The Journal of Chemical Physics

RESEARCH ARTICLE | DECEMBER 17 2025

Machine learning potentials accurately reproduce vibrational dynamics in complex environments

Special Collection: Computational Spectroscopy

Chloe B. Starkey ; Saptarsi Mondal; Carlos R. Baiz

J. Chem. Phys. 163, 234116 (2025)
https://doi.org/10.1063/5.0306072

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

Machine learning potentials accurately reproduce vibrational dynamics in complex environments

Cite as: J. Chem. Phys. 163, 234116 (2025); doi: 10.1063/5.0306072
Submitted: 8 October 2025 • Accepted: 1 December 2025 •
Published Online: 17 December 2025

Chloe B. Starkey, Saptarsi Mondal, and Carlos R. Baiz[a)]

AFFILIATIONS

Department of Chemistry, University of Texas at Austin, 105 E 24th St. A5300, Austin, Texas 78712, USA

Note: This paper is part of the Special Topic on Computational Spectroscopy.
[a)]**Author to whom correspondence should be addressed:** cbaiz@cm.utexas.edu

ABSTRACT

Vibrational spectroscopy provides a bond-specific view of molecular structure and dynamics, but translating spectroscopic observables to local environments requires the development of accurate models to compute spectroscopic observables from simulations. Recent advances in machine learning interatomic potentials (MLIPs) provide *ab-initio*-like accuracy at low computational cost, opening new opportunities for developing general, transferable models that can predict spectroscopic observables without system-specific parameterization. Here, we benchmark the performance of the Universal Model of Atoms (UMA), a recently developed MLIP, for predicting IR absorption spectra and picosecond frequency fluctuations of an ester carbonyl in a range of solvents. UMA results are compared with two established approaches: an empirical frequency map parameterized for the ester carbonyl and the semiempirical tight-binding method, GFN2-xTB. We find that UMA reproduces experimental observables with accuracy comparable to traditional methods, while offering broader generality and efficiency.

I. INTRODUCTION

Vibrational frequency maps provide a quantitative link between molecular structure and spectroscopy by describing how the local environment influences a vibrational mode's frequency. In practice, a vibrational map is a parameterized function that converts instantaneous environments, such as electric fields or potentials, into predicted frequency shifts for a vibrational mode.[1–7] The conceptual foundation for this approach was established in early work connecting OH-stretch frequency fluctuations in water to instantaneous local electric fields and further by parameterizing an electrostatic map for HOD in liquid water to reproduce experimental line shapes and frequency–frequency correlation functions.[2,4] Together, these studies demonstrated that vibrational frequencies can be quantitatively predicted from molecular structure, laying the groundwork for direct translation of molecular dynamics (MD) simulations into vibrational observables that can be compared with ultrafast spectroscopic measurements. By connecting microscopic configurations to spectroscopic observables, vibrational maps offer a framework for connecting molecular interactions to dynamic frequency fluctuations and spectral line shapes.[8]

Two-dimensional infrared (2D IR) spectroscopy directly probes frequency fluctuations by creating a correlation map between excitation and detection frequencies.[9] Conceptually, 2D IR captures correlations between frequencies at different time delays, revealing how quickly a vibrational probe loses memory of its initial environment. The resulting frequency–frequency correlation function (FFCF) defines the time scales of solvent fluctuations, which can be compared directly with MD simulations using vibrational maps.[10–12] As such, the combination of 2D IR spectroscopy and vibrational mapping enables a unified, structure-based understanding of vibrational dynamics in complex liquids, biomolecules, and interfaces.[13]

An effective probe for benchmarking this combined framework is the ester carbonyl stretch, a mode that appears in esters, lipids, and peptides and is dominated by electrostatic interactions with its surrounding environment.[14,15] As a hydrogen-bond acceptor incapable of self-association, it provides a clean and interpretable measure of solvent effects and hydrogen-bond populations. Edington and coworkers developed an empirical frequency map for this mode based on six electrostatic field components at the O–C=O atomic sites, showing strong agreement with experimental center frequencies.[3]

Building on this foundation, the present study further evaluates the accuracy of such maps in predicting not only center frequencies, but also correlation times and associated dynamics.

While the vibrational frequency maps have demonstrated excellent agreement with experiment, their predictive accuracy depends on mode-specific parameterization. Constructing a reliable map typically requires correlation of local electrostatic or structural descriptors with vibrational frequencies obtained from either *ab initio* calculations or experimental data.[3,16,17] These approaches produce accurate models for specific vibrational modes but require significant effort to develop and are not easily transferable across different chemical environments. A more recent approach leverages semiempirical methods designed to approximate Density Functional Theory (DFT)-level accuracy at far lower computational cost by applying a tight-binding method, GFN2-xTB, specifically parameterized for vibrational frequency calculations.[18] In parallel, Kuroda and co-workers introduced the Instantaneous Frequency of Molecules (IFM) method, which applies GFN2-xTB to solutes with a minimal solvation shell to compute instantaneous vibrational frequencies.[19] IFM has been successfully validated against experiment for the amide I mode in N-methylacetamide, showing linear agreement with experimental center frequencies and correlation times. This method has become a useful tool for connecting simulation-derived frequency fluctuations with experimental observables. Here, we test whether GFN2-xTB combined with IFM provides a similarly accurate description of the ester carbonyl stretch.

