ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

Check for updates

A new computational methodology for the characterization of complex molecular environments using IR spectroscopy: bridging the gap between experiments and computations†

Laura X. Sepulveda-Montaño, [a] Johan F. Galindo [b] and Daniel G. Kuroda ⋆[a]

The molecular interactions and dynamics of complex liquid solutions are now routinely measured using IR and 2DIR spectroscopy. In particular, the use of the latter allows the determination of the frequency fluctuation correlation function (FFCF), while the former provides us with the average frequency. In turn, the FFCF can be used to quantify the vibrational dynamics of a molecule in a solution, and the center frequency provides details about the chemical environment, solvatochromism, of the vibrational mode. In simple solutions, the IR methodology can be used to unambiguously assign the interactions and dynamics observed by a molecule in solution. However, in complex environments with molecular heterogeneities, this assignment is not simple. Therefore, a method that allows for such an assignment is essential. Here, a parametrization free method, called Instantaneous Frequencies of Molecules or IFM, is presented. The IFM method, when coupled to classical molecular simulations, can predict the FFCF of a molecule in solutions. Here, *N*-methylacetamide (NMA) in seven different chemical environments, both simple and complex, is used to test this new method. The results show good agreement with experiments for the NMA solvatochromism and FFCF dynamics, including characteristic times and amplitudes of fluctuations. In addition, the new method shows equivalent or improved results when compared to conventional frequency maps. Overall, the use of the new method in conjunction with molecular dynamics simulations allows unlocking the full potential of IR spectroscopy to generate molecular maps from vibrational observables, capable of describing the interaction landscape of complex molecular systems.

Introduction

Complex liquid solutions are multi-component systems with intricate molecular arrangements arising from the diverse interactions among the components in the mixture. Due to the large number of contributions to the potential energy surface (PES), complex molecular systems typically exhibit nanoscopic heterogeneities, especially in cases where the solution contains molecules with contrasting chemical moieties, such as polar and non-polar.[1] Therefore, the complex and non-ideal behaviour of the liquid creates challenges associated with its characterization, both experimentally and computationally. Moreover, the liquid state of the mixture allows its molecular components to diffuse freely (translate and rotate), resulting in a chemical environment that varies as a function of time.

Consequently, one can characterize the structure and dynamics of the complex solution to indirectly assess the PES since these two properties are governed by its energy landscape (Fig. 1).

The development of methods that can generate a map of intermolecular forces based on the dynamics and interactions of each component in the mixture has been a long-standing goal for understanding and later tailoring complex molecular systems. Experimentally, the complete characterization of complex solutions requires the use of techniques capable of resolving the individual interactions occurring in the multi-component system. Among the various experimental methods commonly used to characterize the molecular structure of solutions, one can find NMR, UV-vis, fluorescence, Raman and IR spectroscopies as well as scattering techniques. For example, NMR has been successfully applied to study the diffusion of species in complex solutions, the molecular structure of DNA and RNA in live human cells,[2] and the isomerization dynamics in ionic liquids.[3] However, NMR spectroscopy lacks the required time resolution to characterize solutions undergoing ultrafast dynamics with characteristic times faster than nanoseconds, such as liquids at room temperature, and yields averaged

[a]Department of Chemistry, Louisiana State University, Baton Rouge, Louisiana 70803, USA. E-mail: dkuroda@lsu.edu

[b]Department of Chemistry, Universidad Nacional de Colombia Sede Bogotá, Bogotá 111321, Colombia

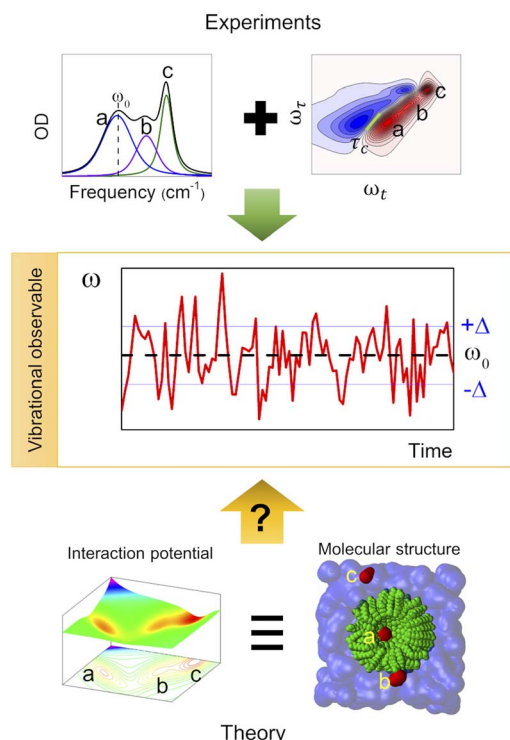† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4sc03219e

Experiments

Fig. 1 Use of vibrational observables to study the PES and chemical environment of a molecular system.

signals for these processes.[4] UV-vis and fluorescence spectroscopies have also shown potential for the analysis of micellar systems,[5] for probing the chemical environment of hydrocarbons,[6] and for exploring the air/liquid interphases.[7] Notwithstanding the capabilities of these two spectroscopies, they are limited by the number of substrates with responses in the UV-vis spectral range, which significantly reduces the number of complex solutions that can be analysed with them.

