The Behavior of Methane-Water Mixtures Under Elevated Pressures Using Many-Body Potentials

Victor Naden Robinson,1 Raja Ghosh,2 Colin K. Egan,1 Marc Riera,2 Christopher Knight,3 Francesco Paesani,2,4,5, a) and Ali Hassana1, b)
1) The “Abdus Salam” International Centre for Theoretical Physics, I-34151 Trieste, Italy
2) Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093, United States
3) Computational Science Division, Argonne National Laboratory, 9700 South Cass Avenue, Lemont, Illinois 60439, United States
4) Materials Science and Engineering, University of California San Diego, La Jolla, California 92093, United States
5) San Diego Supercomputer Center, University of California San Diego, La Jolla, California 92093, United States

Non-polarizable empirical potentials have been shown not to be able to capture the mixing of methane–water mixtures at elevated pressures. Although DFT-based ab initio simulations may circumvent this discrepancy, they are limited in terms of the relevant time and length scales associated with mixing phenomena. Here we show that the many-body MB-nrg potential, designed to reproduce methane–water interactions with coupled cluster accuracy, successfully captures this phenomenon up to 3 GPa and 500 K with varying methane concentration. Two-phase simulations and long time scales that are required to fully capture the mixing, affordable due to the speed and accuracy of the MBX software, are assessed. Constructing the methane–water equation of state across the phase diagram shows that the stable mixtures are denser than the sum of their parts at a given pressure and temperature. We find that many-body polarization plays a central role, enhancing the induced dipole moments of methane by 0.20 D during mixing under pressure. Overall, the mixed system adopts a denser state, which involves a significant enthalpic driving force as elucidated by a systematic many-body energy decomposition analysis.

I. INTRODUCTION

Hydrophobic interactions are one of the most fundamental forces in biophysically relevant processes such as protein folding and protein-protein interactions in water.1 Due to their collective and entropic nature, understanding the microscopic and molecular origins of these forces has been extremely challenging.2 In this regard, many theories have been developed and experiments conducted to understand how hydrophobic interactions change as a function of temperature and pressure.3

One of the apparent hallmarks of hydrophobic interactions is their characteristic length-scale dependent thermodynamics.2 Small hydrophobic solutes such as methane do not lead to a disruption of the water hydrogen-bond network, while for solutes with diameter larger than 1 nm, interfacial effects involving surface tension begin to play a more critical role.5 Perhaps the most well studied hydrophobic interaction is that involving two methane molecules in solution.4,9 Several molecular dynamics (MD) simulations have examined the potential of mean force (PMF) between two methane molecules in water using empirical force fields.4,6–11 These studies have found that the contact pair is thermodynamically stabilized compared to the shallower solvent separated minimum. This confirms our common experience of water and oil not mixing.

An intriguing set of experiments have shown that under very high pressures (larger than 1.0 GPa), it is possible to observe mixing of oil and water.12 The microscopic origins of this phenomenon have remained elusive. Attempts to use empirical force fields in MD simulations did not prove to be successful at reproducing the experimental data.13 Ab initio molecular dynamics (AIMD) simulations based on density functional theory (DFT) recently showed that the polarization of methane molecules at high pressures is a critical ingredient needed to observe high-pressure mixing.14 The effect of this electronic polarization is even more enhanced at pressures above 15 GPa and much higher temperatures.15

The notion that there are subtle electronic effects underly ing the interactions between simple hydrocarbons and water, has long been appreciated. The possibility of creating weak hydrogen bond-like interactions between -CHn groups and water has been observed in several biological contexts16,17 and has been the subject of numerous electronic structure investigations.18 The possibility that electronic structure properties play a role in stabilizing interactions between -CHn groups and water has also recently been invoked to explain the presence of the apparent negative charge observed from electrophoretic experiments.19 However, the precise role of these interactions in macroscopic thermodynamics, remains an open problem.

