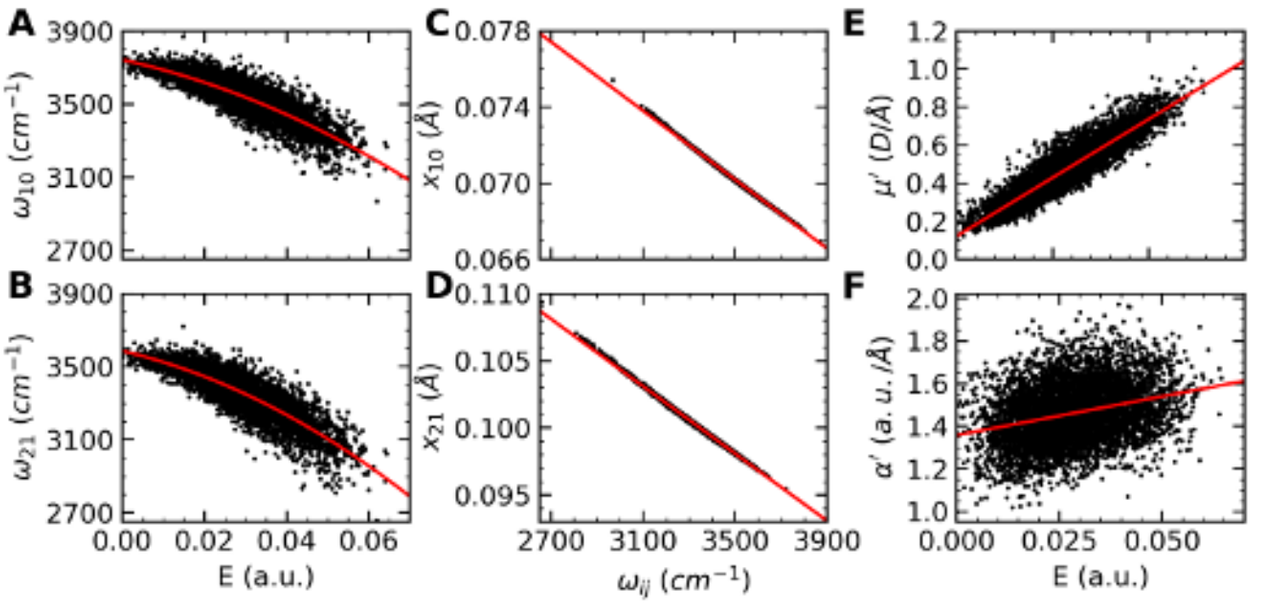

FIG. 1. Fitted parameters for the TIP3P empirical map. Black points are the data calculated from the DVR calculations, and the red lines are the empirical map fit. The transition maps for the ω_{10} and ω_{21} vibrational frequencies are shown in panels (A) and (B), respectively.The position matrix element maps for x_{10} and x_{21} are shown in panels (C) and (D), respectively. For (C) $\omega_{ij} = \omega_{10}$, and for (D), $\omega_{ij} = \omega_{21}$. The dipole derivative map and polarizability derivative map are shown in panels (E) and (F), respectively.

TABLE I. Empirical mapping parameters for the TIP3P water model with a cutoff of 8.31 Å used for the electric field. R is the correlation coefficient, which is most valid for linear regressions. $RMSE$ is the root-mean-squared error of the fit. Units of fitting parameters are enclosed in square brackets. Electric field units are a.u.

Empirical Map	R	$RMSE$
$\omega_{10} = 3742.81\,[\mathrm{cm}^{-1}] - 4884.72\left[\frac{\mathrm{cm}^{-1}}{\mathrm{a.u.}}\right] E - 65278.36\left[\frac{\mathrm{cm}^{-1}}{\mathrm{a.u.}^2}\right] E^2$	0.87	57.160 cm^{-1}
$\omega_{21} = 3586.21\,[\mathrm{cm}^{-1}] - 5251.38\left[\frac{\mathrm{cm}^{-1}}{\mathrm{a.u.}}\right] E - 87664.50\left[\frac{\mathrm{cm}^{-1}}{\mathrm{a.u.}^2}\right] E^2$	0.87	66.718 cm^{-1}
$x_{10} = 0.1019\,[\mathrm{Å}] - 9.0611 \times 10^{-06}\left[\frac{\mathrm{Å}}{\mathrm{cm}^{-1}}\right] \omega_{10}$	1.00	3.3445×10^{-05} Å
$x_{21} = 0.1420\,[\mathrm{Å}] - 1.2533 \times 10^{-05}\left[\frac{\mathrm{Å}}{\mathrm{cm}^{-1}}\right] \omega_{10}$	1.00	6.9342×10^{-05} Å
$\mu' = 0.12\,[\mathrm{D}] + 12.28\left[\frac{\mathrm{D}}{\mathrm{a.u.}}\right] E$	0.92	6.0360×10^{-02} D
$\alpha' = 1.36\,[\mathrm{a.u.}] + 3.63\left[\frac{\mathrm{a.u.}}{\mathrm{a.u.}}\right] E$	0.30	1.3448×10^{-01} a.u.

1. *Construction and Solution of 1D Potential*

The first step of this approach is to run single-point electronic structure calculations on droplet configurations that have been carved out of trajectories from classical molecular dynamics simulations. These droplets are treated such that the inner sphere, which resides around the oscillating atom of interest (for instance the H in HOD), is treated quantum mechanically, and the outer sphere is treated as classical point charges. A series of single-point electronic structure calculations where the vibration of interest is stretched are run at a reasonable level of theory (*i.e.*, B3LYP-D3/6-311G(d,p)), to obtain a 1D potential energy surface.

In the presently developed map for TIP3P, we then ran electronic structure calculations where the inner "quantum" sphere was cutoff at 4.0 Å, and the outer "classical" sphere was cutoff at 8.31 Å which is the approximate location of the final minimum for the TIP3P model's radial distribution function. In practice, the exact value of this cutoff appears to be relatively insignificant as long as

molecules are kept whole across the boundary.

From these electronic structure calculations, the energies, lab-frame dipole moments, and lab-frame isotropic polarizabilities are saved as functions of OH bond length. The latter two of these will be used below to calculate the derivatives of the dipole and polarizability. The final 1D potential is then constructed by fitting the OH bond length versus energy data to a high-order (8th order) polynomial. The sinc-function discrete variable representation (DVR) approach is then applied as described in Ref. 66 for a radial coordinate, where the potential energy contribution to the Hamiltonian comes from the eighth-order polynomial evaluated at DVR grid points.

2. Transition Frequencies

The 0-1 and 1-2 transition frequencies are evaluated through the empirical mapping approach as,

$$\omega_{ji} = \epsilon_j - \epsilon_i \tag{2}$$

where here ϵ_i and ϵ_j are the energy eigenvalues associated with the i^{th} and j^{th} energy levels, respectively. Note that atomic units, where $\hbar = 1$, are used in the DVR approach. Our TIP3P map for the transition frequencies is quadratic as a function of electric field, and is included in Table I and shown, along with the raw frequency data, in Figure 1A and B.

We have calculated the correlation coefficient, R, for both maps and found them to be 0.87 for both which is similar to the maps reported for the SPC/E model.[34] It should be noted that for quadratic fits, the root-mean-squared error (RMSE) is a better measure of agreement of the fit than the correlation coefficient. The RMSE is 57.16 cm^{-1} and 66.72 cm^{-1} for the 1-0 and 2-1 transitions, respectively. These are approximately 20 cm^{-1} smaller than the SPC/E map from Ref. 34. In addition to the presently developed TIP3P map, we likewise have generated a new SPC/E water map using the same number of configurations to which we will compare our calculated spectra.

3. Transition Dipoles

In this approach, the projection of the transition dipole moment onto the laboratory-fixed $\hat{p}$ axis can be approximately reconstructed through the bond dipole model,[68] as

$$\vec{\mu}_{ji}(t) = \mu' x_{ji} (\hat{u} \cdot \hat{p}). \tag{3}$$

Here, $\hat{u}$ is a vector pointing along the OH bond, μ' is the derivative of the dipole moment evaluated at the equilibrium bond length, and x_{ji} is the position matrix element, where again i and j refer to vibrational energy levels.

The position matrix element is calculated through the equation,

$$x_{ji} = \langle \psi_j | r | \psi_i \rangle, \tag{4}$$

where ψ_i and ψ_j are the i^{th} and j^{th} eigenvectors. This matrix element is calculated within the DVR approach, the details of which can be found in Ref. 66. The position matrix elements are typically reported as a linear map as a function of the ji vibrational frequency, ω_{ji}, though in some cases a quadratic map may be more appropriate. In this case, we built a linear map for the position matrix element and have included it in Table I and shown it in Figure 1C and D. The position matrix element maps have incredibly high correlation coefficient values, $R = 1.0$, with RMSE values on the order of 10^{-5} Å. These errors are slightly smaller than those reported for the SPC/E map,[34] likely due to the increased number of configurations used in the present work.