Scattering techniques, such as X-ray and neutron scattering, have also been used to characterize complex solutions such as deep eutectic solvents,[1,8–10] ionic liquids,[11–13] and solutions of proteins and DNA.[14–16] These techniques provide structural information at the atomic level through the scattering profile. For a liquid sample, the peaks of the scattering profile contain different atomic contributions, requiring a nontrivial deconvolution to elucidate the structure from the data.[17] To this end, the scattering experiments usually rely on molecular dynamics (MD) simulations to model and evaluate the different atomic contributions to each peak.[9,12,13] Hence, the correct modelling of scattering data requires a good parameterization of the MD force field since two different atomistic models can generate the same scattering profile.[18]

An alternative characterization method for complex solutions relies on the use of vibrational spectroscopy, which includes both IR and Raman spectroscopies. By analysing the frequency shift of the vibrational bands in the IR or Raman spectrum, it is possible to probe the chemical environment.[19–22] In either case, the vibrational bands are related to the molecular vibrational modes, either through the dipole (IR)[23] or the

polarizability (Raman)[24] of the molecule; consequently, the bands change as a function of the molecular environment.

Normal modes are easily calculated from the motions in a potential energy surface, but the dipole moment associated with a vibrational transition is simpler than the equivalent polarizability tensor, which requires full electronic structure calculations.[25,26] Hence, the changes in the Raman bands are often difficult to relate to specific molecular environments.

Normal modes are mathematical entities related to the diagonalization of the Hessian matrix (*i.e.*, second derivatives of the PES at the minimum) and represent the motion of collections of atoms in a molecule. More importantly, the IR band can be understood in terms of normal modes using the dipole moment formalism.[23] Similar to the PES, the normal mode frequencies (PES curvature) are highly dependent on the molecular speciation and environment (*i.e.*, structure, temperature, pressure, molecular interactions, *etc.*).[19,27,28] Therefore, the molecular environment of a molecule can be, and is, typically deduced from the central frequency of an IR band associated with a particular vibrational mode. In addition, the time evolution (dynamics) of the vibrational modes in a molecular system can be characterized using time resolved vibrational spectroscopy. Among the various possible time resolved methods, 2DIR spectroscopy is the most powerful for describing the dynamics of the system at thermal equilibrium using specific sets of atoms, such as carbonyl groups.[29] Among other features, 2DIR spectroscopy measures the dynamics of the frequency-fluctuation correlation function (FFCF), or equivalently, the time scale at which the system loses its memory of a given molecular environment.[27] In particular, the FFCF provides information about the interaction potential among the molecules through their dynamics.[30] Moreover, the FFCF contains information about the various chemical environments through its amplitude term because it is the variance of the distribution of the frequencies (environments) sensed by the molecules in the sample.[27]

The use of 2DIR spectroscopy to characterize the molecular environments in a complex solution is also challenging due to the lack of prior knowledge of the chemical environment and interactions affecting the molecule under study. Thus, to obtain the location and interaction landscape of the molecular components *via* 2DIR spectroscopy, the modelling of the system using MD simulations is often required. However, the MD simulation does not directly provide the vibrational observables, and instead these are computed using the so-called frequency maps.

Frequency maps transform molecular coordinates into specific variables, such as electric fields, to obtain the instantaneous frequency of a molecule in the MD simulation.[31] Hence, frequency maps have been instrumental in understanding and modelling vibrational spectroscopy data at a low computational cost. However, the applicability of this tool is generally limited to specific molecular systems for which the map was originally parametrized, and its extension to "new" chemical environments often requires further parametrization.[32–35] Moreover, the intricacies in the theoretical description of complex molecular environments with nanoscopic heterogeneities do not always

guarantee the access to such modelling and its transformation into vibrational observables, such as frequency shifts. In order to facilitate the experimental analysis of the IR spectroscopic data (IR position, line shape and dynamics), a new computational methodology is needed. In particular, this methodology should not only accurately translate molecular arrangements (environments) of highly heterogeneous molecular systems into frequency fluctuations but also be computationally efficient.

In this work, a new methodology (Instantaneous Frequencies of Molecules, IFM) is presented. This method uses a quantum mechanical method to compute the instantaneous frequencies of the molecules in a given molecular environment (Fig. 2). The exponential growth of computational power has allowed the direct use of *ab initio* quantum chemical and semiempirical methods to calculate frequencies. However, the proper description of the vibrational frequencies of a solvated molecule usually requires multiple solvation shells beyond the first one.[31] Hence, a low cost computational method is needed to calculate the vibrational frequencies of such large systems, typically containing more than 100 atoms.

GFN2-xTB is a new tight binding approach that uses multipole expansion and density-dependent dispersion fluctuation to describe electrostatic interactions.[36,37] In particular, this semiempirical method is tailored to accurately compute the geometries, non-covalent interactions, and vibrational frequencies of molecular systems and has low computational cost.[36,37] Furthermore, it has recently been shown that GFN2-xTB correctly predicts the vibrational solvatochromism of various molecular probes typically used in IR spectroscopy, depending only on the correct description of the solute and its chemical environment as defined by the classical force field.[20] This makes IFM calculations fully transferable to complex molecular arrangements without the need of new parameterization, beyond the selection of an appropriate classical representation (*i.e.*, force field).