a)Electronic mail: fpaesani@ucsd.edu
b)Electronic mail: ahassana@ictp.it
In the last decade, there has been a tremendous growth in our understanding of the importance of many-body interactions in accurately modeling phenomena associated with water hydrogen-bond networks. In particular, the data-driven many-body MB-pol potential, which is designed to reproduce the many-body expansion of the energy with coupled-cluster accuracy, has been shown to correctly reproduce several thermodynamic, dynamic, and spectroscopic properties across the water phase diagram. Building on the accuracy of MB-pol, data-driven many-body potentials, called MB-nrg potentials, have been developed for modeling CO$_2$/H$_2$O and CH$_4$/H$_2$O mixtures from the gas to the condensed phase. The MB-nrg potentials combine explicit data-driven representations of short-range low-order many-body interactions with an implicit representation of many-body effects based on classical polarization. The MB-nrg potentials for neat methane and methane–water mixtures were shown to accurately predict both structure and energetics of small clusters, the second virial coefficients, and the structure of both liquid CH$_4$ and dilute CH$_4$/H$_2$O mixtures. In this work, the methane–water MB-nrg potential is used in combination with the MB-pol potential for water to examine the mixing of methane and water under pressure.

The recent implementation of the MB-nrg potentials in the MBX software enables simulations across a wide range of the phase diagram at a high level of accuracy and for much longer timescales and larger systems sizes than affordable by AIMD simulations. We begin by constructing the equation of state (EOS) of the methane–water system and find that the mixtures are denser upon mixing. Before examining the collective process of mixing, we determine, using umbrella sampling, the one dimensional PMF between two methane molecules in water at ambient and elevated pressure. Our results are consistent with previous polarizable and classical models. Mixing at higher pressures involves an enhancement of methane’s electronic polarization while leaving water largely unperturbed, along with an enthalpic driving force for the mixed solvated system. We also conduct an energy decomposition analysis of the many-body interaction contributions and illustrate the important role of many-body polarization in mixing under pressure.

II. METHODS

A. Molecular Dynamics Simulations

The accuracy and predictive power of the MB-nrg potentials have been demonstrated in simulations of ion hydration and various fluid mixtures, including CO$_2$-water, ammonia-water, and methane-water mixtures. The MB-nrg potentials are implemented in the open-access MBX software, which is interfaced with LAMMPS and i-PL. The MD simulations of the CH$_4$/H$_2$O mixtures were carried out with LAMMPS in the isobaric-isothermal ensemble using the MB-nrg potential of Ref. 26.

The CH$_4$/H$_2$O MB-nrg potential was rigorously derived from the many-body expansion of the interaction energies calculated at the coupled cluster level of theory including single, double, and perturbative triple excitations, i.e., CCSD(T), the current “gold standard” for molecular interactions. The CH$_4$/H$_2$O MB-nrg potential includes a data-driven representation of 2-body interactions in terms of permutationally invariant polynomials (PIPs) along with an implicit representation of N-body interactions in terms of classical polarization. The CH$_4$/H$_2$O MB-nrg potential was shown to accurately reproduce the second virial coefficient of both neat CH$_4$ and the CH$_4$/H$_2$O mixture, the interaction energies of small (CH$_4$)$_n$ and (CH$_4$)$_n$(H$_2$O)$_m$ clusters, and the density and structure of liquid methane at relatively low temperature and pressure. Specific details about the theoretical and computational framework adopted for the development of the CH$_4$/H$_2$O MB-nrg potential along with a systematic assessment of its accuracy can be found in the original study.

In the present MD simulations carried out at various thermodynamic state points, the geometries of the CH$_4$ and H$_2$O molecules were kept fixed at their average geometries determined from corresponding (short) MD simulations carried out with flexible molecules at the same temperature and pressure conditions. The molecular dipole moments were calculated by summing the permanent and induced components. To characterize the mixing process, simulations were performed with 10, 20, 30, 40, 50, and 60% methane content in a phase separated setup with up to 570 molecules. Each concentration was simulated in the NPT ensemble at pressure and temperature conditions between 0.5-3.0 GPa and 400-500 K using a 2 fs time step. Simulations for the 10% mixture were run for 1 ns while higher concentrations were run for 2 ns in order to converge the mixed state.