The dipole derivative, μ', which is needed to reconstruct the transition dipole moments is obtained by fitting the dipole moments calculated from the electronic structure calculations as a function of OH bond distance to a high-order (*i.e.*, sixth-order) polynomial. Taking advantage of the fact that the derivative of a polynomial can be calculated analytically, this provides a fitting function for the derivative that is of one polynomial order less than the original polynomial (*i.e.*, fifth order). The value of μ' is then evaluated from this polynomial at the equilibrium bond distance and this is used to build a map for μ'. Linear and quadratic maps for μ' have both been used for various water models; Skinner's SPC/E water dipole derivative map was linear in the electric field whereas their TIP4P and E3B3 empirical maps were quadratic. We find that for TIP3P water, a linear map is sufficient (using a quadratic map did not improve the fit significantly, and resulted in the wrong sign for the quadratic term.) The map for μ' is included in Table I and shown in Figure 1 E. As with the other maps, our dipole derivative map has a small RMSE value, 6.0360×10^{-2} D, and a correlation coefficient of 0.92.

In the literature, maps have been reported for μ' in one of two ways, the first is to list μ' in terms of a reference, gas-phase calculation where the same approach is used without the surrounding solvent. In these terms, the map has been reported as μ'/μ'_g, where μ'_g is the gas-phase value of the dipole derivative. The second, which we employ in this work, is to report the μ' map directly. The key difference between the two mapping approaches is that excluding the factor of μ'_g leads to smaller calculated values for the intensity. Since the mapping approach is not well-suited for calculating absolute intensities, the spectra are often normalized based off peak intensity. This factor of μ'_g disappears upon normalization of the spectra, and thus can be safely omitted.

4. Transition Polarizabilities

In a similar manner to the dipole derivative, the isotropic transition polarizability can be reconstructed through the equation

$$\alpha_{10} = \alpha' x_{10}. \tag{5}$$

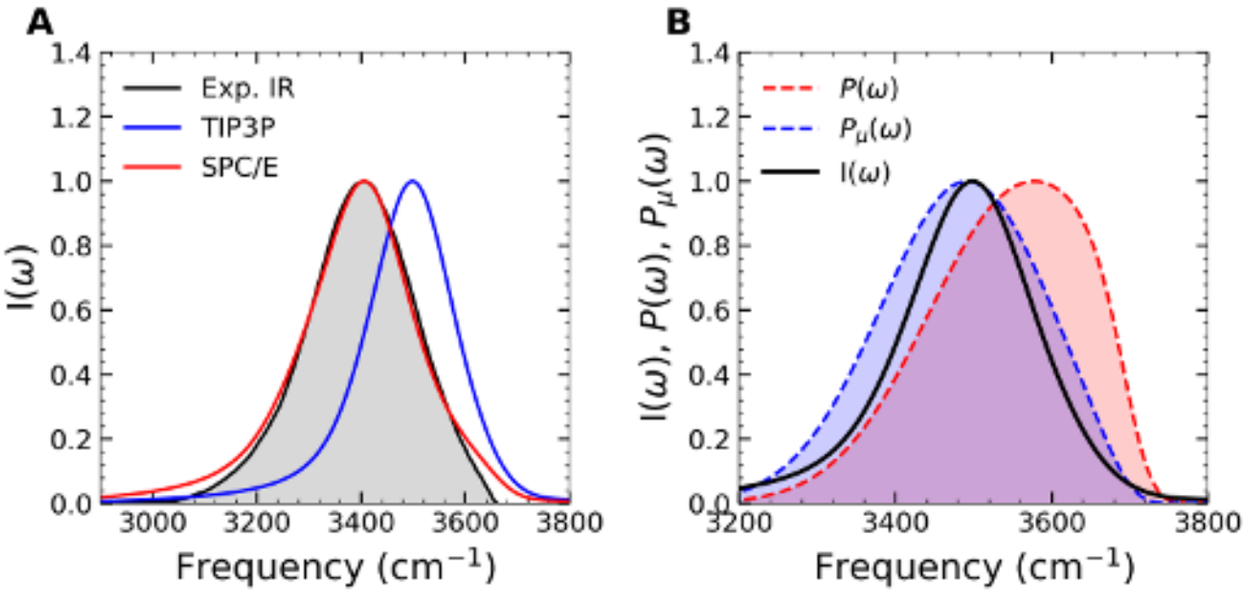

FIG. 2. A) The linear infrared spectrum of the TIP3P model (blue solid line) and the SPC/E model (magenta dashed line) are compared with the experimental HOD/D_2O infrared spectrum (black solid line). The experimental spectral data has been digitized from Ref. 67. B) The TIP3P spectrum is reproduced (as a black solid line), and compared with the frequency distribution (red dashed line), and the spectral density (blue dashed line). In both plots, uncertainties in computational data are smaller than the line thickness and have thus been omitted.

where α_{10} is the isotropic transition polarizability between the ground and excited vibrational states, and α' is the isotropic transition polarizability derivative which is calculated in an identical manner to μ' (using the isotropic polarizability instead of the dipole moment from the electronic structure calculations).

This transition polarizability is often used in the context of the bond polarizability model,[69] which describes the polarizability in terms of one component parallel to the OH bond and two components perpendicular to the bond, to recover components of the polarizability tensor as

$$\alpha_{pq} = (\alpha'_{\parallel} - \alpha'_{\perp})x_{10}(\hat{p}\cdot\hat{u})(\hat{u}\cdot\hat{q}) + \alpha'_{\perp}x_{10}(\hat{p}\cdot\hat{q}) \quad (6)$$

where α_{pq} is a component of the polarizability tensor, and $\alpha_{\parallel}$' and $\alpha'_{\perp}$ are the parallel and perpendicular components of the polarizability derivative, respectively. It is shown in the Supporting Information that these components can be evaluated in terms of the isotropic transition polarizability as

$$\alpha'_{\parallel} = \frac{3}{1+\frac{2}{\gamma}}\alpha'_{iso} \quad (7)$$

and

$$\alpha'_{\perp} = \frac{3}{2+\gamma}\alpha'_{iso}. \quad (8)$$

Here $\gamma = \alpha_{\parallel}/\alpha_{\perp}$ is the ratio of parallel and perpendicular components of the polarizability to the bond vector, and α_{iso} is the isotropic polarizability. As is frequently done, this value has been selected to be 5.6 for HOD/D_2O;[70–72] however, other values have been proposed.[73,74]

Maps for the polarizability derivative are often reported as linear maps with respect to the electric field, which we have mimicked in the present work. The map for α' is included in Table I and shown in Figure 1F. It should be noted, however, that relatively few previous works have included plots like that in Figure 1F due to the poor correlation of the isotropic polarizability derivative with the electric field. This comparatively weaker correlation is often attributed to the fact that the non-Condon effect is much weaker for polarizability than it is for the other mapping parameters. This is reflected in the R^2 value of of 0.30 obtained for our map in this work, which is similar to the value of 0.31 reported by Skinner and co-workers for SPC/E water.[34] Though poorly correlated, the RMSE value of 0.13448 a.u. is relatively small, and the map is included for completeness. For comparison, we have included the equivalent figure and table for the updated SPC/E map in the Supporting Information.

B. Vibrational Spectroscopy of TIP3P Water

The vibrational frequency mapping approach has been used to predict the lineshapes of common vibrational spectroscopies, including linear and two-dimensional infrared spectroscopy, Raman spectroscopy, and sum-frequency generation. For each spectroscopy, we will include a discussion of how the map is used to calculate each spectrum, and how that spectrum compares with both experimental measurements and the results from the SPC/E water model.

1. Linear Infrared Spectroscopy

The infrared lineshape of water can be calculated as the Fourier Transform of the dipole-dipole response function, $\phi(t)$, as

$$I(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i(\omega-\langle\omega\rangle)t}\phi(t)\mathrm{d}t \tag{9}$$

where,

$$\phi(t) = \left\langle \vec{\mu}_{10}(0)\cdot\vec{\mu}_{10}(t)e^{i\int_0^t \delta\omega(\tau)\mathrm{d}\tau}\right\rangle e^{-\frac{|t|}{T_1}}. \tag{10}$$

Here, $\vec{\mu}_{10}$ is calculated through Eq. (3) using the vibrational frequency mapping approach, t is the time, and T_1 is the vibrational lifetime. For the OH stretch of water, a value of 700 fs is commonly taken from experiments used for T_1 as it is only weakly temperature dependent.[67]

In Figure 2 we have reported the linear infrared spectra alongside the experimental HOD/D_2O infrared spectrum which has been digitized from Ref. 67 using `WebPlotDigitizer`.[75]. The TIP3P spectrum, which peaks at 3497 cm^{-1}, is significantly blue shifted compared to the experimental spectral peak at 3405 cm^{-1}, by 92 cm^{-1}. Furthermore, the simulated spectrum is significantly narrowed compared to the experimental spectrum, as can be observed from their full width half maximum (FWHM) values of 187 cm^{-1} and 232 cm^{-1}, respectively. The SPC/E model has a peak located at 3404 cm^{-1}, in good agreement with the experimental spectral peak position; additionally, the FWHM (228 cm^{-1}) is similar to what is observed experimentally.

To explore this further, we have calculated the distribution of frequencies and the spectral density which provide unique information about the origin of the IR spectra. Specifically, the distribution of frequencies, which is calculated as

$$P(\omega) = \langle\delta[\omega-\omega_{10}]\rangle, \tag{11}$$

is a useful quantity as it removes both dynamical effects (*e.g.*, motional narrowing) and non-Condon effects that come from the fact that the transition dipole moment depends on H-bond strength.