The applicability of the IFM method for vibrational spectroscopy is tested on a molecular system composed of *N*-methylacetamide (NMA) in different solvents, ranging from simple to complex. NMA was selected because of its amide I vibrational mode, which is highly sensitive to the chemical environment, as evidenced by its solvatochromic frequency shifts in different solvents.[38–40] In addition, as a model peptide unit, NMA has been extensively studied both experimentally and computationally. From a theoretical perspective, there are now many MD simulation studies dedicated to the solvation structure and dynamics of this molecule and some provide frequency maps to correlate with vibrational dynamics studies.[32,41–44] From an experimental perspective, the vibrational dynamics of NMA has been thoroughly investigated using time resolved vibrational spectroscopy,[32,33,41,45–47] making it an excellent reference for the IFM method. To this end, the following systems have been selected: NMA in heavy water (D_2O), chloroform (CLF), dimethylsulfoxide (DMSO), tetrahydrofuran (THF), and toluene (TOL), as well as mixtures of DMSO and THF and of DMSO and D_2O. The suitability of the computational methodology combining MD simulations and the IFM method (MD-IFM) was evaluated by computing and comparing with experimental data the vibrational frequency shifts (solvatochromism), the FFCF dynamics (solvent motions), and the amplitudes of the frequency fluctuations (distribution of environments) of NMA in different molecular environments.

Methods

The experimental measurements to test the IFM method in the present work include linear and non-linear IR spectroscopy. The procedures and specifications can be found in the ESI.†

The schematic computational methodology for the IFM calculation is described in Fig. 2. However, a brief description is given in the following sections, and an example of the code used here can be found free of charge at **https://github.com/dkurodalab/IFM**.

Molecular dynamics

MD simulations were carried out for the systems of NMA in D_2O, CLF, DMSO, THF, TOL, DMSO : THF (1 : 3), and DMSO : D_2O (1 : 1). In the first and last cases (D_2O and D_2O : DMSO (1 : 1) as solvents), the hydrogen atom attached to the nitrogen atom of NMA was replaced by a deuterium. Solvent boxes of 40 Å with or without the solute were built using PACKMOL.[48] The quality of the force field was evaluated from the densities determined from an MD simulation in the pure system. MD simulations were performed using the AMBER 18 computational package.[49] For water and all the other solvents, the TIP3P model and Generalized Amber Force Field (GAFF) were used, respectively.[50] The NMA molecule was parameterized with the ff14SB force field.[51] The MD simulation was performed by starting with an energy minimization, followed by a heating step from 0 to 300 K
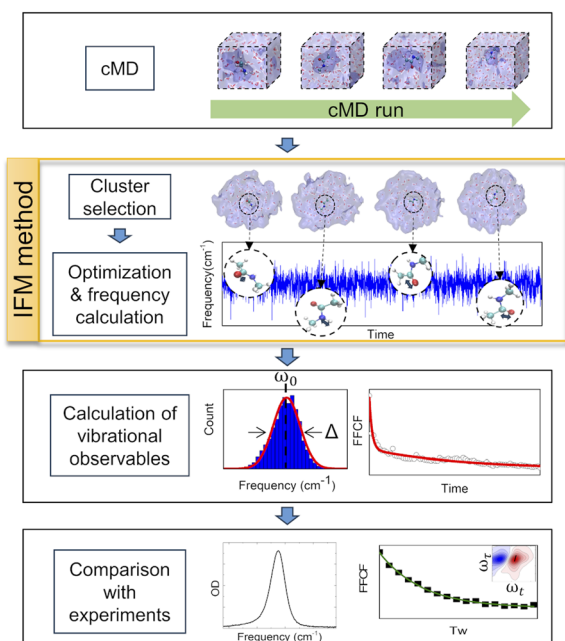

Fig. 2 Procedure followed for the development and testing of the IFM method.

over 20 ps; the production run was carried out for 250 ps in the NPT ensemble after an equilibration run of 20 ns. A time step of 0.5 fs was used for all simulation stages. For the NMA/TOL system, a 1 ns production run was performed to accomplish the necessary averaging in the FFCF calculation.

Frequency calculations and FFCF

Snapshots were selected every 50 fs from the MD production run and subsequently trimmed to the number of solvation shells needed for optimization and frequency calculations. At least two solvation shells were selected in each case, as they are sufficient to yield reliable frequency values (see the ESI†). To preserve the distribution of states from the MD, the geometry energy minimizations were performed only for the NMA molecule embedded in the selected solvation shells using the GFN2-xTB package. The geometry optimization and frequency calculations were carried out at the GFN2-xTB level of theory. The central frequencies for the amide I normal mode of NMA in different solvents represent the intensity weighted average. Note that the selection of GFN2-xTB semiempirical method is based on the low computational cost when compared to other *ab initio* methods, such as DFT (see the ESI†).

The FFCF was calculated from the instantaneous amide I frequencies over time: $C(t) = \delta\omega(t)\delta\omega(0)$, where $\delta\omega(t)$ is the delta between the frequency at time t and the average frequency ($\omega(t)-\langle\omega\rangle$).[27] The correlation time (τ_c) was obtained by modelling the calculated FFCF with the Kubo formalism (*i.e.*, the bi-exponential decay function, see the ESI† for the fits of all molecular systems). The amplitude of the fluctuations was obtained by analysing the frequency distributions with Gaussian profiles.[27]