On the frontier of computational methods, machine learning (ML) has emerged as a powerful alternative for obtaining *ab-initio*-like accuracy at much reduced computational expense. ML-based interatomic potentials (MLIPs) represent a rapidly advancing frontier in computational chemistry, which has the potential to largely replace expensive electronic-structure methods as the reduced computational costs allow for broader applications, such as on-the-fly computation of vibrational frequencies and molecular simulations. MLIPs are trained on massive and chemically diverse datasets, granting them wide transferability across molecular systems. Their near *ab initio* accuracy has already transformed applications in molecular dynamics, catalysis, and materials discovery.[20–22] For vibrational spectroscopy, MLIPs hold promise as a general-purpose model for predicting frequency shifts and dynamical observations across complex condensed-phase environments.

Recently, Meta FAIR introduced the Universal Model for Atoms (UMA), a suite of machine learning interatomic potential (MLIP) models trained on an unprecedented ~500 × 10^6 atomic structures to predict a wide range of chemical properties, including vibrational frequencies.[23] UMA achieves near-DFT accuracy with linear scaling (n), enabling rapid simulations at scales inaccessible to conventional electronic structure methods, including GFN2-xTB, which scales as (n^2). In this work, we apply the IFM methodology with the UMA-small (UMA-sm) model to compute instantaneous frequencies and frequency-fluctuation correlation functions from molecular dynamics trajectories. The results are compared to experimentally measured FTIR spectra and frequency-fluctuation correlation functions extracted from measured 2D IR spectra.

Ethyl acetate (EtOAc) is chosen, as its ester carbonyl group serves as a relatively simple probe, minimally perturbed by intramolecular interactions and unable to hydrogen-bond with itself. Furthermore, its solubility across a broad range of solvents makes it ideally suited for both experimental and computational studies of condensed-phase vibrational solvatochromism.[24,25] To our knowledge, this represents the first application of UMA to vibrational solvatochromism, providing a new framework for rapid, experimentally comparable predictions.

II. METHODS

A. FTIR spectroscopy

Fourier Transform Infrared (FTIR) spectra of ethyl acetate in a range of solvents have been published previously. Here, we use the spectra as published.[3] Briefly, in the previous study, ethyl acetate (EtOAc) solutions were prepared at 20 mg/ml concentration (0.23M) in D_2O, methanol (MeOH), ethanol (EtOH), *n*-butanol (ButOH), *n*-hexanol (HexOH), diethylene glycol (DEG), dimethyl sulfoxide (DMSO), tetrahydrofuran (THF), and acetonitrile (ACN). For measurement, 50–100 ml of each solution was placed between two CaF_2 windows with a 50 mm PTFE spacer. Spectra were collected at 2 cm^{-1} resolution using a Bruker Vertex 70 FTIR spectrometer.

B. Two-dimensional infrared spectroscopy

The detailed experimental design of the current setup was reported earlier.[26] In brief, mid-IR pulses of ~100 fs duration and ~1 kHz repetition rate centered at 5.8 μm were generated via a combination of optical parametric amplification using type II BBO crystals and difference frequency generation using $AgGaS_2$. A standard pump–probe geometry was employed to acquire 2D IR spectra. Two pump pulses arranged in a collinear pump sequence with a controlled delay generated by a pulse shaper (PhaseTech Spectroscopy) excited the sample, and a weaker, time-delayed probe pulse sampled the induced vibrational coherence. The absorptive 2D IR signal was detected on a dispersive 128 × 128 element mid-IR MCT array. By scanning the pump–probe delay (population waiting time t_2) from 0.1 ps up to 3 ps and Fourier-transforming the signal along the coherence time, we obtained a series of 2D spectra (ω_1, ω_2; t_2). The 0–1 and 1–2 transitions of the C=O stretch mode appear as a pair of lobes (absorptive line shape) along the diagonal. 2D IR spectra were recorded in the same solvent compositions as the FTIR spectra using the same sample cell and spacers.

Experimental spectra are included in Sec. S1. The spectral diffusion was quantified via the CLS method.[27] In brief, for each waiting time, t_2, a "centerline" was obtained by tracing the maximum of the 2D peak in $S(\omega_1, \omega_2; t_2)$ along the detection axis ω_3 (Fig. 1). The slope of this line provides a normalized frequency–frequency correlation function $C(t)$, which reflects the extent of frequency memory retention. The CLS values (Sec. S2) were plotted as a function of t_2 and fit to a single-exponential decay function to obtain the spectral diffusion/frequency–frequency correlation timescale, τ, as follows:

$$C(t) = Ae^{t/\tau}.$$

C. Molecular dynamics simulation

Molecular dynamics (MD) simulations were performed using the CHARMM36m force field (with TIP3P water) with