Likewise, the spectral density, which provides information about the non-Condon effects in the transition dipole moment while still excluding dynamical considerations, can be calculated as

$$P_\mu(\omega) = \langle|\mu_{10}|^2\delta[\omega-\omega_{10}]\rangle. \tag{12}$$

We have included both $P(\omega)$ and $P_\mu(\omega)$ in Figure 2 for comparison with the IR spectrum. The spectral density is only slightly broader than the IR spectrum, with a full-width half-maximum of 248 cm^{-1}. Since the spectral density effectively reports the impact of the transition dipole moment on the frequency distribution without incorporating the impact of dynamics, this suggests that the narrowness of the simulated spectra cannot only originate from motional narrowing. Importantly, the ratio of the spectral density FWHM to the IR spectra FWHM is 1.326 for TIP3P, which is nearly identical to that of the SPC/E model, which we find to be 1.333 (based on a spectral density FWHM of 304 cm^{-1}). This suggests that, surprisingly, there is limited additional motional narrowing present within the TIP3P model compared to the SPC/E model, despite their vastly different dynamics.[43,76] Instead, the relative additional narrowness of the TIP3P spectra originates from the narrower distribution of frequencies compared to SPC/E (248 cm^{-1} *vs.* 351 cm^{-1}).

2. Raman Spectroscopy

The Raman spectrum, in a clear parallel with the Infrared spectrum, can be calculated as the Fourier Transform of the isotropic polarizability-polarizability response function, $\phi_{iso}(t)$, as

$$I_{iso}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i(\omega-\langle\omega\rangle)t}\phi_{iso}(t)\mathrm{d}t \tag{13}$$

where,

$$\phi_{iso}(t) = \left\langle \alpha_{10}(0)\cdot\alpha_{10}(t)e^{i\int_0^t \delta\omega(\tau)\mathrm{d}\tau}\right\rangle e^{-\frac{|t|}{T_1}}. \tag{14}$$

In an analogous way, α_{10} is calculated using the bond polarizability model, as described above in Eq. (5).

We have calculated the isotropic Raman spectra using our new TIP3P water map, and have included a comparison of this spectra with the experimental Raman spectra measured by Smith and co-workers in Figure 3.[77] As was the case with the IR spectrum, the TIP3P Raman spectrum is both significantly narrowed and blue-shifted compared to the experimental spectrum.

Importantly, the Raman spectrum takes on a very different shape than what is observed for both the SPC/E and experimental spectra, where the shoulder present in those spectra is not apparent in the TIP3P spectra. This shoulder originates from water OHs that are involved in

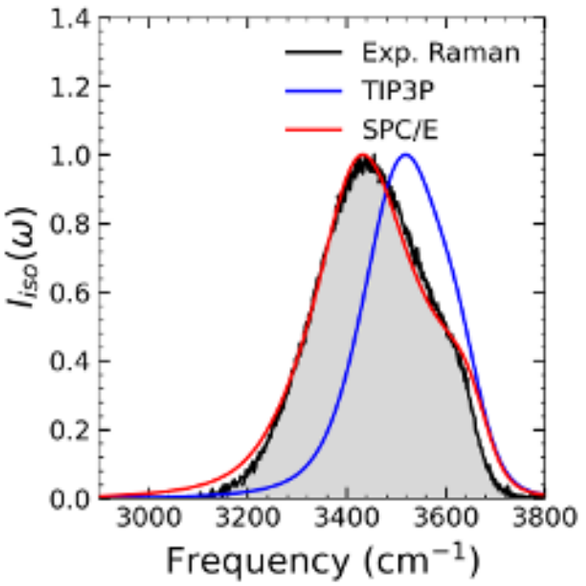

FIG. 3. The linear Raman spectrum of the TIP3P model (blue solid line) and the SPC/E model (magenta dashed line) are compared with the experimental HOD/D_2O Raman spectrum (black solid line). The experimental spectral data has been digitized from Ref. 77. Uncertainties in computational data are smaller than the line thickness and have thus been omitted.

weak hydrogen bonds, which leads to the blue-shifted frequencies observed. This, along with fact that hydrogen bonds in the TIP3P model are significantly more labile than those observed in other models,[76] is likely behind both the dissapearance of this shoulder and the overall shift to higher energies of the OH vibrations in the Raman spectrum. The main peak of the Raman spectrum is blueshifted from the experimental value of 3441 cm^{-1} by 79 cm^{-1} to 3520 cm^{-1} for the TIP3P model. As the experimental spectrum includes more noise than the IR spectra, the maximum was determined by fitting the data around the peak to a quadratic function. The SPC/E model retains a peak position of 3432 cm^{-1}, in good agreement with experiments. It is worth noting that this is not a general failure of three-point water models, our newly generated SPC/E empirical map also visible in Figure 3, is able to nearly perfectly reproduce the HOD/D_2O experimental Raman spectrum with only a slight broadening at high and low frequencies.

3. Two-Dimensional Infrared Spectroscopy

As a three-pulse echo experiment, the response functions for two-dimensional infrared (2DIR) spectroscopy take on a significantly more complicated form.[78,79] In short, these response functions take the form $R_j(t_1, T_w, t_3)$ where $j = 1 \cdots 6$. Here, t_1 is the delay between the first and second pulses, T_w is the waiting time between the second and third pulses, and t_3 is the time between the third pulse and the echo.

These six time-correlation functions can be separated into rephasing and non-rephasing contributions, written as

$$R_r(t_1, T_w, t_3) = \sum_{j=1}^{3} R_j(t_1, T_w, t_3) \tag{15}$$

and

$$R_{nr}(t_1, T_w, t_3) = \sum_{j=4}^{6} R_j(t_1, T_w, t_3) \tag{16}$$

respectively.

The form of these time-correlation functions are reproduced in full from Ref. 80 in the Supporting Information.

The 2DIR spectrum is calculated from the real part of the double-Fourier transform as

$$\begin{aligned} I_{2DIR}(\omega_1, \omega_3; T_w) \;=\; & Re\left\{ \int_0^\infty \mathrm{d}t_3 e^{i\omega_3 t_3} \int_0^\infty \mathrm{d}t_1 \right. \\ & \left[R_r(t_1, T_w, t_3) e^{-i\omega_1 t_1} \right. \\ & + \left. \left. R_{nr}(t_1, T_w, t_3) e^{i\omega_1 t_1} \right] \right\} \end{aligned} \tag{17}$$

where here, T_w is the waiting time, and ω_1 and ω_3 are the pump and probe frequencies.[80]

In Figure 4 we present a comparison between the 2DIR spectra of the TIP3P and SPC/E water models. A few key differences are clearly apparent, the TIP3P spectra are significantly narrower than their SPC/E counterparts. Furthermore, as was the case with the linear spectra, the TIP3P 2DIR spectra are blueshifted not only in ω_1, but also ω_3 as well. With that said, the lineshapes are similar between the models and the loss of correlation in the center of lines slope between the two models is similar. Values of this slope are included in the Supporting Information, and as will be discussed below, follow the trends observed for the spectral diffusion times calculated below.

4. Spectral Diffusion Dynamics

In liquids, spectral diffusion is a measure of how frequencies evolve in time and thus is sensitive primarily to the evolution of the local solvation environment surrounding a particular mode. From 2DIR measurements (and calculations) the spectral diffusion timescale can be estimated through the center-lines slope (CLS), which is pictured in Figure S2. However, from simulations, it is often more convenient to calculate the spectral diffusion timescale from the frequency-frequency correlation function, as

$$C_\omega(t) = \frac{\langle \delta\omega_{10}(0)\delta\omega_{10}(t) \rangle}{\langle \delta\omega_{10}(0)^2 \rangle} \tag{18}$$

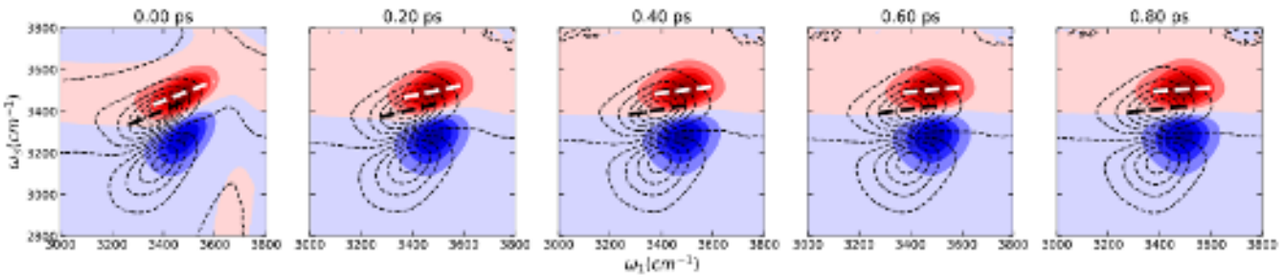

FIG. 4. The 2DIR spectra of the TIP3P water model has been included. From left to right, panels correspond to waiting times of 0.0 ps, 0.2 ps, 0.4 ps, 0.6 ps, and 0.8 ps. White dashed lines are used to illustrate the center-of-lines slope. The SPC/E 2D-IR spectra is included as black dashed lines, with the center-of-lines slope included in black.