Results and discussion

The IFM method predicts Gaussian distributions for the amide I frequencies of NMA in different solvents, consistent with the stochastic motions of the molecular environment (see the ESI†).[52] Because of the nature of the amide I vibrational mode, these distributions contain information about the molecular environment through their central frequencies; *i.e.*, sol-vatochromism.[28,43,53] The central frequencies predicted by the IFM method correctly reproduce the experimental sol-vatochromic shift of NMA amide I mode in different solvents including mixtures. Moreover, a linear relationship is observed between the computed and experimental values (Fig. 3). A linear model of frequencies also shows that the IFM method over-estimates their values by ∼60 cm^{-1}. This deviation from experimental values is associated with the semiempirical method.[20] A Pearson correlation coefficient (R^2) of 0.874 is observed, indicating not only the strong linear correlation between the predicted and experimental frequency values but also the suitability of the IFM model for computing instantaneous frequencies. Notably, the NMA/TOL system deviates significantly from the linear trend. The linear model without NMA–toluene has a significantly better predictive power (*i.e.*, R^2 increases to 0.990, see the ESI†). Note that the deviation in the
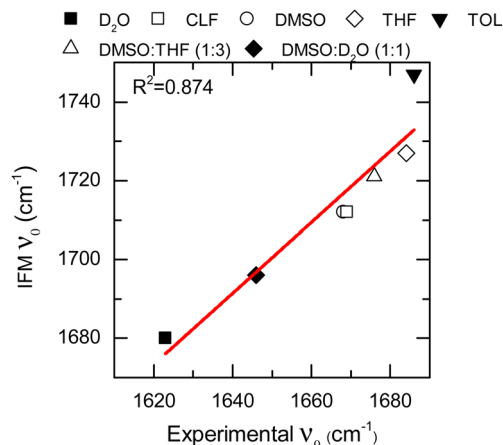

Fig. 3 Solvatochromic shift for the NMA amide I mode in different solvents calculated through the IFM method *vs.* experimentally through FTIR (data points). Linear fit correspond to the red line.

prediction of the vibrational solvatochromism of amide I of NMA in toluene is likely due to a deficiency in the force field describing toluene in the MD simulation and not in the method. This assertion based on the GFN2-xTB method correctly characterizes other complex solvents, such as DMSO : D$_2$O, where various and competing interactions are found.[28,45,54–58]

To further assess the suitability of the computational methodology (MD-IFM), the FFCF was computed from the instantaneous frequencies predicted by the IFM method. Assuming a Kubo stochastic model,[27] the FFCF includes both correlation times (τ_c) and amplitude of frequency fluctuations (Δ) of the form: $c(t) = \sum_i \Delta_i^2 e^{-t/\tau_{ci}}$. The computed τ_c as a function of the chemical environment also shows a linear relationship with the experimental values (Fig. 4, Table 1, and the ESI†). In general, the predicted correlation times are in reasonable agreement with the experiments. In addition, NMAD/D$_2$O is a particularly well studied system and has frequency maps to
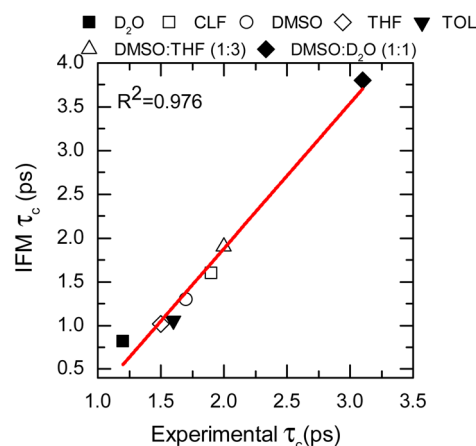

Fig. 4 Correlation times calculated from the IFM method *vs.* the experimental values determined from 2DIR (data points). Linear fit corresponds to the red line.

Table 1 Central frequencies, correlation times, and fluctuation amplitudes determined from the IFM method, experimentally, and from frequency maps

Solvent	$\tilde{V}$ (cm^{-1})		τ_c (ps)			Δ (cm^{-1})			
	IFM	Exp.	MD-IFM	Exp.	Freq. maps	IFM	Exp.	Freq. maps[28]	From IR
D$_2$O	1680	1623	0.82 ± 0.03	1.2 (ref. 64)	0.76,[56] 0.75 (ref. 56)	20.9 ± 0.3	19.75,[28] 9 (ref. 63)	17[a]	8.2
CLF	1712	1669	1.6 ± 0.7	1.9 ± 0.1	2.9 (ref. 28)[a]	12.1 ± 0.3	13.57 (ref. 28)	8.8[a]	7.7
DMSO	1712	1668	1.3 ± 0.2	1.7 ± 0.1	1.6 (ref. 28)[a]	10.95 ± 0.08	11.55 (ref. 28)	7.8[a]	6.1
THF	1727	1684	1.02 ± 0.06	1.5 ± 0.1	N/A	9.1 ± 0.1	N/A	N/A	3.1
TOL	1747	1686	1.06 ± 0.07	1.6 ± 0.2	N/A	7.0 ± 0.2	N/A	N/A	3.4
DMSO : THF	1721	1676	1.9 ± 0.1	2.0 ± 0.2	N/A	10.2 ± 0.1	N/A	N/A	6.1
DMSO : D$_2$O	1696	1646	3.8 ± 0.2	3.1 ± 0.3	N/A	14.6 ± 0.2	N/A	N/A	11.1

[a] Averaged value.

predict the amide I mode (*e.g.* the electric field and electrostatic potential models).[56] Comparison of the new method with the previously reported results from frequency maps shows very similar results. Specifically, the correlation time obtained from frequency maps was ~0.75 ps,[56] while the new method based on GNF2-xTB predicted a value of 0.82 ps. A similar result is obtained for the NMA/DMSO system, where τ_c values of 1.6 ps and of 1.3 ps are predicted from the frequency map and the new method, respectively. In contrast, NMA in CLF presents different correlation times for the FFCF dynamics when computed using either frequency maps[28] or the new method. In this case, the frequency map and the new method predicts τ_c values of 2.9 ps and of 1.6 ps, respectively; while the experimental result is 1.9 ps.[28] Hence, it appears that the new method has a smaller error than the frequency map, which is likely due to the IFM method taking into account more detailed interactions affecting the frequency of the amide I mode, such as weak hydrogen bonds.[20]