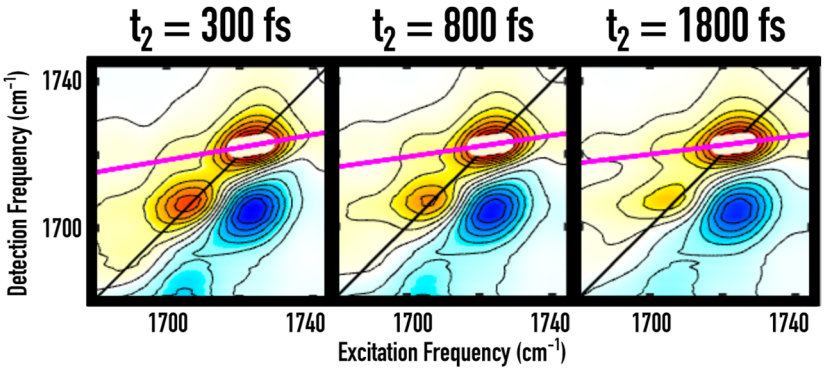

FIG. 1. Example 2D IR spectra of EtOAc in EtOH. Three different waiting times, t_2, are shown. The centerline is shown in magenta, and the centerline slope (CLS) analysis was performed at frequencies marked in white. The slope of this line determines the CLS value for a specific t_2 value.

topologies generated by CHARMM-GUI and executed in GROMACS 2024.[28–30] Each system contained a single EtOAc molecule constrained at the center of a 4 nm cubic box, solvated with solvent molecules to match solvent densities at room temperature. The initial configurations were generated using PACKMOL.[31] Energy minimization was performed using the steepest descent algorithm, followed by 25 ps of NVT equilibration with a velocity-rescaling thermostat and 1 ns of NPT equilibration with a Nose–Hoover thermostat and Parrinello–Rahman barostat.[32,33] All bonds were constrained with the LINCS algorithm.[34] Production simulations were run for 1 ns in the NPT ensemble at 303.15 K and 1 atm with a 2 fs time step. Snapshots were saved every 50 fs for subsequent frequency calculations.

D. Vibrational map for ester carbonyl

Instantaneous frequencies were computed using a vibrational spectroscopic map for the ester carbonyl developed by Baiz and co-workers.[3] The map correlates oscillator frequency with the surrounding electrostatic environment via six electric field parameters centered on the O–C=O atoms, optimized by comparison of FTIR spectra with MD simulations. Electrostatic calculations included all atoms in the 4 nm simulation box except non-parameterized atoms within EtOAc.

E. GFN2-xTB model and the instantaneous frequencies of molecules method

The GFN2-xTB semi-empirical tight-binding model was used in conjunction with the IFM method.[18,19] For each MD snapshot, the first solvation shell around EtOAc was extracted using MDAnalysis.[35,36] The solvation shell was defined based on the radial distribution function (RDF) between the solute and solvent, applying a cutoff corresponding to the first minimum of the RDF. Depending on the solvent, the first shell contained ~6–12 solvent molecules. EtOAc geometry was optimized while constraining the solvent molecules, and the ester carbonyl frequency was calculated at the GFN2-xTB level of theory.

F. Universal model for atoms (UMA)

The UMA family of machine learning interatomic potentials (MLIPs) developed by Meta FAIR was used to calculate ester carbonyl frequencies.[23] For each MD snapshot, the first solvation shell around EtOAc was extracted, geometry was optimized with no constraints, and the numerical Hessian was calculated by finite differences of forces using the uma-s-1p1 model as implemented in the atomic simulation environment package.[37]

G. Spectral calculation

Spectra were calculated for all three models using the following optical response function:[38]

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

where

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

Angled brackets represent an ensemble average where a 1 ns trajectory is divided into 500 circularly shifted replicas of the frequency trajectory. The absorption spectrum, $A(\omega)$, is computed as the imaginary part of a rotating-frame single-sided Fourier transform of the optical response function, $U(\tau)$. Homogeneous broadening due to population relaxation was included through an exponential decay term with time constant, T_1, which was set equal to the vibrational lifetime extracted from analysis of the corresponding experimental 2D IR spectra.

III. RESULTS AND DISCUSSION

Across all approaches, predicted center frequencies show a linear relationship with the experiment (Fig. 2). UMA demonstrates strong solvent-dependent trends comparable to both the frequency map and GFN2-xTB, establishing MLIPs as reliable tools for connecting simulations to experimental observables. Similar to many DFT-based methods, UMA tends to systematically overestimate absolute frequencies, but applying a uniform scaling factor (0.978) yields near one-to-one agreement with experiment. After correction, UMA achieves a correlation with experiment comparable to GFN2-xTB (R^2 = 0.75 and 0.74, respectively).

Beyond absolute frequencies, UMA reproduces hydrogen-bond (HB)-dependent shifts with clear separation between 0, 1, and 2 HB ensembles. For the 0 and 1 HB populations, UMA outperforms GFN2-xTB, capturing frequency differences with improved accuracy. Deviations arise for the 2 HB ensemble, where UMA slightly