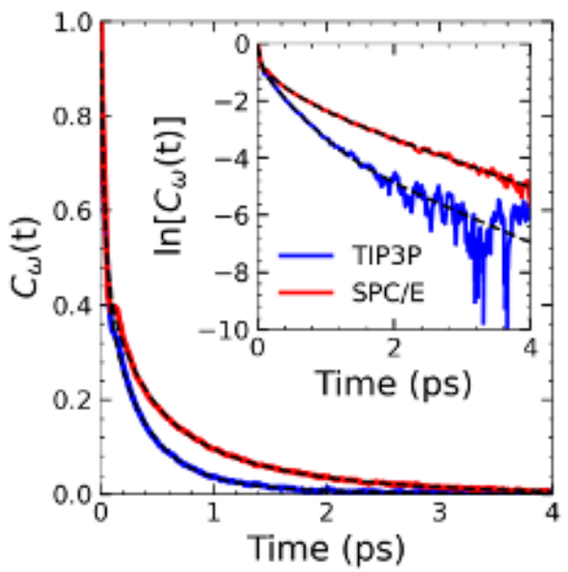

FIG. 5. The frequency-frequency correlation function of TIP3P water and SPC/E water are included in blue and red, respectively. Fits to a tri-exponential fitting function are included as black dashed lines. Data are presented on linear, and semi-log scale in the main panel and inset, respectively.

where $\delta\omega_{10} = \omega_{10} - \langle\omega_{10}\rangle$. This time-correlation function is often fit to a tri-exponential fitting function, of the form,

$$C_\omega(t) = A_1 e^{-\frac{t}{\tau_1}} + A_2 e^{-\frac{t}{\tau_2}} + A_3 e^{-\frac{t}{\tau_3}} \qquad (19)$$

where $A_i, i = 1 \cdots 3$ are the amplitudes corresponding to the three timescales, τ_i. Here, we included the constraint that $A_3 = 1 - A_1 - A_2$. These timescales, in order from fastest to slowest, are typically attributed to inertial motion of the unbroken hydrogen bond, librational motion (rocking within an unbroken hydrogen bond), and spectral diffusion. This latter timescale, which is typically on the order of a picosecond, is used as a measure for the timescale on which an OH at a particular frequency loses memory of that frequency.

In practice, under normal sampling conditions these are the only timescales that are easily observable from molecular simulations; and thus for many years the longest timescale was thought to be the hydrogen-bond exchange jump timescale. However, it has been shown previously that with high levels of sampling, the jump timescale appears as a fourth timescale (using a quad-exponential fit) in the decay as an approximately 1% contribution to the overall decay of the FFCF (and thus, as-of-yet has not been observed experimentally). Thus the spectral diffusion time, while being a measure of the changing hydrogen bonding environment surrounding the OH, is not a direct measure of that OH's hydrogen bond exchange time.[45]

We have calculated the FFCF for both TIP3P and SPC/E water, and have included plots of these, along with their fits, in Figure 5. We observe that the FFCF decays much more quickly for TIP3P water than it does for SPC/E water, suggesting that spectral diffusion occurs more quickly for the former than the latter.

Fitting the FFCF to the triple exponential fitting function described by Eq. 19 provides useful insight into the differences between the dynamics of these models. The inertial and librational timescales in the FFCF are similar for both TIP3P and SPC/E. The inertial timescales of these models are 36.3 fs ± 1.8 fs and 33.9 fs ± 0.7 fs, respectively. The librational timescales are 314 fs ± 35 fs and 333 fs ± 21 fs, respectively. Finally, we find that the spectral diffusion timescale is 1.00 ps (± 0.41 ps) and 1.19 (± 0.14 ps), respectively. The faster decay for the TIP3P model compared to SPC/E is consistent with what has been observed in the CLS slope of the 2DIR spectra (see Supporting Information).

While resolving spectral diffusion timescales outside of a 95% confidence interval proved challenging due to the significantly larger fluctuations in the TIP3P FFCF surrounding the long-time decay, its important to note that the TIP3P spectral diffusion timescale has a significantly lower contribution (5.3 ± 3.6 %) to the FFCF decay than observed for SPC/E (18.9 ± 3.0%). This suggests that the changing hydrogen-bond environment of the water molecules in proximity to a particular OH plays a less

significant role overall in driving changes in that OH's frequency dynamics.

5. Sum-Frequency Generation Spectroscopy

Sum-frequency generation spectroscopy is a surface-sensitive spectroscopy in which a beam of infrared light polarized in the $\hat{k}$ direction with frequency ω_{IR} and a visible beam polarized in the $\hat{j}$ direction with frequency ω_{VIS} are shined on a surface. This results in scattering of light with a frequency of $\omega_S = \omega_{IR} + \omega_{VIS}$, with an intensity that is proportional to the square of the magnitude of the second-order nonlinear susceptibility $\chi_{ijk}(\omega_{IR}, \omega_{VIS})$.

The SFG intensity can be expressed as,

$$I(\omega) \propto \left|\chi^R_{ijk}(\omega_{IR}) + \chi^{NR}_{ijk}\right|^2 \tag{20}$$

where a fixed visible frequency has been assumed. ω_{IR} is the IR frequency, χ^R_{ijk} is the resonant contribution of the nonlinear susceptibility, and χ^{NR}_{ijk} is the non-resonant contribution. As discussed in Ref. 35, the non-resonant component of the susceptibility is often treated as an adjustable constant, rather than evaluating it directly. Importantly, the imaginary part of the resonant second-order susceptibility, $\mathrm{Im}\left[\chi^R(\omega)\right]$ can be measured directly through experiment.[81–86]

From a simulation, the resonant non-linear susceptibility is calculated in much the same way as has been done for the IR and Raman spectra above. To express this in a parallel form, we can write

$$\chi^R_{ijk}(\omega) \propto i\int_0^\infty e^{i\omega t}\phi_{SFG}(t)\mathrm{d}t \tag{21}$$

where $\phi_{SFG}(t)$ is the dipole-polarizability response function, which can be expressed as

$$\phi_{SFG}(t) = \left\langle \alpha_{ij}(t)\cdot\mu_k(0)e^{i\int_0^t \delta\omega(\tau)\mathrm{d}\tau}\right\rangle e^{-\frac{|t|}{2T_1}}. \tag{22}$$

Here, $\alpha_{ij}(t)$ is an element of the transition polarizability tensor, and μ_k is the k^{th} element of the transition dipole moment tensor.

As SFG is sensitive to disruptions of centrosymmetry, how contributions of water molecules far from the interface are considered can have a significant impact on such calculations. In the present work, we use the approach of Byrnes and co-workers, which found that dividing contributions from the top-half and bottom-half of the interface is best done based on the location of the water oxygen atom.[87] Other approaches, for instance using the hydrogen atom instead of the oxygen atom can have significant impacts on the line shape. It should be noted, as was described in that work, that the present selection is appropriate for *ssp*, and *sps* polarizations, *pss* polarizations instead should use a location located 0.15 A from the O atom along the OH bond.

In the present work, we have used sum-frequency generation spectroscopy to study two interfaces of common interest. The first is the air-water interface, which has received much attention and for which high-quality experimental data exists as a comparison point, and the second is the lipid-water interface, for which some prior computational studies exist; however, there is relatively little experimental data thus far.

Air-Water Interface:

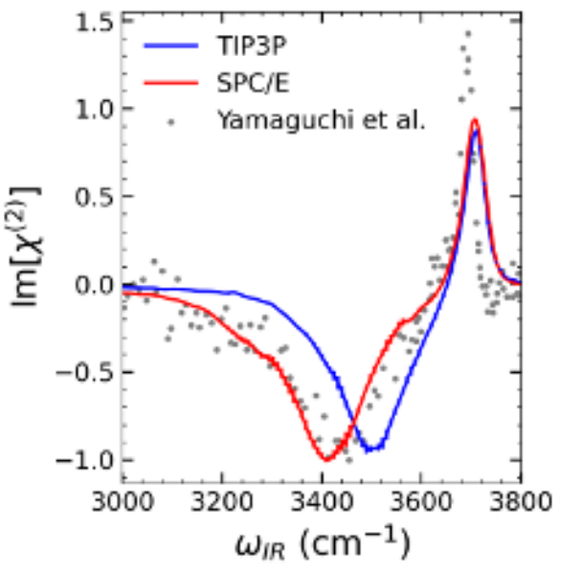

FIG. 6. The liquid/vapor interface HOD/D_2O SFG spectrum calculated using the TIP3P and SPC/E water models are included in blue and red, respectively. The experimental air-water spectrum from Ref. 88 is included as a gray circles (data digitized using `WebPlotDigitizer`).[75]

As described above, we calculated the SFG spectra for an air-water interface in a slab geometry using both the TIP3P and SPC/E water models. The air-water interface has received significantly more attention from both computation[90–92] and experiment[89,93,94] than the lipid-water interface, and thus is an ideal test-case for understanding the ability of gaining insight from TIP3P SFG calculations.