Simulations of the 30% mixture were extended up to 10 ns to investigate whether complete mixing could be achieved over a long timescale. In total, around 400 ns of simulation was performed. Equilibrium quantities (e.g., density) were averaged over the final 0.1 ns of each simulation. Coordination numbers were calculated by integrating up the first minima in a given pair distribution function (PDF) or by counting the average number of neighbors up to a distance cutoff defined by the first PDF minima. Time resolved PDFs along with the final snapshots from the MD simulations illustrating the state of mixing are given in the supplementary information (SI Fig. 3). For comparison, simulations at the same pressure and temperature conditions were performed for 500 ps for two neat systems of CH$_4$ and H$_2$O, each containing 256 molecules.
B. Decomposition of the MB-nrg Energy

The MB-nrg energy, $E_{\text{TOT}}$, is defined as:

$$E_{\text{TOT}} = E_{\text{IB}} + E_{\text{PERM}} + E_{\text{POL}} + E_{\text{DISP}} + E_{\text{PIP}}$$  \hspace{1cm} (1)

where the first term describes the 1-body (1B) monomer distortion energies (which is constant for rigid monomer simulations), the next three terms describe the long-range interactions corresponding to permanent electrostatics, polarization, and dispersion, and $E_{\text{PIP}}$ corresponds to PIPs describing close-range interactions such as Pauli repulsion, while also correcting deficiencies in the other terms, for example charge penetration effects that are not well described by $E_{\text{PERM}}$ alone. In the water potential (MB-pol), $E_{\text{PIP}}$ includes a PIP for both 2-body (2B) water-water and 3-body (3B) water–water–water interactions, while the methane and water–methane potentials include 2B methane–methane and water–methane PIPs.

In order to estimate the energetic driving forces in the simulated mixing processes, $\Delta E$ values for each interaction term were calculated for the simulations of the 10% and 20% mixtures at 400 K at each pressure. The first 1 ps of each simulation was discarded as the initial relaxation time period prior to mixing. The average of each energy term over the next 25 ps (in 0.5 ps increments) was taken to be the initial energy value, the corresponding average over the last fifth of each simulation was taken to be the final energy value, and $\Delta E$ was taken to be the difference. We emphasize that the reported $\Delta E$ values are not equilibrium properties due to limited sampling (particularly in the initial energy), as well as a lack of equilibration of the pre-mixed states of the system.

Further decomposition into subsystem energies allows for the total system energy to be partitioned into $\text{H}_2\text{O}$ interactions, $\text{CH}_4$ interactions, and $\text{H}_2\text{O}–\text{CH}_4$ interactions. Subsystem energies are explicitly defined by the MB-nrg potential for all terms except for $E_{\text{POL}}$ which is a fundamentally many-body interaction, so $E_{\text{POL}}$ must be evaluated for $\text{H}_2\text{O}$ and $\text{CH}_4$ separately (for example by setting the charges and polarizabilities of the other species to 0), and then the cross term ($\text{H}_2\text{O}–\text{CH}_4$), describing the mutual polarization between $\text{H}_2\text{O}$ and $\text{CH}_4$, is taken to be the remainder after subtracting the $\text{H}_2\text{O}$ and $\text{CH}_4$ $E_{\text{POL}}$ energies from that of the total system.

C. $\text{CH}_4–\text{CH}_4$ potential of mean force

The $\text{CH}_4–\text{CH}_4$ PMF at 298.15 K and 1 atm was calculated from umbrella sampling MD simulations carried out for 2 $\text{CH}_4$ molecules dissolved in a box containing 256 $\text{H}_2\text{O}$ molecules. A harmonic bias (umbrella potential) with a force constant of 5.0 kcal/mol/Å$^2$ was applied to 29 equally spaced $\text{CH}_4–\text{CH}_4$ distances, between 2.5 Å and 9.5 Å with an interval of 0.25 Å. The simulations with the MB-nrg potential were carried out with a time step of 0.2 fs using the i-PI software\textsuperscript{41} using flexible $\text{CH}_4$ and $\text{H}_2\text{O}$ molecules. For comparison, analogous simulations were carried out with LAMMPS using the OPLS force field for methane\textsuperscript{45} combined with the TIP4P/2005 water model.\textsuperscript{46} In this case, the equations of motion were integrated with a time step of 2 fs since both the OPLS and TIP4P/2005 models consider rigid methane and water molecules. In both i-PI and LAMMPS simulations, PLUMED\textsuperscript{47} was used to apply the umbrella potential, and each umbrella window was sampled for 500 ps. The actual PMF was then calculated using the weighted histogram analysis method (WHAM)\textsuperscript{48,49} as implemented in the open-access WHAM software.\textsuperscript{50}