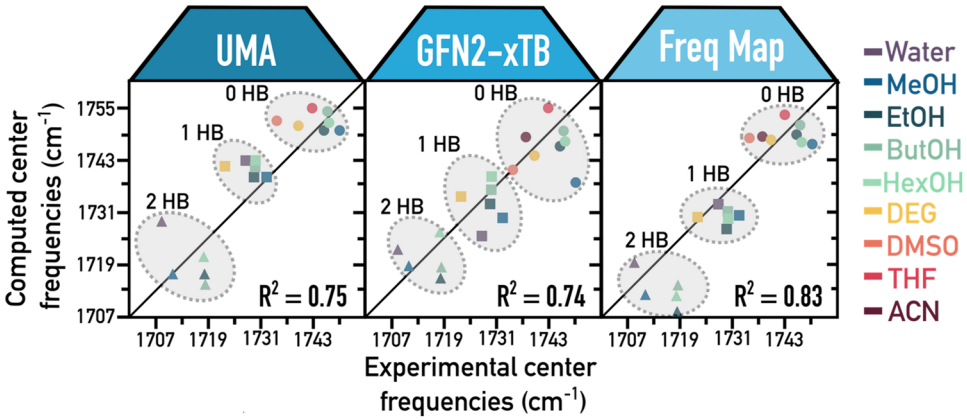

FIG. 2. Comparison between computed and experimental ester carbonyl center frequencies for ethyl acetate in eight selected solvents. HB populations are grouped for clarity. The solvents are labeled as follows: MeOH = methanol, EtOH = ethanol, ButOH = butanol, DEG = diethylene glycol, DMSO = dimethyl sulfoxide, THF = tetrahydrofuran, and ACN = acetonitrile.

overestimates the 1 HB → 2 HB redshift, suggesting challenges in describing multiple hydrogen bond contributions simultaneously. In contrast, GFN2-xTB better captures these strongly solvated environments, although it fails to clearly separate water HB populations. The frequency map remains the most accurate overall, yielding the highest R^2 and well-defined HB-dependent shifts of ~16 to 20 cm^{-1}, consistent with prior work.[3]

One important exception is acetonitrile: after scaling, UMA predicts a substantially lower center frequency than observed experimentally. This discrepancy may reflect the underrepresentation of nitrile systems in the ML training set.[39] However, this limitation does not compromise UMA's ability to reproduce solvent-dependent shifts or dynamic observables, which are less sensitive to absolute frequency. Taken together, these results suggest that (1) center frequency offsets have limited impact on predicted dynamics, and (2) UMA remains a robust and transferable approach across solvents, with acetonitrile highlighting an area for further refinement.

The correlation time (τ_c) reflects the timescale of the fluctuations of the environment around a vibrational probe. A short τ_c indicates rapid fluctuations, where the probe loses "memory" of its initial state quickly, while a long τ_c corresponds to more persistent, longer-lived interactions with the environment. Figure 3 compares the τ_c values predicted from the three computational approaches against experimental benchmarks obtained from 2D IR centerline slope (CLS) decays. All three methods reproduce the solvent-dependent trends in dynamics with good overall agreement with the experiment ($R^2 > 0.70$). Among them, UMA performs comparably to both the empirical frequency map and the semiempirical GFN2-xTB method, establishing MLIPs as a viable new option for capturing vibrational dynamics. While the frequency map yields slightly better quantitative agreement overall, UMA demonstrates competitive accuracy without requiring probe-specific parameterization. This result highlights the promise of MLIPs as general-purpose tools for spectroscopic applications.

In several cases, UMA provides unique advantages. For acetonitrile, where GFN2-xTB has been shown to perform poorly due to limitations in its parameterization, UMA reproduces the experimental correlation time with near-perfect accuracy (0.23 vs 0.22 ps).[40] This suggests that MLIPs may overcome specific weaknesses of semiempirical approaches in chemically diverse solvents. Conversely, for solvents such as DEG and DMSO, the frequency map shows larger deviations, consistent with previous observations that electrostatic field-based maps can be less reliable when force-field charges are underestimated.[3] UMA and GFN2-xTB instead yield more accurate correlation times across these systems when compared to the frequency map.

As with previous studies, correlation times for protic solvents remain more challenging for all approaches.[3,19] Large

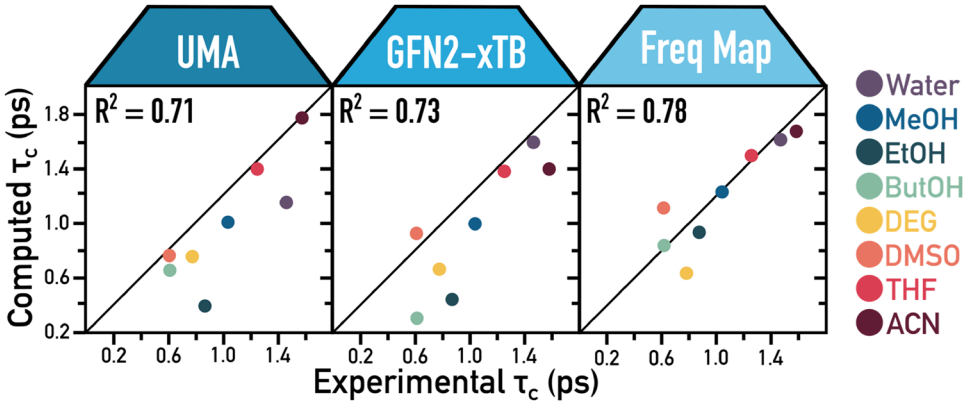

FIG. 3. Comparison between computed and experimental correlation times for ethyl acetate in eight selected solvents. HB populations are grouped for clarity. The solvents are labeled as follows: MeOH = methanol, EtOH = ethanol, ButOH = butanol, DEG = diethylene glycol, DMSO = dimethyl sulfoxide, THF = tetrahydrofuran, and ACN = acetonitrile.