These results are included in Figure 6. The maximum in the imaginary spectrum (located at 3,709 cm^{-1} for TIP3P and 3,705 cm^{-1} for SPC/E) is consistent between both models, and is typically interpreted as OH groups that are "dangling" at the interface, and not participating in hydrogen bonds. On the other hand, the negative peak (located at 3,493 cm^{-1} for TIP3P and 3,412 cm^{-1} for SPC/E) is the contribution that comes from hydrogen bonded water molecules at the interface, and corresponds to a blue-shift from the SPC/E result of similar magnitude observed for the IR and Raman spectrum, of about 85 cm^{-1}. This is consistent with the fact that this negative peak corresponds to the contributions of hydrogen

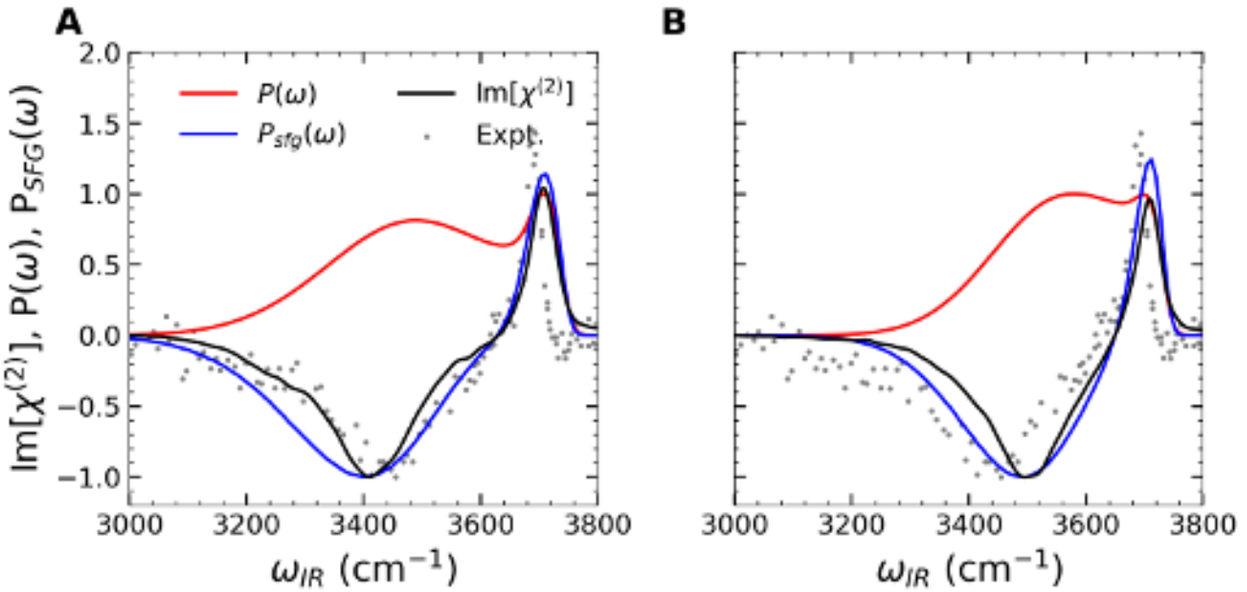

FIG. 7. The decomposition of the liquid/vapor interface HOD/D_2O SFG spectrum calculated using the (A) SPC/E and (B) TIP3P water models into the imaginary spectrum, spectral density, and interfacial frequency distribution is shown in black, blue, and red, respectively. The experimental result from Ref. 88 is included as gray circles in both figures.[89]

bonded waters at the interface.

Interestingly, the experimental results from Ref. 89 indicate that the negative peak falls in between these two results, at 3,454 cm^{-1}; however, it should be noted that other experiments (with pure H_2O, as well as isotopically dilute HOD in D_2O) have found differing locations for the minimum.[93–96] A comparison with these experiments, which agree better with the minimum location of the SPC/E model, is included in the Supporting Information. Variations in the techniques used to generate the SFG spectrum have recently been demonstrated to have significant influences on the overall peak positions and relative intensities as is discussed in detail in Ref. 97

As was the case with the infrared spectrum, more information can be gained about the origin of the two peaks in the sum-frequency generation spectrum by examining the spectral density and the frequency distribution associated with SFG spectroscopy. The former, as it depends on the *signed* polarizability, is clearly defined for the interface; however, the latter which represents the distribution of all OH frequencies in the system also includes the response of the bulk OHs, thus, it is necessary to select OHs only in the interfacial region.

The position-resolved frequency distribution can be written as,

$$P(\omega_{IR}; z) = \langle \delta(\omega - \omega_{IR}) \delta(z - z_c) \rangle \,, \tag{23}$$

where here z and z_c are the z-position of the OH, and the z-position of the center of the box, respectively. This gives a two-dimensional distribution which describes the OH frequency distribution as a function of distance from the center of the slab. The full distribution, calculated from Eq. (23) is included in the Supporting Information in Figure S9.

Likewise, the total spectral density associated with sum-frequency generation can be written as,

$$P_{SFG}(\omega_{IR}) = \langle \mu_{10} \alpha_{ij} \delta(\omega - \omega_{IR}) \rangle \,. \tag{24}$$

For completeness, it should also be noted that the spectral density can likewise be separated onto a z-position domain as well as has been done by Auer and Skinner,[98] as

$$P_{SFG}(\omega_{IR}; z) = \langle \mu_{10} \alpha_{ij} \delta(\omega - \omega_{IR}) \delta(z - z_c) \rangle \,. \tag{25}$$

As with the two-dimensional frequency distribution, the result using Eq. (25) may be found in the Supporting Information in Figure S10.

In Figure 7 we have included a comparison of the imaginary part of the SFG spectrum along with the spectral density calculated using Eq. (24), and a slice from the 2D frequency distribution calculated using Eq. (23) selected at the interface calculated for both the SPC/E and TIP3P water models. This slice, calculated in a 2 Å bin, was selected based on the position of the maximum in the 2D spectral density. Similar to the IR spectrum, the (negative) hydrogen-bonded peak of the SFG spectrum for both models is red-shifted in comparison to the frequency distribution at the interface. The shoulder, which separates the hydrogen-bonded and dangling OH peaks

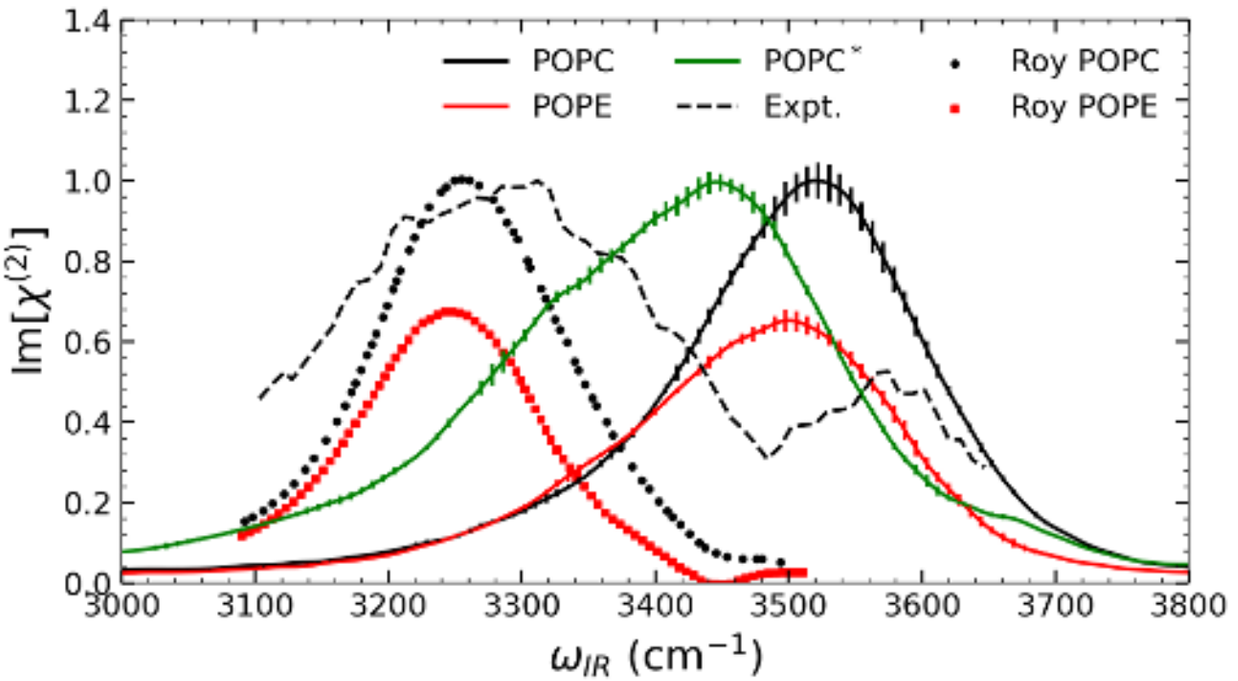

FIG. 8. The HOD/D_2O sum-frequency generation spectra at a POPC and POPE interface (calculated using the TIP3P water model) are included in black and red, respectively. The experimental POPC sum-frequency generation spectrum from Ref. 27 is included as a black dashed line. Previous results from Roy and Skinner using the SPC/E water model and the Berger lipid model are includes as black circles and red squares for POPC, and POPE, respectively.[24] A modified simulation, where the SPC/E model was used for both the simulation and the mapping calculation is included in green.

at the interface, is observed in the SPC/E and experimental SFG imaginary spectra and also reflected in the SPC/E spectral density. This shoulder is missing from both the TIP3P imaginary spectrum and its spectral density.

The frequency distribution, however, suggests that the origin of this shoulder in the SPC/E spectrum comes from the relative red-shift of the hydrogen-bonded peak compared to the dangling OH peak. In the TIP3P spectrum, where the two peaks have significantly more overlap this shoulder disappears. Thus, the qualitative differences observed in the TIP3P spectrum again originates from its poorer description of hydrogen-bonded configurations. A more detailed analysis of the spectral density is included in the Supplementary Information, as are the two-dimensional distributions of frequencies and spectral density calculated from Eqs. (23) and (25), respectively.