III. RESULTS

A. Equation of State

When a binary mixture forms under pressure, it is expected to be denser than the sum of its parts\textsuperscript{55} due to the significant PV contribution to the free energy, though reducing the internal energy and maximizing the entropy of the system are also key factors. The equations of state (EOS) for methane-water mixture as well as for the pure $\text{CH}_4$ and $\text{H}_2\text{O}$ liquids derived from the MB-nrg simulations are shown in Fig. 1 along with the corresponding experimental data.\textsuperscript{51,52} The available experimental data for methane at 405 K, and water at 373 K and 423 K data, which were interpolated to 405 K, were fitted to the logarithmic function $a \log(x/b) + 1 + c$ which was then used to extrapolate to lower and higher pressures (solid blue and red lines). The MB-nrg EOS can be compared with an approximate EOS that assumes a linear mixture combination of pure water and methane according to:

$$1/\rho_{\text{mix}} = x_{\text{CH}_4}/\rho_{\text{CH}_4} + x_{\text{H}_2\text{O}}/\rho_{\text{H}_2\text{O}}$$  \hspace{1cm} (2)

where $x_{\text{CH}_4}$ and $x_{\text{H}_2\text{O}}$ are the mass fractions of methane and water, respectively. The results obtained with the MB-nrg potential of methane and the MB-pol potential of water can be compared with the available experimental data, which is noted to have an uncertainty of ±5% from speed of sound measurements. We find that MB-pol at 400 K displays excellent agreement with the interpolated experimental data, as well as the extrapolation from 1.0 to 0.25 GPa. This level of agreement provides confidence for the reliability of the simulation results at 450 K and 500 K. For methane, the MB-nrg potential predicts a slightly denser liquid than observed in experiment, although the simulation results are within the experimental uncertainty.

In the upper part of Fig. 1 we find that the 400 K mixture simulated with the MB-nrg/MB-pol potentials is generally much denser than the experimental linear mixture density at 405 K (up triangle symbols are always above the dotted lines). This is expected for a mixture stabilized at high pressure, though this is possibly enhanced as MB-nrg methane is denser than the EOS of Li
et al.\textsuperscript{52} as seen in the pure systems. Taking a linear mixture of the pure systems described by the MB-nrg and MB-pol potentials, we can compare the mixture densification for one model, which is shown in the lower panel of Fig. 1. We find that the mixed system is always denser than the sum of its parts for the same temperature and pressure.

We note that at concentrations of 10 and 20\% the systems are fully mixed and are thus expected to have the greatest relative density increase as seen in Fig. 1. For the 30-60\% mixtures, the systems partially mix within 2 ns (see snapshots in the SI Fig. 3). This is also expected given the results for the lower concentration, and so there is still a partial increase in the density. If there was no mixing, then we would expect the same or a lower density than the pure systems. Finally, a refined EOS can be calculated by starting from a mixed condition and equilibrating, improving data for the 30 and 40\% mixtures. The mixture EOS is in consistent with DFT results for a 14\% mixture\textsuperscript{14}.

We also performed phase-separated slab simulations at ambient pressure and 300 K and found that the system does not mix, while the methane region expands, as expected at these conditions.

### B. Methane-Methane PMFs

The interaction of two methanes in water has been the subject of numerous theoretical and numerical studies.\textsuperscript{4,6–11} Specifically, the thermodynamics associated with this process has been a classical model system for understanding hydrophobic interactions in water that is relevant for understanding biophysical phenomena such as protein folding\textsuperscript{2,54} and protein aggregation\textsuperscript{55}.

![FIG. 1.](image1.png)

**FIG. 1.** (Top) Methane-water equation of state simulated with MBX averaging the density over the final 0.1 ns assuming a mixed state was achieved (solid lines). Interpolated experimental data to 405 K from (Ab. and Li.)\textsuperscript{51,52} for water and methane is included for comparison (red and blue curves), as well their linear mixture density from 10-60\% (coloured dotted lines). (Bottom) Percentage of density increase compared with the MBX linear mixing density.

![FIG. 2.](image2.png)

**FIG. 2.** PMFs for methane-methane with MBX and OPLS-AA at various pressures, as well as literature data from\textsuperscript{8,9}. PMFs were normalized by setting zero to the PMF value at greatest value of $r$.

Although the bare van der Waals interaction between two methane molecules has a very weak attraction arising from dispersion forces, the PMF in water exhibits more structure. The physical origin of this arises from the
water-reorganization, which makes it more entropically favorable to reduce the solvent exposed surface area when the two methanes are in contact. Specifically, several simulations using classical non-polarizable potentials\textsuperscript{4,10} have shown that the PMF between two methanes is characterized by a contact minimum (CM) and a solvent separated minimum (SSM). The CM involves direct contact between the methanes, while the SSM involves water molecules that form a hydrogen-bonded network around two solvated methanes.