The goodness of the MD-IFM method for computing instantaneous frequencies was also further supported by computation of NMA amide I in other chemical environments. For example, the method predicts τ_c values of 1.02 ps and 1.06 ps for the NMA/THF and NMA/TOL systems, respectively, which are in good agreement with the experimental values of 1.5 ps (NMA/THF) and 1.6 ps (NMA/TOL). Overall, the use of the new method in single solvents allows us to demonstrate not only the validity of the MD-IFM method for predicting the dynamics of the FFCF when compared to other methods but also the transferability of the IFM method to different solvents without the need for modification or reparametrization.

The adequate predictions made by the MD-IFM method allow us to extend its application to complex molecular systems, such as mixtures of solvents. To this end, two different mixtures were selected: DMSO : THF (1 : 3 molar ratio) and DMSO : D$_2$O (1 : 1 molar ratio). These two particular systems represent solvents with competitive dipole–dipole interactions as well as hydrogen bond and picosecond dynamics.[28,40,56,59,60] The characteristic time of the FFCF for the amide I vibrational mode in these two solvents was found to be 1.9 ps (DMSO : THF) and 3.8 ps (DMSO : D$_2$O), which are in very good agreement with the

observed experimental values of 2.0 and 3.2 ps, respectively (see Fig. 4, Table 1, and the ESI†).

These last two cases are particularly important because of the complexity of the molecular environment or, analogously, the chemical interaction landscape. The selected molecular landscapes make it particularly challenging to parametrize a single frequency map capable of representing specific probes in various solvents. Consequently, a single frequency map is usually not capable of describing the solvatochromism of a probe on both the individual solvents and their mixtures when the solvents are very different. This is the case for DMSO and D$_2$O, where not only the hydrogen bond donating and accepting capabilities are changed but also their dipoles. The correct description of both single solvents and mixtures provides strong support for the broad applicability of the new methodology in complex environments.[31,61] Overall, it is shown that the IFM-MD methodology can be applied to compute with good agreement with the experiment not only the vibrational solvatochromism of molecules in solution ($R^2 = 0.874$, Fig. 3) but also the FFCF decorrelation times ($R^2 = 0.976$, Fig. 4). More importantly, this is achieved without any additional parameterization or modelling beyond the MD simulation force field.

It is derived from the data that central frequencies and correlation times of the amide I mode in different solvents are not correlated ($R^2 = -0.192$, see the ESI†). This result demonstrates that the two vibrational observables are independent and that both are needed to obtain a correct description of the molecular system. However, the central frequencies and correlation times are not sufficient to completely determine the molecular environment observed by a molecule. A good example of the ambiguity in determining the molecular environment from vibrational observables can be seen in the cases of NMA/DMSO and NMA/CLF. In these two systems, the central frequencies for the amide I mode (1712 cm^{-1} for both) and the correlation times (1.3 ± 0.2 and 1.6 ± 0.7 ps, respectively) are similar. Therefore, these predictions, while accurate, are not sufficient to determine the chemical environment of the studied molecule and require a third observable.

Thus far, only the central frequencies and correlation times have been used to assess the molecular environment of a molecule using IR spectroscopy. Another important parameter

that defines the line shape of a vibrational transition is the amplitude of the frequency fluctuations. This vibrational observable contains information about the ensemble of molecular environments that the molecule can observe thermally. Consequently, the amplitudes of frequency fluctuations (Δ) are strongly dependent on the chemical nature of the system.[27] Experimentally, these amplitudes are obtained by modelling the FTIR and 2DIR spectra.[62] Particularly, the computed amplitudes are found to be 12.1 ± 0.3, 10.95 ± 0.08 and 20.9 ± 0.3 cm^{-1} for NMA in CLF, DMSO and D$_2$O, respectively, and compare very well with those derived from the experiment: 13.57 cm^{-1}, 11.55 cm^{-1}, and 19.75 cm^{-1}, respectively.[28]

Nonetheless, it is important to note that the NMA in D$_2$O experimental Δ has been reported by others to be 9 cm^{-1}.[63] Therefore, it was excluded from the overall analysis of the frequency fluctuation amplitude.

The IFM-based amplitudes of the remaining systems were compared with the experimental values by using a more approximate method to compute the fluctuation amplitudes from the FTIR and 2DIR spectra.[62] The experimental amplitudes, Δ, show a good agreement with those obtained by the IFM method (Fig. 5 and Table 1). The correlation between the experiment and computation is very good, as seen by the linear correlation between the experimental and computed values ($R^2 = 0.882$). Interestingly, the amplitude for the NMA/TOL system has the largest deviation from the linear trend observed for the other samples. In the case of NMA/TOL, the result is not surprising, given that the method is not particularly good at predicting the vibrational solvatochromism of the same system, as previously shown. Nonetheless, the high Pearson correlation coefficient between the experimental and computation amplitudes indicates that the IFM-based method predicts reasonably well the amplitudes of the frequency fluctuations in different systems. More importantly, the strong correlation between the experiment and theory shows that the IFM method captures more detailed interactions, which directly translates into more accurate frequency fluctuations, even when compared to other maps.