hydrogen-bond fluctuations in protic environments drive substantial frequency shifts, which are more difficult to capture accurately. UMA, such as GFN2-xTB, shows a slight tendency to overestimate τ_c values in these cases, suggesting both methods somewhat exaggerate solvent–solute interactions. Longer simulations may mitigate this limitation by improving sampling of hydrogen-bond dynamics. For EtOAc in HexOH, all models substantially underestimate the experimental correlation time of 2.9 ps, and this case was excluded from statistical comparison. The discrepancy likely reflects microheterogeneous solvent environments that are poorly sampled in MD.[3,41,42] As the largest alcohol studied here, HexOH may also present challenges for all methods in accurately reproducing slower, long-range interactions with the probe.

Altogether, these results demonstrate that UMA provides solvent-dependent vibrational dynamics in close agreement with experiment and performs on par with established computational strategies. Importantly, UMA achieves this without system-specific tuning, underscoring the potential of MLIPs as broadly applicable and efficient resources for vibrational spectroscopy.

Accurate reproduction of vibrational observables depends not only on the frequency prediction method but also on realistic sampling of hydrogen-bond (HB) ensembles in the underlying MD simulations used to generate snapshots.[1] Differences in HB populations can strongly influence both center frequencies and correlation times, as both depend directly on the ensemble composition. Figure 4 compares the HB populations obtained from 1 ns MD trajectories with experimental populations derived from Lorentzian fits to FTIR spectra corrected by previously reported oscillator strengths for each HB ensemble.[43] The largest discrepancy occurs for water, where MD underestimates the 2 HB population and overestimates 0 HB, a limitation arising from the CHARMM force field rather than the vibrational models themselves. Given that HB populations are not often used as benchmarks in the development of force-fields, these discrepancies are often expected.[6,44,45] These discrepancies emphasize the necessity of accurately representing HB populations to model frequencies. Even with advanced vibrational models such as GFN2-xTB or UMA, deviations in HB populations will propagate into inaccuracies in the line shapes and potentially the frequency fluctuations, and the populations are different. Earlier work has shown that improved parameterizations and consideration of effects such as Fermi resonances can enhance agreement with experiment, but these challenges highlight that the treatment of hydrogen bonding remains a central limitation across all methods.[46,47]

Figure 5 shows a comparison between calculated and experimental spectra for water, methanol, and THF across the three computational methods. For protic solvents, spectra were generated by weighing each HB ensemble according to FTIR-derived population (see Fig. 4). Among the methods, the frequency map provides the best overall agreement with experiment, which is expected given its parameterization against FTIR data. UMA predicts larger frequency shifts in water, resulting in a clear, bimodal feature, while GFN2-xTB compresses the frequency range of the ensembles, producing a more Gaussian-like peak. The frequency map balances these extremes by resolving distinct ensemble contributions while still reproducing the broadened profile observed experimentally. In the aprotic case of THF, UMA and GFN2-xTB perform similarly, capturing the reduced broadening but with a slight redshift relative to experiment. The frequency map more accurately reproduces the center frequency but slightly overestimates the width. Overall, all three approaches can capture solvent dependent spectral trends, but their relative strengths and weaknesses differ depending on hydrogen-bonding character. Complete spectra for all solvents are provided in Sec. S3.

From a computational standpoint, UMA requires a geometry optimization and a full Hessian for each snapshot, making it more computationally demanding than the empirical frequency map but competitive with semiempirical electronic-structure based methods. Generally, the cost of using general methods is substantially higher than frequency maps, which rely on a simple Coulomb sum over atoms. However, the main advantage of UMA is its parameterization-free application to a wide range of vibrational

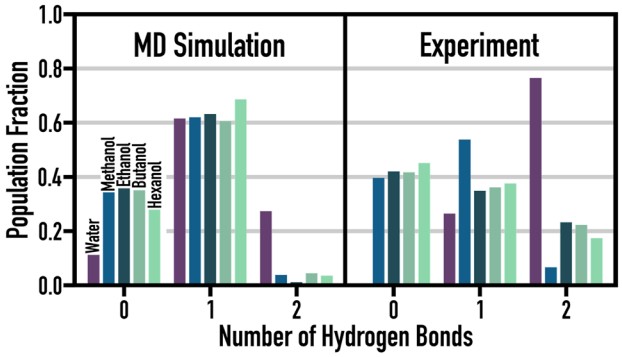

FIG. 4. Relative hydrogen bond populations from CHARMM-based MD simulations (left) and from integrated IR absorption peaks using previously published spectra (right).