Hummer and co-workers examined how these 1-dimensional PMFs evolve as a function of pressure.\textsuperscript{11} Using the OPLS-AA and SPC/E models for methane and water, respectively, they conducted umbrella sampling simulations varying the pressure between -16 and 725 MPa. Increasing the pressure within this range stabilized the solvent separated species, consistent with the observation of pressure induced unfolding in proteins.\textsuperscript{56} However, it should be stressed that the 1-dimensional PMFs do not directly probe collective effects underlying biophysical processes.

Over a decade ago, Car and co-workers re-visited the PMFs at ambient pressure using DFT-based AIMD simulations.\textsuperscript{8} Contrary to results obtained from previous simulations with empirical potentials, the AIMD simulations did not show an SSM, but instead presented a much deeper attractive CM. These results suggested that the empirical potentials could be significantly underestimating the hydrophobic attraction between hydrocarbons.

With the many-body MB-nrg potential in hand describing the interactions between methane and water, we first set out to examine the 1-dimensional PMF between two methane molecules in a box of water molecules. Constrained MD simulations were conducted using umbrella sampling, from which the unbiased free energies were extracted using WHAM. In addition, we also repeated our calculations with the OPLS-AA potential and TIP4P/2005 water model. All these simulations were performed at 300 K and at two pressures, namely 1 atmosphere and 0.5 GPa.

Fig. 2 shows the PMFs obtained with the MB-nrg potential of methane and the MB-pol potential of water (panel a), and the OPLS-AA model of methane and the TIP4P/2005 model of water (panel b). The MB-nrg/MB-pol PMFs at ambient pressure are very similar to those obtained with the OPLS-AA/TIP4P/2005 models and consistent with previous results obtained with empirical models.\textsuperscript{4,6–11} Moving towards the higher 0.5 GPa pressure appears to increase the stability of the solvent separated state for both models consistent with previous studies by Hummer and co-workers.\textsuperscript{11}

The PMFs obtained with the MB-nrg/MB-pol potentials are quite different from those obtained by Car and co-workers, as shown more clearly in the bottom panel (panel c) of Fig 2. These earlier AIMD studies were performed with the PBE functional without dispersion corrections\textsuperscript{27} and therefore likely created a much more structured water hydrogen bond network which over-stabilizes the hydrophobic effect. Recently, Akin-Ojo and co-workers have developed polarizable models for methane and water which were used to calculate the methane-methane PMF.\textsuperscript{9} The MB-nrg/MB-pol results closely track the results obtained in Ref. 9. As we will see in the following sections, although many-body polarization does not appear to be so critical at ambient pressure, it is clearly important to reproduce the high pressure mixing as observed in some of our previous theoretical studies discussed in the introduction.\textsuperscript{14,15}

C. Pressure Dependence of Methane-Water Mixing

The mixing of methane and water in simulations can be quantified by several pair coordination numbers namely carbon-oxygen (CO), carbon-carbon (CC), and oxygen-oxygen (OO). Earlier DFT-based simulations from some of us showed that achieving mixing when starting from phase separated initial conditions requires rather long simulations to overcome the kinetic barrier that separates the unmixed and mixed states. Here we examine the mixing convergence with time and summarize the mixed state’s coordination obtained from simulations with the MB-nrg potentials.

The coordination numbers as a function of time are shown in Fig. 3 which highlights the convergence of the mixing process for the 10 and 20% mixtures. The coordination numbers, most noticeable for the CO and CC pairs, appear stable after 0.5 and 1.0 ns for the 10% and 20% mixture, respectively. Although the system sizes simulated here is almost double that of our previous DFT-based AIMD simulations,\textsuperscript{14} the MB-nrg simulations clearly show that the equilibration timescales for the mixing process takes at least 0.5 ns.

The first-shell coordination numbers, calculated by integrating up to the first minimum of each PDF, are shown for the mixed and pure simulations in Fig. 4. These plots were constructed using the last 0.1 ns of the trajectory, removing the first parts as equilibration (Fig. 3). As a reference, the top panel of Fig. 4 shows the coordination numbers for the pure systems which are consistent with previous results.\textsuperscript{14,58} For both the 10 and 20% mixtures (excluding 0.5 GPa) we find little temperature dependence on the pair coordination numbers. Compared with the data extracted from the 14% methane neutron diffraction experiments (shown by stars), the MB-nrg CO value for the 10% mixture is greater and the 20% mixture is similarly lower, as expected. In general, all the coordination numbers between 1.5 and 2.5 GPa are in good agreement with the available experimental data collected at 413 K.