Despite these differences, the TIP3P model is able to qualitatively capture these two regions in the spectra that are observed in other models and in the experiments, and thus at least for this system is sufficient to provide qualitative insight.

Lipid-Water Interface:

A key draw, however, is the ability to use the TIP3P model as a probe of water spectroscopy at biochemical interfaces. To that end, we have calculated the SFG spectra of water at the interface of two different lipids, POPC and POPE. The SFG spectra for this system has been included in Figure 8, along with the experimental POPC spectra. Unlike the other spectra calculated, which consist of only water and thus all had a similar spectral blueshift of ∼90 cm^{-1}, here the shift compared to experiment is much larger. The low-frequency experimental maximum is located at ∼ 3310 cm^{-1} (3281 cm^{-1} when fitted by quadratic), whereas the simulated peak for POPC is located at 3522 cm^{-1}. Thus, the shift is somewhere in the range between 212-241 cm^{-1}.

POPC and POPE spectra have been calculated previously, by Roy and Skinner, for a smaller lipid system using the SPC/E model with the Berger lipid forcefield which is less frequently used in the membrane simulation community.[23,24,42] Importantly, their work was able to reproduce the experimental peak position of the POPC low-frequency peak. Furthermore, they demonstrated that the POPE spectra was less intense than the POPC spectra, due to the decreased ordering of water molecules at the interface. We have digitized the results from Ref. 24 and have included them in Fig. 8 for comparison. In the present work, despite again having a significant blueshift compared to the experimental and SPC/E spectra

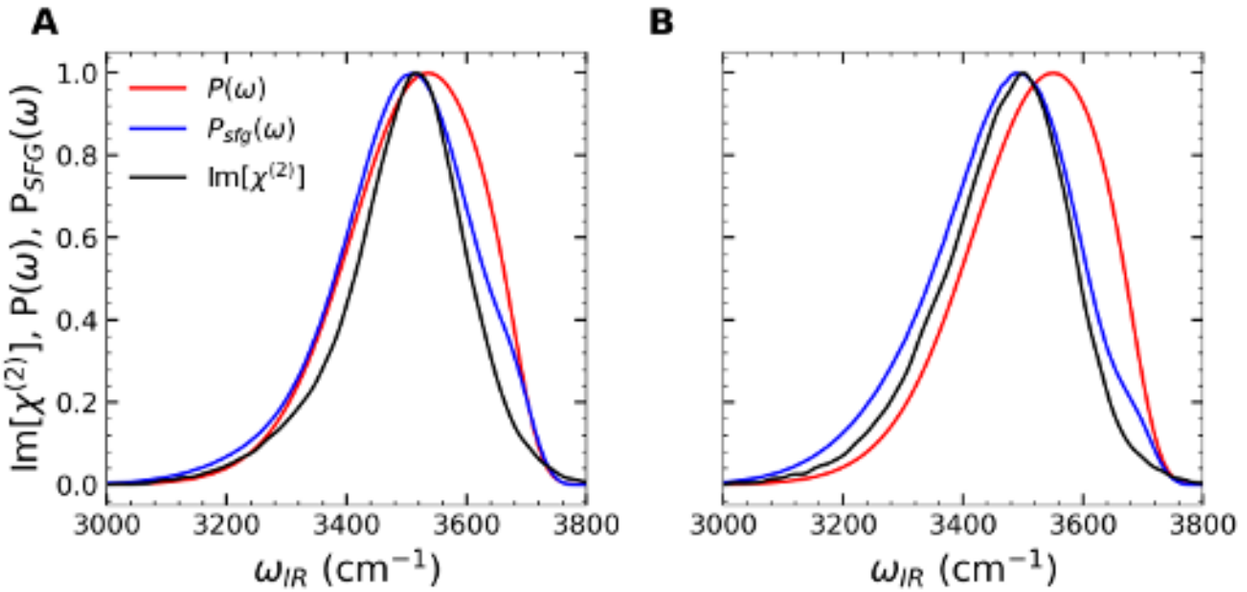

FIG. 9. The decomposition of the lipid/water interface HOD/D_2O SFG spectrum calculated for the (A) POPC and (B) POPE lipid membranes into the imaginary spectrum, spectral density, and interfacial frequency distribution is shown in black, blue, and red, respectively. All calculations are for the TIP3P water model.

of Roy and Skinner, we find the same relative magnitudes of intensities between the POPC and POPE systems that they previously observed. As will be discussed in more depth in Section IV, we believe the larger magnitude of blue shift for TIP3P for lipids originates from the description of water-lipid interactions.

As was the case for the air-water interface, we have calculated the two-dimensional spectral density and frequency distribution (which may be found in the Supporting Information, along with images of the simulation cell). In Figure 9, we have plotted the interfacial frequency distribution, alongside the spectral density and imaginary components of the SFG spectrum. Unlike the air-water interface, for which the separation of two peaks in the overall SFG spectrum explained the differences between the models, for the membrane systems the singular spectral peak is matched by singular peaks in the frequency distribution and in the spectral density.

IV. DISCUSSION

The vibrational spectroscopy of the TIP3P water model reported in the previous section exhibits significant quantitative differences from experiment, and that of the SPC/E model, which has typically been used in materials applications such as nanoporous silica.[99,100] However, in the biochemical simulation community TIP3P is a widely used model due to its persistence in the CHARMM36 force-field. Thus, a key aim of the present work has been to identify where the vibrational spectroscopy of the TIP3P model can be trusted to provide qualitative information about spectroscopy in biochemical environments, and where other better models should be explored.

TIP3P will never be a replacement for other, better models for bulk liquid water such as TIP4P/2005 which can accurately describe features such as liquid structure, dynamics, and temperature-dependent anomalies of liquid water. However, the TIP3P model was able to fairly accurately predict the shape of the infrared spectrum (though, as discussed in the above, not the position of the peak); whereas, the model offers a poor description of the shape of the Raman spectrum being both blue-shifted and missing the shoulder at high frequencies. Likewise, while the dynamics of spectral diffusion observed through both the center-of-lines slope approach and the frequency-frequency correlation function are faster for the TIP3P model, the difference is smaller than what is observed for other dynamical properties like diffusion and reorientation timescales.[76]

Where things become more complicated, however, is when there is a combination of a hydrogen-bonded band and a dangling OH band as can be found in hydrophobic environments. Due to the blue-shift of the TIP3P model spectra in H-bonded environments, this can result in significant differences between model and experimental spectra. Thus, for calculating the spectroscopy of TIP3P in biochemical environments one needs to consider whether two (or more) OH populations are expected

to be present, and how that might influence the overall observed spectroscopy.

Often, one question of interest when there are problems with a particular force-field is whether one can just swap out the mis-behaving model. Indeed, there have been a number of recent studies that have explored what happens to CHARMM36 simulations when one replaces TIP3P water with a better model.[101] While a full-exploration of the impact of swapping the water model is far beyond the scope of the present work, it was of interest to know whether replacing TIP3P with SPC/E could "correct" for the blue shift in the sum-frequency generation. Thus, we ran a POPC membrane system using the same conditions as described in methods, with the exception that we simulated the membrane system with SPC/E and then calculated the sum-frequency generation spectroscopy (using our newly generated SPC/E map). While agreement is not perfect between the experimental peak position and the new results (also included in Figure 8) it does lead to an overall red-shift towards the experimental result from 3,522 cm^{-1} to 3,445 cm^{-1}, for a total red shift of 77 cm^{-1}. This is significant, as this is nearly the amount of the blue shift that the TIP3P model exhibits in bulk spectroscopy compared to experiments; however, there is a significant blue-shift compared to the experimental spectrum that remains.

It should be noted that only the SPC/E water model paired with CHARMM36 lipids is able to reproduce the double-peaked structure observed in the experimental spectrum, though the relative intensity of the peaks is not accurately predicted. Neither Roy and Skinner's previous simulations with the SPC/E water and Berger lipid combination, nor our TIP3P water CHARMM36 lipid combination were able to reproduce this two-peaked structure. Typically, the most intense experimental peak has been assigned to water molecules influenced by the phosphate and choline groups, whereas the less intense peak has been suggested to be populations of water molecules hydrogen-bonded to the lipid carbonyl group.[48] Ishiyama and co-workers have previously suggested that the polarity of the carbonyl group in the CHARMM36 force-field may need to be reduced to produce this second peak in computed spectra using the Charge Response Kernel approach.[48] We find that replacing TIP3P with SPC/E in a CHARMM36 simulation appears to recover the signatures of this peak. In part, this is likely due to the improved description of hydrogen bonding within the SPC/E model when compared with TIP3P. However, future studies will be needed to determine when such a swap is acceptable, and what, if any, impacts it has on the local membrane environment.

Overall, the empirical mapping approach continues to provide significant insights into the spectroscopy of a wide variety of systems due to its versatile ability to describe a wide-range of spectroscopies. While this is the case, it is by no means the only technique through which spectroscopy can be achieved computationally. Recent advances in neural network potentials have been used successfully to study small systems, for instance bulk water and the air-water interface, at *ab initio* accuracy.[90] Approaches for learning the dipole moments and polarizabilities from these simulations have enabled spectroscopic studies like the ones described herein, without the use of empirical mapping. Other studies have moved beyond simple electric-field maps to use additional information to machine learn relationships between local environment, frequency, and dipole moment.[90] For large, biochemical systems, empirical mapping is expected to remain an important way for the calculation of spectroscopic observables as the aforementioned techniques are still prohibitively expensive for the large system sizes and timescales involved in many biochemical simulations.