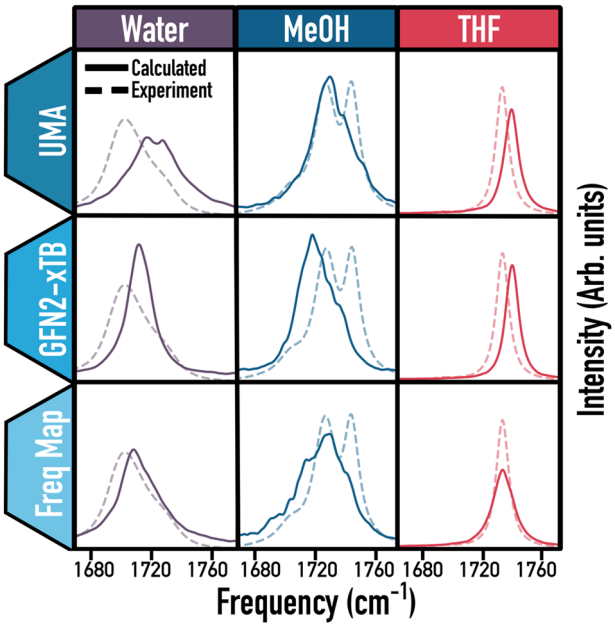

FIG. 5. Calculated (solid) and experimental (dashed) linear absorption spectra of ethyl acetate in water, methanol (MeOH), and tetrahydrofuran (THF) across all three methods.

probes. Its favorable scaling relative to DFT and its applicability across diverse chemical environments make it especially well-suited for more complex and heterogeneous environments, such as lipid membranes or electrochemical interfaces, where parameterized maps may be less reliable, while maps may continue to perform well for applications such as computing the amide I modes in proteins, which require a large number of oscillators but with more consistent environments that are well captured by the map.[13,48]

In this study, UMA performs comparably to GFN2-xTB and, in several cases, better reproduces experimental solvent-dependent shifts. Importantly, UMA achieves this without probe-specific parameterization, highlighting the potential of MLIPs as a new paradigm for vibrational spectroscopy. While we employed the UMA-small architecture here as a proof of principle, larger variants such as UMA-medium may provide improved accuracy for more complex applications.[23]

Overall, our results suggest complementary roles for the three methods. The frequency map remains the most efficient, offering reliable and inexpensive predictions for well-parameterized probes. GFN2-xTB balances physical descriptions of atoms and electrons with moderate computational cost, reproducing many experimental features across a range of solvents. UMA, by contrast, represents a forward-looking solution: a general, transferable, and systematically improvable framework that can extend vibrational spectroscopy modeling to increasingly complex environments.

IV. CONCLUSIONS AND OUTLOOK

This study compared three computational approaches—an empirical frequency map, the semiempirical method GFN2-xTB, and the machine learning potential UMA—for predicting vibrational frequency shifts and solvent dynamics of ethyl acetate in diverse solvents. Although each method differs in its underlying principles, all reproduce ensemble-averaged observables and solvent-dependent correlation times with good agreement with experiment.

The frequency map remains the most efficient strategy, providing accurate probe-specific predictions once parameterized. GFN2-xTB offers a more rigorous quantum-chemical foundation at moderate cost but shows limitations for certain solvent classes. UMA achieves accuracy comparable to both methods without system-specific parameterization, demonstrating the capability of machine learning interatomic potentials to serve as general and transferable models.

Looking ahead, machine learning potentials represent an important step forward for the field. Their scalability, efficiency, and transferability open the door to modeling increasingly complex chemical environments, including mixed solvents and biological interfaces. By establishing UMA's performance alongside established methods, this work highlights the potential of machine learning to reshape how vibrational observables are predicted and integrated with experiment.

SUPPLEMENTARY MATERIAL

The supplementary material includes the following: (1) Experimental 2D IR spectra, (2) CLS decays and exponential fits, (3) computed frequency fluctuation correlation functions and fits, (4) experimental and computed FTIR spectra, (5) computed frequency distributions, (6) additional comparisons between UMA models, and (7) radial distribution functions.

ACKNOWLEDGMENTS

This project was funded by the National Institutes of Health (Grant No. R35GM133359) as well as the Welch Foundation (Grant No. F-1891). Simulations were carried out at the Texas Advanced Computing Center (TACC).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Chloe B. Starkey: Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Software (equal); Writing – original draft (equal); Writing – review & editing (equal). **Saptarsi Mondal**: Data curation (supporting); Formal analysis (supporting); Writing – original draft (supporting). **Carlos R. Baiz**: Conceptualization (lead); Data curation (supporting); Investigation (supporting); Methodology (equal); Resources (lead); Software (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material.

REFERENCES

[1] 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, "Vibrational spectroscopic map, vibrational spectroscopy, and intermolecular interaction," Chem. Rev. **120**(15), 7152–7218 (2020).

[2] T. La Cour Jansen and J. Knoester, "A transferable electrostatic map for solvation effects on amide I vibrations and its application to linear and two-dimensional spectroscopy," J. Chem. Phys. **124**(4), 044502 (2006).

[3] S. C. Edington, J. C. Flanagan, and C. R. Baiz, "An empirical IR frequency map for ester C=O stretching vibrations," J. Phys. Chem. A **120**(22), 3888–3896 (2016).

[4] S. A. Corcelli, C. P. Lawrence, and J. L. Skinner, "Combined electronic structure/molecular dynamics approach for ultrafast infrared spectroscopy of dilute HOD in liquid H_2O and D_2O," J. Chem. Phys. **120**(17), 8107–8117 (2004).

[5] A. A. Kananenka, K. Yao, S. A. Corcelli, and J. L. Skinner, "Machine learning for vibrational spectroscopic maps," J. Chem. Theory Comput. **15**(12), 6850–6858 (2019).