Overall, it was demonstrated that the method based on the IFM frequencies produces a reliable prediction of the solvatochromism (ω_0), dynamics (τ_c) and distribution of environments (Δ) from different samples without the need for reparametrization, which is usually required when using frequency maps. These metrics, obtained for a molecule (NMA) solvated in pure solvents and their mixtures, describe the effect of the dynamics and interactions of the molecular environment on the investigated molecule. Therefore, the method presented here allows unambiguous assignment of experimental observables in the vibrational spectra (linear and non-linear), unlocking the full potential of vibrational observables to create molecular maps capable of describing the interaction landscape of complex molecular systems.

Conclusions

A new method for frequency calculation is presented. This method takes advantage of the existing and new computational tools (MD simulations and semi-empirical methods), as well as of current computational power to obtain the frequencies of solvated molecules in large clusters. The new methodology in conjunction with MD simulations allows us to determine vibrational solvatochromism, dynamics, and distribution of environments. The new method is successfully tested on a well-studied system: *N*-methylacetamide in polar solvents and their mixtures. In addition, the transferability of the method to more complex binary solutions was examined. The correct description of the vibrational solvatochromism in both simple and complex chemical environments shows the utility of the method to predict the location of a probe in a solution. Moreover, it is observed that the combination of frequencies, decorrelation times and fluctuation amplitudes is sufficient to fully characterize a specific system. The use of such a tool is particularly important for vibrational spectroscopy to unlock the full potential of the IR methodology when probing highly heterogeneous environments.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author contributions

L. X. S.-M.: methodology, data curation, validation, formal analysis, investigation, writing – original draft, visualization; J. F. G.: methodology, writing – original draft, writing – review and editing, supervision; D. G. K.: conceptualization, methodology, writing – review and editing, resources, supervision, project administration, funding acquisition.

Conflicts of interest
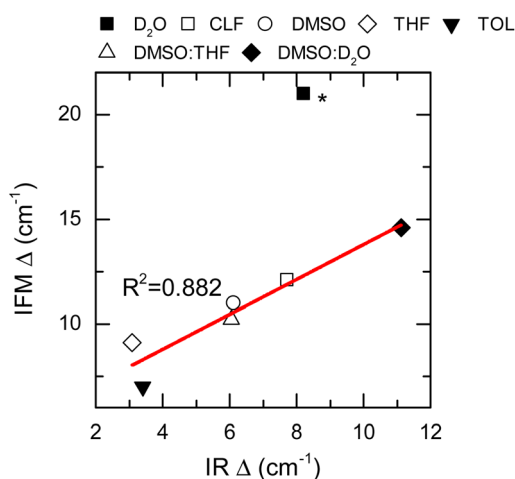
There are no conflicts to declare.

Fig. 5 Frequency fluctuation amplitudes (Δ) calculated by the IFM *vs.* determined from the IR spectra (data points). Linear fit corresponds to the red line. Star (*) NMA in the D$_2$O Δ value was not included into the fit.

Acknowledgements

References

1 Y. Cui, J. C. Rushing, S. Seifert, N. M. Bedford and D. G. Kuroda, Molecularly heterogeneous structure of a nonionic deep eutectic solvent composed of n-methylacetamide and lauric acid, *J. Phys. Chem. B*, 2019, **123**, 3984–3993.

2 Y. Yamaoki, A. Kiyoishi, M. Miyake, F. Kano, M. Murata, T. Nagata and M. Katahira, The first successful observation of in-cell NMR signals of DNA and RNA in living human cells, *Phys. Chem. Chem. Phys.*, 2018, **20**, 2982–2985.

3 R. Clark, M. A. Nawawi, A. Dobre, D. Pugh, Q. Liu, A. P. Ivanov, A. J. White, J. B. Edel, M. K. Kuimova and A. J. McIntosh, The effect of structural heterogeneity upon the microviscosity of ionic liquids, *Chem. Sci.*, 2020, **11**, 6121–6133.

4 R. G. Bryant, The NMR time scale, *J. Chem. Educ.*, 1983, **60**, 933.

5 R. Sabaté, M. Gallardo, A. de la Maza and J. Estelrich, A spectroscopy study of the interaction of pinacyanol with n-dodecyltrimethylammonium bromide micelles, *Langmuir*, 2001, **17**, 6433–6437.

6 B. Narayan, K. Nagura, T. Takaya, K. Iwata, A. Shinohara, H. Shinmori, H. Wang, Q. Li, X. Sun and H. Li, The effect of regioisomerism on the photophysical properties of alkylated-naphthalene liquids, *Phys. Chem. Chem. Phys.*, 2018, **20**, 2970–2975.

7 C. Rubia-Payá, G. De Miguel, M. T. Martín-Romero, J. J. Giner-Casares and L. Camacho, UV–Vis Reflection–Absorption Spectroscopy at air–liquid interfaces, *Adv. Colloid Interface Sci.*, 2015, **225**, 134–145.

8 I. Manasi, R. Schweins, K. Ma and K. J. Edler, Nanostructure in Amphiphile-Based Deep Eutectic Solvents, *Langmuir*, 2023, **39**, 16776–16784.

9 A. Malik and H. K. Kashyap, Heterogeneity in hydrophobic deep eutectic solvents: SAXS prepeak and local environments, *Phys. Chem. Chem. Phys.*, 2021, **23**, 3915–3924.

10 M. Sakuragi, S. Tsutsumi and K. Kusakabe, Deep eutectic solvent-induced structural transition of microemulsions explored with small-angle X-ray scattering, *Langmuir*, 2018, **34**, 12635–12641.