V. CONCLUSIONS

In the present work, we have developed an empirical vibrational frequency map for theTIP3P water model which is the predominant water model used in many biochemical simulations, as well as an updated map for the SPC/E water model for the HOD/D_2O vibration. We demonstrate that spectral features related to the hydrogen bonding environment of the TIP3P model are blue-shifted significantly, on the order of 100 cm^{-1} due to the formation of weaker hydrogen bonds by the model, when compared to the corresponding spectra in experiment and SPC/E as is observed through the distribution of frequencies. This blue-shift is observed throughout the vibrational spectroscopy of the model, including in the infrared, Raman, 2D-IR, and Sum-Frequency Generation spectra. In the latter, we demonstrate that due to the presence of both a hydrogen-bonded peak and a dangling OH peak for the air-water interface, a qualitative change in the overall shape of the imaginary part of the spectrum is observed as a result of enhanced spectral overlap between these two populations.

Overall, it appears that the TIP3P model is sufficient to obtain *some* qualitative insight into biochemical systems; however, for any such consideration it is important to recognize that peaks associated with hydrogen-bonded environments are likely to be significantly blue-shifted compared to experiment. For bulk water and other aqueous solutions, it is ideal to use a better water model that can provide a better description of the local solvation environment, for which a plethora of vibrational frequency maps exist. For biochemical simulations, however, this choice is not always available. Thus, we believe the presently developed vibrational frequency map can provide useful phenomenological insight into such cases; however, the presently developed approaches should not be used for calculations of exact peak positions or peak widths and instead should be used to understand things like relative intensities and general, qualitative spectral features. The packages used to create the map,[102] calculate the electric field,[59] and evaluate the spectroscopic quantities[61] have been made available to the greater sci-

entific community. While in this work, we considered the vibrational spectroscopy of HOD in D_2O, similar approaches could be used to develop maps for HOD in H_2O, and pure H_2O (with the inclusion of coupling terms as is described elsewhere).[103] These tools are also readily extended to develop spectroscopic maps for water models in the framework of more advanced polarizable force fields,[104,105] which have been shown to be particularly important to the proper description of biological problems that implicate buried charges, dipoles and water molecules.[106,107]

SUPPORTING INFORMATION

The supporting information includes a derivation of the bond-polarizability model used for the vibrational frequency mapping approach, simulation details for the SPC/E simulations, calculations of missed isosbestic points for the TIP3P water model, and the 2DIR response functions.

ACKNOWLEDGMENTS

The authors would like to thank Prof. Ward H. Thompson from the University of Kansas, Dr. Santanu Roy from Oak Ridge National Laboratory, Prof. Damien Laage from École Normale Supériure, Prof. Carlos Baiz from UT Austin, Reilly Osadchey Brown from Boston University, and Giulia Frigero from Università degli Studi di Milano-Bicocca for useful discussions related to this manuscript. The authors would like to additionally thank Dr. Hasini Senanayake, Dr. Ashley Borkowski, and Prof. Ward Thompson for sharing their empirical mapping and spectroscopy codes, from which the newly developed software was adapted.[80,100] Bulk spectroscopy calculations and map-building was funded by the National Science Foundation under Grant No. CHE-2001611, the NSF Center for Sustainable Nanotechnology. Membrane simulations and spectroscopy was funded by the Army Research Office through grant number #W911NF-22-1-0073. This work used the Expanse Cluster at the San Diego Supercomputing Center through allocation CHE-230043 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by National Science Foundation grants #2138259, #2138286, #2138307, #2137603, and #2138296. The authors are also pleased to acknowledge that the computational work reported on in this paper was performed on the Shared Computing Cluster which is administered by Boston University's Research Computing Services.

REFERENCES

[1]J. C. Flanagan, M. L. Valentine, and C. R. Baiz, Acc. Chem. Res. **53**, 1860 (2020).
[2]K. Lorenz-Ochoa and C. Baiz, "Ultrafast Spectroscopy Reveals Slow Water Dynamics in Biocondensates," (2023).
[3]M. K. Petti, J. P. Lomont, M. Maj, and M. T. Zanni, J Phys Chem B **122**, 1771 (2018).
[4]R. Fritzsch, S. Hume, L. Minnes, M. J. Baker, G. A. Burley, and N. T. Hunt, Analyst **145**, 2014 (2020).
[5]K. R. Webb, K. A. Hess, A. Shmidt, K. D. Segner, and L. E. Buchanan, Biophys. J. **122**, 1491 (2023).
[6]H. Binder, Eur. Biophys. J. **36**, 265 (2007).
[7]C.-M. Saak, L. B. Dreier, K. Machel, M. Bonn, and E. H. G. Backus, Faraday Discuss. **249**, 317 (2024).
[8]M. C. Bellissent-Funel, A. Hassanali, M. Havenith, R. Henchman, P. Pohl, F. Sterpone, D. Van Der Spoel, Y. Xu, and A. E. Garcia, Chem. Rev. **116**, 7673 (2016).
[9]P. Ball, Chem. Rev. **108**, 74 (2008).
[10]H. Frauenfelder, G. Chen, J. Berendzen, P. W. Fenimore, H. Jansson, B. H. McMahon, I. R. Stroe, J. Swenson, and R. D. Young, Proc. Natl. Acad. Sci. U.S.A. **106**, 5129 (2009).
[11]Y. Levy and J. N. Onuchic, Annu. Rev. Biophys. **35**, 389 (2006).
[12]C. I. Lynch, S. Rao, and M. S. P. Sansom, Chem. Rev. **120**, 10298 (2020).
[13]A. Fersht, *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (W.H. Freeman and Company, 1999).
[14]R. Baron and J. A. McCammon, Annu. Rev. Phys. Chem. **64**, 151 (2013).
[15]B. Z. Zsido and C. Hetenyi, Curr. Opin. Struct. Biol. **67**, 1 (2021).
[16]M. L. Samways, R. D. Taylor, H. E. Bruce Macdonald, and J. W. Essex, Chem. Soc. Rev. **50**, 9104 (2021).
[17]B. Bagchi, *Water in Biological and Chemical Processes: From Structure and Dynamics to Function* (Cambridge University Press, New York, 2013).
[18]A. C. Fogarty and D. Laage, J. Phys. Chem. B **118**, 7715 (2014).
[19]P. Jungwirth, J. Phys. Chem. Lett. **6**, 2449 (2015).
[20]D. Laage, T. Elsaesser, and J. T. Hynes, Chem. Rev. **117**, 10694 (2017).
[21]F. Sterpone, G. Stirnemann, and D. Laage, J. Am. Chem. Soc. **134**, 4116 (2012).
[22]D. M. Leitner, C. Hyeon, and K. M. Reid, J. Chem. Phys. **152**, 240901 (2020).
[23]S. Roy, S. M. Gruenbaum, and J. L. Skinner, J. Chem. Phys. **141**, 22D505 (2014).
[24]S. Roy, S. M. Gruenbaum, and J. L. Skinner, J. Chem. Phys. **141**, 18C502 (2014).
[25]E. E. Fenn, D. B. Wong, and M. D. Fayer, J. Chem. Phys. **134**, 054512 (2011).
[26]R. Costard, C. Greve, N. Levinger, E. Nibbering, and T. Elsaesser, EPJ Web of Conferences **41**, 06003 (2013).
[27]J. A. Mondal, S. Nihonyanagi, S. Yamaguchi, and T. Tahara, J. Am. Chem. Soc. **134**, 7842 (2012).
[28]O. Kel, A. Tamimi, M. C. Thielges, and M. D. Fayer, J. Am. Chem. Soc. , 12 (2013).
[29]H. Ishikawa, I. J. Finkelstein, S. Kim, K. Kwak, J. K. Chung, K. Wakasugi, A. M. Massari, and M. D. Fayer, Proc. Natl. Acad. Sci. U.S.A. **104**, 16116 (2007).
[30]M. J. Ryan, L. Gao, F. I. Valiyaveetil, A. A. Kananenka, and M. T. Zanni, J. Am. Chem. Soc. **146**, 1543 (2024).
[31]M. L. McDermott, H. Vanselous, S. A. Corcelli, and P. B. Petersen, ACS Cent. Sci. **3**, 708 (2017).
[32]S. A. Corcelli, C. P. Lawrence, and J. L. Skinner, J. Chem. Phys. **120**, 8107 (2004).
[33]S. A. Corcelli and J. L. Skinner, J. Phys. Chem. A **109**, 6154 (2005).