The middle panel of Fig. 4 shows that simulations with the MB-nrg/MB-pol potentials predict a mixing temperature dependence at 0.5 GPa, with the system at 500 K being the most mixed (greatest CO coordination number) and the system at 400 K being the least mixed. We also note that the coordination numbers as a function
of pressure show a slight pressure dependence for the 10 and 20% mixture, also observed in experiments carried out for a 14% mixture.

D. Electronic Properties

While mixing of the methane–water system is driven by both volume reduction (PV) and the entropy of mixing (TS), changes to the molecular interactions (U) are also relevant. These have been previously probed\textsuperscript{14,15} showing an enhanced dipole moment for methane and a band gap closure at extreme temperatures and pressure. Here we examine the induced dipoles of the pure and mixed systems, a decomposition of the energetics, and their time evolution with the MB-nrg potential.

The dipole moment time evolution is shown in Fig. 5 as the system mixes. The water dipole remains mostly constant, while that of methane is clearly growing as molecules migrate into the water solution. It is interesting to note that the methane dipoles continue to increase on a much slower timescale even after mixing has been achieved (black dashed lines in Fig. 5). This indicates that there are clearly some very slow timescales on the order of several nanoseconds, likely involving the reorganization of the water hydrogen-bond network and the packing of the solvated methane molecules. These long timescales demonstrate the challenges of using AIMD simulations to explore the free energy landscapes of mixing phenomena. The hydrogen-bonding as a function of time is also shown in the SI.

In a previous study using AIMD simulations, we showed that despite the mixing of methane under pressure, the water hydrogen-bond network essentially remains intact. In other words, there is no significant disruption in the hydrogen-bond structure. In solid methane-water clathrates\textsuperscript{14} it has been shown that water forms a complete hydrogen-bond network involving water molecules that obey the ice rules. These water molecules surround the guest methane molecules forming closed ring structures.\textsuperscript{14,60} In the top panel of Fig. 6 we find that the hydrogen-bond network in the mixture is largely maintained with respect to the pure water system at the same pressure, e.g. 96-97% and 93-94% per water molecule in the 10% and 20% mixtures respectively, in agreement with a previous result of 97%.\textsuperscript{14}

The induced molecular dipoles for methane and water are shown in the lower and middle panels of Fig. 6 as a function of pressure. Though the molecules in the present simulations with the MB-nrg/MB-pol potentials are treated as rigid, the atom-centered point-dipole po-

FIG. 3. Coordination numbers (CO,CC,OO) as a function of time as the slab configuration mixes at 400 K, 2 GPa, and 10% concentration (above) and 500 K, 3 GPa, and 20% concentration (below). A fixed radial cut-off was used as the g(r) minima position is almost constant.

FIG. 4. Summary of coordination numbers as a function of P,T shown for fully mixed concentrations (10,20%), as well as pure systems (above), from the final 0.1 ns of each simulation. Stars denote coordination numbers for a 0.14 mixture at 413 K with neutron diffraction experiments\textsuperscript{14}.  
larizabilities built into the MB-nrg models\textsuperscript{26} allow for many-body polarization through the induction of molecular dipole moments. For water, we find that the dipole moment increases with pressure as expected, although less-so than that of pure water at 400 K, and this difference is enhanced with increasing methane content. This is likely connected to the slightly lower number of hydrogen bonds in the mixed system. We find that methane has a dipole enhancement of around 0.2 D in the mixture compared to the pure system at a given pressure, noting that the 20\% mixture at 0.5 GPa has not fully mixed.

![Graph showing average induced molecular dipoles for CH\textsubscript{4} and H\textsubscript{2}O as a function of time for 0.10 (left) and 0.20 concentration (right) simulation at 3 GPa and 400 K. The moving average is over 50 points (solid lines) and the dashed black lines indicate the time of converged coordination numbers.]

**FIG. 5.** Average induced molecular dipoles (blue line) for CH\textsubscript{4} and H\textsubscript{2}O as a function of time for the 0.10 (left) and 0.20 concentration (right) simulation at 3 GPa and 400 K. The moving average is over 50 points (solid lines) and the dashed black lines indicate the time of converged coordination numbers.