11 T. Murphy, R. Atkin and G. G. Warr, Scattering from ionic liquids, *Curr. Opin. Colloid Interface Sci.*, 2015, **20**, 282–292.

12 H. K. Kashyap, C. S. Santos, H. V. Annapureddy, N. S. Murthy, C. J. Margulis and E. W. Castner Jr, Temperature-dependent structure of ionic liquids: X-ray scattering and simulations, *Faraday Discuss.*, 2012, **154**, 133–143.

13 A. McGrogan, J. Lafferty, L. O'Neill, L. Brown, J. M. Young, P. Goodrich, M. J. Muldoon, L. Moura, S. Youngs and T.-L. Hughes, Liquid Structure of Ionic Liquids with [NTf2]$^-$ Anions, Derived from Neutron Scattering, *The Journal of Physical Chemistry B*, 2024, **128**, 3220–3235.

14 A. H. Larsen, Y. Wang, S. Bottaro, S. Grudinin, L. Arleth and K. Lindorff-Larsen, Combining molecular dynamics simulations with small-angle X-ray and neutron scattering data to study multi-domain proteins in solution, *PLoS Comput. Biol.*, 2020, **16**, e1007870.

15 T. W. Gräewert and D. I. Svergun, Structural modeling using solution small-angle X-ray scattering (SAXS), *J. Mol. Biol.*, 2020, **432**, 3078–3092.

16 S. Yamamoto, F. Kono, K. Nakatani, M. Hirose, K. Horii, Y. Hippo, T. Tamada, Y. Suenaga and T. Matsuo, Structural characterization of human *de novo* protein NCYM and its complex with a newly identified DNA aptamer using atomic force microscopy and small-angle X-ray scattering, *Front. Oncol.*, 2023, **13**, 1213678.

17 A. K. Soper, The radial distribution functions of water as derived from radiation total scattering experiments: Is there anything we can say for sure?, *Int. Scholarly Res. Not.*, 2013, **2013**, 279463.

18 T. Head-Gordon and G. Hura, Water structure from scattering experiments and simulation, *Chem. Rev.*, 2002, **102**, 2651–2670.

19 I. R. Whittall, A. M. McDonagh, M. HAMPHREY and M. Samoc, Organometallic complexes in nonlinear optics I: second-order nonlinearities, *ChemInform*, 1998, **29**, 199830303.

20 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*, 2023, 7955–7963.

21 M. G. Romei, E. V. von Krusenstiern, S. T. Ridings, R. N. King, J. C. Fortier, C. A. McKeon, K. M. Nichols, L. K. Charkoudian and C. H. Londergan, Frequency changes in terminal alkynes provide strong, sensitive, and solvatochromic Raman probes of biochemical environments, *J. Phys. Chem. B*, 2022, **127**, 85–94.

22 S. Bagchi, S. D. Fried and S. G. Boxer, A solvatochromic model calibrates nitriles' vibrational frequencies to electrostatic fields, *J. Am. Chem. Soc.*, 2012, **134**, 10373–10376.

23 E. B. Wilson, J. C. Decius and P. C. Cross, *Molecular vibrations: the theory of infrared and Raman vibrational spectra*, Courier Corporation, 1980.

24 H. A. Szymanski, *Raman spectroscopy: theory and practice*, Springer Science & Business Media, 2012.

25 S. Y. Lee, Placzek-type polarizability tensors for Raman and resonance Raman scattering, *J. Chem. Phys.*, 1983, **78**, 723–734.

26 L. A. Nafie, Theory of Raman scattering, *Pract. Spectrosc.*, 2001, **28**, 1–10.

27 P. Hamm and M. Zanni, *Concepts and methods of 2D infrared spectroscopy*, Cambridge University Press, 2011.

28 M. DeCamp, L. DeFlores, J. McCracken, A. Tokmakoff, K. Kwac and M. Cho, Amide I vibrational dynamics of N-methylacetamide in polar solvents: The role of electrostatic interactions, *J. Phys. Chem. B*, 2005, **109**, 11016–11026.

29 H. Bakker and J. Skinner, Vibrational spectroscopy as a probe of structure and dynamics in liquid water, *Chem. Rev.*, 2010, **110**, 1498–1517.

30 K. Kwac and M. Cho, Two-color pump – probe spectroscopies of two-and three-level systems: 2-dimensional line shapes and solvation dynamics, *J. Phys. Chem. A*, 2003, **107**, 5903–5912.

31 C. R. Baiz, B. Błasiak, J. Bredenbeck, M. Cho, J.-H. Choi, S. A. Corcelli, A. G. Dijkstra, C.-J. Feng, S. Garrett-Roe and N.-H. Ge, Vibrational spectroscopic map, vibrational spectroscopy, and intermolecular interaction, *Chem. Rev.*, 2020, **120**, 7152–7218.

32 E. Małolepsza and J. E. Straub, Empirical maps for the calculation of amide I vibrational spectra of proteins from classical molecular dynamics simulations, *J. Phys. Chem. B*, 2014, **118**, 7848–7855.

33 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*, 2011, **115**, 3713–3724.

34 O. O. Mesele and W. H. Thompson, A "universal" spectroscopic map for the OH stretching mode in alcohols, *J. Phys. Chem. A*, 2017, **121**, 5823–5833.

35 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*, 2016, **120**, 3888–3896.

36 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.*, 2019, **15**, 1652–1671.

37 C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher and S. Grimme, Extended tight-binding quantum chemistry methods, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2021, **11**, e1493.