[34]B. Auer, R. Kumar, J. R. Schmidt, and J. L. Skinner, Proc. Natl. Acad. Sci. U.S.A. **104**, 14215 (2007).
[35]B. M. Auer and J. L. Skinner, J. Chem. Phys. **129**, 214705 (2008).
[36]S. M. Gruenbaum, C. J. Tainter, L. Shi, Y. Ni, and J. L. Skinner, J. Chem. Theory Comput. **9**, 3109 (2013).
[37]E. Małolepsza and J. E. Straub, J. Phys. Chem. B **118**, 7848 (2014).
[38]L. Wang, C. T. Middleton, M. T. Zanni, and J. L. Skinner, J. Phys. Chem. B **115**, 3713 (2011).
[39]T. Takayama, T. Otosu, and S. Yamaguchi, J. Chem. Phys. **158**, 136101 (2023).
[40]Y. Jiang and L. Wang, J. Phys. Chem. B **123**, 5791 (2019).
[41]C. R. Baiz, B. Błasiak, J. Bredenbeck, M. Cho, J.-H. Choi, S. A. Corcelli, A. G. Dijkstra, C.-J. Feng, S. Garrett-Roe, N.-H. Ge, M. W. D. Hanson-Heine, J. D. Hirst, T. L. C. Jansen, K. Kwac, K. J. Kubarych, C. H. Londergan, H. Maekawa, M. Reppert, S. Saito, S. Roy, J. L. Skinner, G. Stock, J. E. Straub, M. C. Thielges, K. Tominaga, A. Tokmakoff, H. Torii, L. Wang, L. J. Webb, and M. T. Zanni, Chem. Rev. **120**, 7152 (2020).
[42]O. Berger, O. Edholm, and F. Jähnig, Biophys. J. **72**, 2002 (1997).
[43]P. Mark and L. Nilsson, J. Phys. Chem. A **105**, 9954 (2001).
[44]E. E. S. Ong and J.-L. Liow, Fluid Phase Equilib. **481**, 55 (2019).
[45]Z. A. Piskulich, D. Laage, and W. H. Thompson, J. Phys. Chem. A **125**, 9941 (2021).
[46]D. J. Price and C. L. Brooks, J. Chem. Phys. **121**, 10096 (2004).
[47]T. Ishiyama and A. Morita, J. Chem. Phys. **131**, 244714 (2009).
[48]T. Ishiyama, D. Terada, and A. Morita, J. Phys. Chem. Lett. **7**, 216 (2016).
[49]T. Ishiyama and A. Morita, Annu. Rev. Phys. Chem. **68**, 355 (2017).
[50]J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. Brooks, A. D. MacKerell, J. B. Klauda, and W. Im, J. Chem. Theory Comput. **12**, 405 (2016).
[51]S. Jo, T. Kim, V. Iyer, and W. Im, J. Comput. Chem. **29**, 1859 (2009).
[52]M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindah, SoftwareX **1–2**, 19 (2015).
[53]S. Nosé and M. Klein, Mol. Phys. **50**, 1055 (1983).
[54]W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).
[55]B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, J. Comput. Chem. **18**, 1463 (1997).
[56]S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).
[57]P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).
[58]A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[59]Z. A. Piskulich and Q. Cui, "Electric Field Cell List https://doi.org/10.5281/zenodo.13869139," (2024).
[60]W. Mattson and B. M. Rice, Comput. Phys. Commun. **119**, 135 (1999).
[61]Z. A. Piskulich, A. Borkowski, H. S. Senanayake, W. H. Thompson, and Q. Cui, "Generalized Water Spectroscopy Code https://doi.org/10.5281/zenodo.13868846," (2024).
[62]Y. Li, J. Liu, and J. C. Gumbart, Methods Mol. Biol. **2302**, 237 (2021).
[63]"CHARMM-GUI," https://www.charmm-gui.org/?doc=input/membrane.bilayer&step=6.
[64]A. A. Kananenka, K. Yao, S. A. Corcelli, and J. L. Skinner, J. Chem. Theory Comput. **15**, 6850 (2019).
[65]D. Konstantinovsky, E. A. Perets, T. Santiago, K. Olesen, Z. Wang, A. V. Soudackov, E. C. Yan, and S. Hammes-Schiffer, J. Phys. Chem. B **127**, 2418 (2023).
[66]D. T. Colbert and W. H. Miller, J. Chem. Phys. **96**, 1982 (1992).
[67]C. J. Fecko, J. J. Loparo, S. T. Roberts, and A. Tokmakoff, J. Chem. Phys. **122**, 054506 (2005).
[68]D. W. Oxtoby, D. Levesque, and J. J. Weis, J. Chem. Phys. **68**, 5528 (1978).
[69]D. Long, Proc. R. Soc. Lond. A **217**, 203 (1953).
[70]A. C. Belch and S. A. Rice, J. Chem. Phys. **78**, 4817 (1983).
[71]J. R. Scherer and R. G. Snyder, J Chem. Phys. **67**, 4794 (1977).
[72]J. E. Bertie, M. K. Ahmed, and H. H. Eysel, J. Phys. Chem. **93**, 2210 (1989).
[73]V. Buch, T. Tarbuck, G. L. Richmond, H. Groenzin, I. Li, and M. J. Shultz, J. Chem. Phys. **127**, 204710 (2007).
[74]H. Torii, J. Phys. Chem. A **110**, 9469 (2006).
[75]A. Rohatgi, "WebPlotDigitizer," (2024).
[76]Z. A. Piskulich and W. H. Thompson, J. Chem. Theor. Comput. **17**, 2659 (2021).
[77]J. D. Smith, C. D. Cappa, K. R. Wilson, R. C. Cohen, P. L. Geissler, and R. J. Saykally, Proc. Natl. Acad. Sci. U.S.A. **102**, 14171 (2005).
[78]J. R. Schmidt, S. A. Corcelli, and J. L. Skinner, J. Chem. Phys. **123**, 044513 (2005).
[79]S. M. Gruenbaum, P. A. Pieniazek, and J. L. Skinner, J. Chem. Phys. **135**, 164506 (2011).
[80]A. K. Borkowski, N. I. Campbell, and W. H. Thompson, J. Chem. Phys. **158**, 064507 (2023).
[81]V. Ostroverkhov, G. A. Waychunas, and Y. R. Shen, Phys. Rev. Lett. **94**, 046102 (2005).
[82]Y. R. Shen and V. Ostroverkhov, Chem. Rev. **106**, 1140 (2006).
[83]N. Ji, V. Ostroverkhov, C. S. Tian, and Y. R. Shen, Phys. Rev. Lett. **100**, 096102 (2008).
[84]Y. Shen, Annual Review of Physical Chemistry **64**, 129 (2013).
[85]S. Nihonyanagi, J. A. Mondal, S. Yamaguchi, and T. Tahara, Annu. Rev. Phys. Chem. **64**, 579 (2013).
[86]S. Yamaguchi and T. Otosu, Phys. Chem. Chem. Phys. **23**, 18253 (2021).
[87]S. J. Byrnes, P. L. Geissler, and Y. R. Shen, Chem. Phys. Lett. **516**, 115 (2011).
[88]S. Yamaguchi, T. Takayama, Y. Goto, T. Otosu, and T. Yagasaki, J. Phys. Chem. Lett. **13**, 9649 (2022).
[89]S. Sun, R. Liang, X. Xu, H. Zhu, Y. R. Shen, and C. Tian, J. Chem. Phys. **144**, 244711 (2016).
[90]M. de la Puente, A. Gomez, and D. Laage, J. Phys. Chem. Lett. **15**, 3096 (2024).
[91]J. Noah-Vanhoucke, J. D. Smith, and P. L. Geissler, J. Phys. Chem. B **113**, 4065 (2009).
[92]Y. Ni and J. L. Skinner, J. Chem. Phys. **145**, 031103 (2016).
[93]E. A. Raymond, T. L. Tarbuck, M. G. Brown, and G. L. Richmond, J. Phys. Chem. B **107**, 546 (2003).
[94]S. Nihonyanagi, R. Kusaka, K.-i. Inoue, A. Adhikari, S. Yamaguchi, and T. Tahara, J. Chem. Phys. **143**, 124707 (2015).
[95]W. J. Smit, J. Versluis, E. H. G. Backus, M. Bonn, and H. J. Bakker, J. Phys. Chem. Lett. **9**, 1290 (2018).
[96]K.-i. Inoue, M. Ahmed, S. Nihonyanagi, and T. Tahara, Nat. Commun. **11**, 5344 (2020).
[97]K. Niu, H.-f. Wang, and R. A. Marcus, Proc. Natl. Acad. Sci. U.S.A. **121**, e2402550121 (2024).
[98]B. M. Auer and J. L. Skinner, J. Chem. Phys. **129**, 214705 (2008).
[99]P. C. Burris, D. Laage, and W. H. Thompson, J. Chem. Phys. **144**, 194709 (2016).
[100]H. S. Senanayake, J. A. Greathouse, and W. H. Thompson, J. Chem. Phys. **157**, 034702 (2022).
[101]D. R. Nutt and J. C. Smith, J. Chem. Theory Comput. **3**, 1550 (2007).
[102]Z. A. Piskulich, W. H. Thompson, and Q. Cui, "Empirical Mapping Code - Water https://doi.org/10.5281/zenodo.13869000," (2024).
[103]L. Shi, S. M. Gruenbaum, and J. L. Skinner, J. Phys. Chem. B **116**, 13821 (2012).
[104]J. A. Lemkul, J. Huang, B. Roux, and A. D. MacKerell Jr., Chem. Rev. **116**, 4983 (2016).
[105]Z. Jing, C. Liu, S. Y. Cheng, R. Qi, B. D. Walker, J.-P. Piquemal, and P. Ren, Annu. Rev. Biophys. **48**, 371 (2019).

[106]J. Deng and Q. Cui, J. Am. Chem. Soc. **144**, 4594 (2022).

[107]J. Deng and Q. Cui, J. Chem. Theory Comput. **20**, 1897 (2024).