The observed changes in the dipole moment of both the methane and water molecules as a function of mixing as well as the significant changes in the induction energy for the two molecules motivated us to examine in more detail the various energetic contributions to the many-body energy as a function of pressure. The dependence of \( \Delta E_{POL} \) and \( \Delta E_{TOT} \) on the pressure is shown in Figure 7, and all other interaction terms are shown in the SI. Time series for each energy term are also reported in the SI.

**FIG. 6.** Average induced molecular dipoles and hydrogen-bonds for the mixed (dashed lines) and neat (solid lines) CH\textsubscript{4} and H\textsubscript{2}O systems as a function of pressure and temperature. Hydrogen-bonds we calculated with a 3.5 Å cutoff and 30° as originally proposed by Luzar and Chandler\textsuperscript{61}.

**E. Energy Decomposition**

The observed changes in the dipole moment of both the methane and water molecules as a function of mixing as well as the significant changes in the induction energy for the two molecules motivated us to examine in more detail the various energetic contributions to the many-body energy as a function of pressure. The dependence of \( \Delta E_{POL} \) and \( \Delta E_{TOT} \) on the pressure is shown in Figure 7, and all other interaction terms are shown in the SI. Time series for each energy term are also reported in the SI.

Fig. 7 shows that the trends in the \( \Delta E \) values tend to be monotonic with respect to pressure for the mixing systems (all pressures above 0.5 GPa), with the exception of H\textsubscript{2}O \( \Delta E_{POL} \) in the 10\% system at 2 GPa, which demonstrates the uncertainty in our \( \Delta E \) estimation. The H\textsubscript{2}O and H\textsubscript{2}O–CH\textsubscript{4} \( \Delta E_{POL} \) values tend to decrease with increasing pressure, while the CH\textsubscript{4} \( \Delta E_{POL} \) increase with increasing pressure. Due to the use of rigid monomers, the nonpolar CH\textsubscript{4} molecules only weakly polarize each other, so the CH\textsubscript{4} induced dipoles (Figure 6) are mostly due to their interactions with the surrounding H\textsubscript{2}O molecules, which corresponds to the H\textsubscript{2}O–CH\textsubscript{4} \( E_{POL} \) term. Furthermore, the total system \( \Delta E_{POL} \) is dominated by H\textsubscript{2}O and H\textsubscript{2}O–CH\textsubscript{4} contributions, and therefore follows the same pressure dependence. In all simulations, the \( E_{POL} \) energy remains negative for all subsystem contributions (see SI), and therefore the positive \( \Delta E_{POL} \) corresponds to less attractive \( E_{POL} \), while negative changes correspond to an enhancement of attractive polarization interactions.

For all energy terms in the MB-nrg binding energy decomposition (Eq. 1), the 20\% mixtures apparently have stronger pressure dependence than those for the 10\% mixtures (i.e., larger slopes in \( \Delta E \) plotted against pressure). This difference is not surprising since compositions with more CH\textsubscript{4} molecules lead to larger structural changes as a result of mixing (notably, a larger disruption of the water hydrogen-bond network, as can be seen in the time series of average number of water hydrogen...
bonds shown in the SI).

Surprisingly, the H$_2$O $\Delta E_{\text{POL}}$ values are mostly negative for the 10% mixtures, and positive for the 20% mixtures. Mixing results in a decrease in the average number of water hydrogen bonds for both concentrations (see Figure 33 in the SI), so negative H$_2$O $\Delta E_{\text{POL}}$ values for the 10% mixtures indicate that the H$_2$O polarization energy becomes more attractive despite a decrease in the number of hydrogen bonds. This suggests that the mixing process in the 10% mixtures results in subtle structural changes (that cannot be seen in a simple average of the number of hydrogen-bonds) optimizing the H$_2$O polarization. The 20% mixtures show the expected behavior of decreasing numbers of hydrogen bonds (due to mixing) resulting in decreasing attractive polarization interactions (positive H$_2$O $\Delta E_{\text{POL}}$). The mixing process reduces the number of hydrogen bonds in the 20% mixtures more than in the 10% mixtures (see Figure 33 in the SI) which probably overshadows any optimization of the H$_2$O polarization energy occurring in the 20% mixtures.