[6] L. Wang, C. T. Middleton, M. T. Zanni, and J. L. Skinner, "Development and validation of transferable amide I vibrational frequency maps for peptides," J. Phys. Chem. B **115**(13), 3713–3724 (2011).

[7]J.-H. Choi, K.-I. Oh, H. Lee, C. Lee, and M. Cho, "Nitrile and thiocyanate IR probes: Quantum chemistry calculation studies and multivariate least-square fitting analysis," J. Chem. Phys. **128**(13), 134506 (2008).
[8]Z. A. Piskulich and W. H. Thompson, "Temperature dependence of the water infrared spectrum: Driving forces, isosbestic points, and predictions," J. Phys. Chem. Lett. **11**(18), 7762–7768 (2020).
[9]P. Hamm and M. Zanni, *Concepts and Methods of 2D Infrared Spectroscopy* (Cambridge University Press, 2011).
[10]T. L. C. Jansen, "Computational spectroscopy of complex systems," J. Chem. Phys. **155**(17), 170901 (2021).
[11]C. R. Baiz, K. J. Kubarych, E. Geva, and E. L. Sibert, "Local-mode approach to modeling multidimensional infrared spectra of metal carbonyls," J. Phys. Chem. A **115**(21), 5354–5363 (2011).
[12]C. A. Daly, E. J. Berquist, T. Brinzer, S. Garrett-Roe, D. S. Lambrecht, and S. A. Corcelli, "Modeling carbon dioxide vibrational frequencies in ionic liquids: II. Spectroscopic map," J. Phys. Chem. B **120**(49), 12633–12642 (2016).
[13]M. Reppert and A. Tokmakoff, "Computational amide I 2D IR spectroscopy as a probe of protein structure and dynamics," Annu. Rev. Phys. Chem. **67**(1), 359–386 (2016).
[14]I. M. Pazos, A. Ghosh, M. J. Tucker, and F. Gai, "Ester carbonyl vibration as a sensitive probe of protein local electric field," Angew. Chem., Int. Ed. **53**(24), 6080–6084 (2014).
[15]A. C. Guerin, K. Riley, K. Rupnik, and D. G. Kuroda, "Determining the energetics of the hydrogen bond through FTIR: A hands-on physical chemistry lab experiment," J. Chem. Educ. **93**(6), 1124–1129 (2016).
[16]T. Hayashi, T. La Cour Jansen, W. Zhuang, and S. Mukamel, "Collective solvent coordinates for the infrared spectrum of HOD in D_2O based on an ab initio electrostatic map," J. Phys. Chem. A **109**(1), 64–82 (2005).
[17]M. Reppert and A. Tokmakoff, "Communication: Quantitative multi-site frequency maps for amide I vibrational spectroscopy," J. Chem. Phys. **143**(6), 061102 (2015).
[18]C. Bannwarth, S. Ehlert, and S. Grimme, "GFN2-xTB—An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions," J. Chem. Theory Comput. **15**(3), 1652–1671 (2019).
[19]L. X. Sepulveda-Montaño, J. F. Galindo, and D. G. Kuroda, "A new computational methodology for the characterization of complex molecular environments using IR spectroscopy: Bridging the gap between experiments and computations," Chem. Sci. **15**(35), 14440–14448 (2024).
[20]S. Zhang, M. Z. Makoś, R. B. Jadrich, E. Kraka, K. Barros, B. T. Nebgen, S. Tretiak, O. Isayev, N. Lubbers, R. A. Messerly, and J. S. Smith, "Exploring the frontiers of condensed-phase chemistry with a general reactive machine learning potential," Nat. Chem. **16**(5), 727–734 (2024).
[21]D. M. Anstine and O. Isayev, "Machine learning interatomic potentials and long-range physics," J. Phys. Chem. A **127**(11), 2417–2431 (2023).
[22]E. C.-Y. Yuan, Y. Liu, J. Chen, P. Zhong, S. Raja, T. Kreiman, S. Vargas, W. Xu, M. Head-Gordon, C. Yang, S. M. Blau, B. Cheng, A. Krishnapriyan, and T. Head-Gordon, "Foundation models for atomistic simulation of chemistry and materials," arXiv:2503.10538 (2025).
[23]B. M. Wood, M. Dzamba, X. Fu, M. Gao, M. Shuaibi, K. Abdelmaqsoud, V. Gharakhanyan, J. R. Kitchin, D. S. Levine, A. Sriram, T. Cohen, A. Das, A. Rizvi, S. J. Sahoo, W. Ulissi, and C. L. Zitnick, "UMA: A family of universal models for atoms," arXiv:2506.23971 (2025).
[24]L. De Marco, M. Thämer, M. Reppert, and A. Tokmakoff, "Direct observation of intermolecular interactions mediated by hydrogen bonding," J. Chem. Phys. **141**(3), 034502 (2014).
[25]L. Chuntonov, I. M. Pazos, J. Ma, and F. Gai, "Kinetics of exchange between zero-one-and two-hydrogen-bonded states of methyl and ethyl acetate in methanol," J. Phys. Chem. B **119**(12), 4512–4520 (2015).
[26]S. C. Edington, A. Gonzalez, T. R. Middendorf, D. B. Halling, R. W. Aldrich, and C. R. Baiz, "Coordination to lanthanide ions distorts binding site conformation in calmodulin," Proc. Natl. Acad. Sci. U. S. A. **115**(14), E3126 (2018).