38 P. Mukherjee, I. Kass, I. T. Arkin and M. T. Zanni, Structural disorder of the CD3ζ transmembrane domain studied with 2D IR spectroscopy and molecular dynamics simulations, *J. Phys. Chem. B*, 2006, **110**, 24740–24749.

39 C. Fang, A. Senes, L. Cristian, W. F. DeGrado and R. M. Hochstrasser, Amide vibrations are delocalized across the hydrophobic interface of a transmembrane helix dimer, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 16740–16745.

40 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.*, 2006, **124**, 044502.

41 S. Ham, J.-H. Kim, H. Lee and M. Cho, Correlation between electronic and molecular structure distortions and vibrational properties. II. Amide I modes of NMA–n D_2O complexes, *J. Chem. Phys.*, 2003, **118**, 3491–3498.

42 J.-H. Choi, S. Ham and M. Cho, Local amide I mode frequencies and coupling constants in polypeptides, *J. Phys. Chem. B*, 2003, **107**, 9132–9138.

43 K. Kwac and M. Cho, Differential evolution algorithm approach for describing vibrational solvatochromism, *J. Chem. Phys.*, 2019, 151.

44 P. Bouř and T. A. Keiderling, Empirical modeling of the peptide amide I band IR intensity in water solution, *J. Chem. Phys.*, 2003, **119**, 11253–11262.

45 Y.-S. Lin, J. Shorb, P. Mukherjee, M. Zanni and J. Skinner, Empirical amide I vibrational frequency map: application to 2D-IR line shapes for isotope-edited membrane peptide bundles, *J. Phys. Chem. B*, 2009, **113**, 592–602.

46 E.-L. Karjalainen, T. Ersmark and A. Barth, Optimization of model parameters for describing the amide I spectrum of a large set of proteins, *J. Phys. Chem. B*, 2012, **116**, 4831–4842.

47 M. Reppert and A. Tokmakoff, Electrostatic frequency shifts in amide I vibrational spectra: Direct parameterization against experiment, *J. Chem. Phys.*, 2013, **138**, 04B603.

48 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.*, 2009, **30**, 2157–2164.

49 D. A. Case, T. Darden, T. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, M. Crowley, R. C. Walker and W. Zhang, *Amber 10*, University of California, 2008.

50 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.*, 2004, **25**, 1157–1174.

51 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.

52 R. Kubo, A stochastic theory of line shape, *Adv. Chem. Phys.*, 1969, **15**, 101–127.

53 K. Kwac and M. Cho, Machine learning approach for describing vibrational solvatochromism, *J. Chem. Phys.*, 2020, **152**, 174101.

54 M. A. Krestyaninov, D. V. Ivlev, A. A. Dyshin, D. M. Makarov, M. G. Kiselev and A. M. Kolker, Complex investigation of H-bond in Water-N-methylacetamide system: Volumetric properties, DFT, IR, MD analysis, *J. Mol. Liq.*, 2022, 119533.

55 C. Lawrence and J. Skinner, Vibrational spectroscopy of HOD in liquid D_2O. III. Spectral diffusion, and hydrogen-bonding and rotational dynamics, *J. Chem. Phys.*, 2003, **118**, 264–272.

56 J. Schmidt, S. Corcelli and J. Skinner, Ultrafast vibrational spectroscopy of water and aqueous N-methylacetamide: Comparison of different electronic structure/molecular dynamics approaches, *J. Chem. Phys.*, 2004, **121**, 8887–8896.

57 A. Chand and S. Chowdhuri, Effects of dimethyl sulfoxide on the hydrogen bonding structure and dynamics of aqueous N-methylacetamide solution, *J. Chem. Sci.*, 2016, **128**, 991–1001.

58 R. Schweitzer-Stenner, H. Carson and D. DiGuiseppi, Probing the replacement of water by dimethyl sulfoxide in the hydration shell of N-methylacetamide by FTIR-spectroscopy, *Vib. Spectrosc.*, 2017, **92**, 251–258.

59 L. P. DeFlores, Z. Ganim, S. F. Ackley, H. S. Chung and A. Tokmakoff, The anharmonic vibrational potential and relaxation pathways of the amide I and II modes of N-methylacetamide, *J. Phys. Chem. B*, 2006, **110**, 18973–18980.

60 S. R. Galle Kankanamge, J. Ma, R. T. Mackin, F. M. Leonik, C. M. Taylor, I. V. Rubtsov and D. G. Kuroda, Proving and Probing the Presence of the Elusive C–H⋯O Hydrogen Bond in Liquid Solutions at Room Temperature, *Angew. Chem.*, 2020, **132**, 17160–17165.

61 K.-I. Oh and C. R. Baiz, Empirical S= O stretch vibrational frequency map, *J. Chem. Phys.*, 2019, 151.

62 D. G. Kuroda, D. Y. Vorobyev and R. M. Hochstrasser, Ultrafast relaxation and 2D IR of the aqueous trifluorocarboxylate ion, *J. Chem. Phys.*, 2010, **132**, 044501.

63 S. Woutersen, R. Pfister, P. Hamm, Y. Mu, D. S. Kosov and G. Stock, Peptide conformational heterogeneity revealed from nonlinear vibrational spectroscopy and molecular-dynamics simulations, *J. Chem. Phys.*, 2002, **117**, 6833–6840.

64 A. Ghosh and R. M. Hochstrasser, A peptide's perspective of water dynamics, *Chem. Phys.*, 2011, **390**, 1–13.