The largest contributions to the changes in the binding energies, $\Delta E_{\text{TOT}}$, tend to be $\Delta E_{\text{DISP}}$ and $\Delta E_{\text{PIP}}$ (SI), with $E_{\text{DISP}}$ being an exclusively attractive interaction, and $E_{\text{PIP}}$ mostly corresponding to repulsive close-range interactions, so $\Delta E_{\text{TOT}}$ for the most part results from the balancing of these two competing terms.

Mixing results in a larger number of H$_2$O–CH$_4$ contacts, so negative H$_2$O–CH$_4$ $\Delta E_{\text{TOT}}$ values are mostly due to $\Delta E_{\text{DISP}}$ overpowering $\Delta E_{\text{PIP}}$. Mixing also results in fewer H$_2$O–H$_2$O and CH$_4$–CH$_4$ contacts, so positive $\Delta E_{\text{TOT}}$ values are similarly due to $\Delta E_{\text{DISP}}$ overpowering $\Delta E_{\text{PIP}}$. Notably, the 10% mixtures at high pressures have attractive H$_2$O and CH$_4$ $\Delta E_{\text{TOT}}$ values, likely due to the same structural optimizations reflected in the attractive $\Delta E_{\text{POL}}$ in the same systems.

All in all, the changes in the dipole moments of both methane and water as well as the importance of the polarization energy as a function of pressure confirm the need to go beyond the standard empirical potentials which only include permanent electrostatics.

**IV. CONCLUSIONS**

In conclusion, in this work we have used the recently developed MB-nrg potential to examine the mixing of oil and water under pressure. We have examined the EOS of the dense methane–water system and find that the mixture to be always denser than the sum of its parts, resulting in an enthalpic driving force to the high pressure mixing. Re-examining the one-dimensional PMF associated with two methane molecules in water, we find that the MB-nrg potential is consistent with both both empirical potentials as well as polarizable models.

Overall, the results obtained with the MB-nrg potential are consistent with previous DFT based AIMD simulations. However, we note that particularly at higher concentrations of methane, there are significant kinetic barriers involving collective reorganization of both the water and methane. Thus, the mixing phenomena takes place on the timescale of nanoseconds. The simulations here thus required long timescales and large system sizes, which go beyond what was possible with *ab initio* methods.

Finally, we also demonstrate an enhancement in both the polarizability of water and methane under pressure, which were found to be mediated by methane–water interactions. This is also reflected in the energy decomposition analysis, where we find an important role of many-body polarization interactions of both the methane and water molecules. These effects appear to become more enhanced at high pressures and likely have important implications for simulating biological systems.

**FIG. 7.** $\Delta E_{\text{POL}}$ as a function of pressure for (a) H$_2$O, (b) CH$_4$, (c) H$_2$O–CH$_4$, and (d) the total system, and $\Delta E_{\text{TOT}}$ as a function of pressure for (e) H$_2$O, (f) CH$_4$, (g) H$_2$O–CH$_4$, and (h) the total system, as described in the main text.
Our results also provide further evidence for the accuracy and predictive power of the MB-nrg potentials which may open the door to quantitatively reliable studies of the thermophysical properties of fluid mixtures and underlying molecular mechanisms over a wide range of thermodynamic conditions.

V. SUPPLEMENTARY MATERIAL

The supplementary materials include time evolved PDFs for all mixing simulations, final snapshots of selected mixtures, and time evolved energy decomposition data.

VI. ACKNOWLEDGMENTS

R.G., M.R., and F.P. were supported by the U.S. Department of Energy, through Award no. DE-SC0019490. Computing resources were provided by the National Energy Research Scientific Computing Center (NERSC), which is supported by the Office of Science of the U.S. Department of Energy under Contract DE-AC02-05CH11231, the Triton Shared Computing Cluster (TSCC) at the San Diego Supercomputer Center (SDSC), and CINECA for computing resources on Marconi. C.K. was supported by the Argonne Leadership Computing Facility, which is a U.S. Department of Energy Office of Science User Facility operated under contract DE-AC02-06CH11357.

VII. AUTHOR DECLARATIONS

The authors have no conflicts to disclose.

VIII. DATA AVAILABILITY STATEMENT

MD inputs and trajectories, analysis scripts, and all results presented in this study are available from the authors upon request.


IX. GOVERNMENT LICENSE

The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory (“Argonne”). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan. http://energy.gov/downloads/doe-public-access-plan