[27]Q. Guo, P. Pagano, Y.-L. Li, A. Kohen, and C. M. Cheatum, "Line shape analysis of two-dimensional infrared spectra," J. Chem. Phys. **142**(21), 212427 (2015).
[28]K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. Mackerell, "CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields," J. Comput. Chem. **31**(4), 671–690 (2010).
[29]S. Jo, T. Kim, V. G. Iyer, and W. Im, "CHARMM-GUI: A web-based graphical user interface for CHARMM," J. Comput. Chem. **29**(11), 1859–1865 (2008).
[30]M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," SoftwareX **1–2**, 19–25 (2015).
[31]L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez, "PACKMOL: A package for building initial configurations for molecular dynamics simulations," J. Comput. Chem. **30**(13), 2157–2164 (2009).
[32]D. J. Evans and B. L. Holian, "The Nose–Hoover thermostat," J. Chem. Phys. **83**(8), 4069–4074 (1985).
[33]M. Parrinello and A. Rahman, "Polymorphic transitions in single crystals: A new molecular dynamics method," J. Appl. Phys. **52**(12), 7182–7190 (1981).
[34]B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, "LINCS: A linear constraint solver for molecular simulations," J. Comput. Chem. **18**(12), 1463–1472 (1997).
[35]R. Gowers, M. Linke, J. Barnoud, T. Reddy, M. Melo, S. Seyler, J. Domański, D. Dotson, S. Buchoux, I. Kenney, and O. Beckstein, "MDAnalysis: A python package for the rapid analysis of molecular dynamics simulations," in *Proceedings of the 15th Python in Science Conference* (SciPy, Austin, TX, 2016), pp. 98–105.
[36]N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, and O. Beckstein, "MDAnalysis: A toolkit for the analysis of molecular dynamics simulations," J. Comput. Chem. **32**(10), 2319–2327 (2011).
[37]A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, "The atomic simulation Environment—A python library for working with atoms," J. Phys.: Condens. Matter **29**(27), 273002 (2017).
[38]M. D. Stephens, J. G. Saven, and J. L. Skinner, "Molecular theory of electronic spectroscopy in nonpolar fluids: Ultrafast solvation dynamics and absorption and emission line shapes," J. Chem. Phys. **106**(6), 2129–2144 (1997).
[39]D. S. Levine, M. Shuaibi, E. W. C. Spotte-Smith, M. G. Taylor, M. R. Hasyim, K. Michel, I. Batatia, G. Csányi, M. Dzamba, P. Eastman, N. C. Frey, X. Fu, V. Gharakhanyan, A. S. Krishnapriyan, J. A. Rackers, S. Raja, A. Rizvi, A. S. Rosen, Z. Ulissi, S. Vargas, C. L. Zitnick, S. M. Blau, and B. M. Wood, "The open molecules 2025 (OMol25) dataset, evaluations, and models," arXiv:2505.08762 (2025).
[40]L. X. Sepulveda-Montaño, J. F. Galindo, and D. G. Kuroda, "Infrared spectroscopy of liquid solutions as a benchmarking tool of semiempirical QM methods: The case of GFN2-xTB," J. Phys. Chem. B **127**(37), 7955–7963 (2023).
[41]F. G. Sanchez, A. N. Díaz, and J. L. Ramírez, "Environmentally shifted fluorescence lifetimes for multicomponent phase-modulation fluorimetric analysis," J. Fluoresc. **4**(1), 11–16 (1994).
[42]K.-I. Oh and C. R. Baiz, "Molecular heterogeneity in aqueous cosolvent systems," J. Chem. Phys. **152**(19), 190901 (2020).
[43]M. L. Valentine, M. K. Waterland, A. Fathizadeh, R. Elber, and C. R. Baiz, "Interfacial dynamics in lipid membranes: The effects of headgroup structures," J. Phys. Chem. B **125**(5), 1343–1350 (2021).
[44]A. D. MacKerell, N. Banavali, and N. Foloppe, "Development and current status of the CHARMM force field for nucleic acids," Biopolymers **56**(4), 257–265 (2000).
[45]J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, "Development and testing of a general amber force field," J. Comput. Chem. **25**(9), 1157–1174 (2004).
[46]B. Fang, T. Wang, X. Chen, T. Jin, R. Zhang, and W. Zhuang, "Modeling vibrational spectra of ester carbonyl stretch in water and DMSO based on molecular dynamics simulation," J. Phys. Chem. B **119**(38), 12390–12396 (2015).
[47]A. Ghosh, B. Cohn, A. K. Prasad, and L. Chuntonov, "Quantifying conformations of ester vibrational probes with hydrogen-bond-induced fermi resonances," J. Chem. Phys. **149**(18), 184501 (2018).
[48]K. E. Van Adrichem and T. L. C. Jansen, "AIM: A mapping program for infrared spectroscopy of proteins," J. Chem. Theory Comput. **18**(5), 3089–3098 (2022).