POLITECNICO DI TORINO
Repository ISTITUZIONALE

Ephemeral ice-like local environments in classical rigid models of liquid water

Original
Ephemeral ice-like local environments in classical rigid models of liquid water / Capelli, R.; Muniz-Miranda, F.; Pavan, G. M.. - In: THE JOURNAL OF CHEMICAL PHYSICS. - ISSN 0021-9606. - ELETTRONICO. - 156:21(2022), p. 214503. [10.1063/5.0088599]

Availability:
This version is available at: 11583/2968499 since: 2022-07-20T15:26:32Z

Publisher:
American Institute of Physics Inc.

Published
DOI:10.1063/5.0088599

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
AIP postprint/Author's Accepted Manuscript e postprint versione editoriale/Version of Record

(Article begins on next page)

10 August 2024

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

The Journal of Chemical Physics

Ephemeral Ice-Like Local Environments in Classical Rigid Models of Liquid Water

Riccardo Capelli,[1, a)] Francesco Muniz-Miranda,[2] and Giovanni M. Pavan[1, 3, b)]

[1)] *Department of Applied Science and Technology, Politecnico di Torino, Corso Duca degli Abruzzi 24, I-10129 Torino, Italy*

[2)] *Department of Chemical and Geological Sciences, University of Modena and Reggio-Emilia, Via Campi 103, I-41125 Modena, Italy*

[3)] *Department of Innovative Technologies, University of Applied Sciences and Arts of Southern Switzerland, Polo Universitario Lugano, Campus Est, Via la Santa 1, CH-6962 Lugano-Viganello, Switzerland*

(Dated: 16 April 2022)

AIP Publishing

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

The Journal of Chemical Physics

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

Despite great efforts over the past 50 years, the simulation of water still presents significant challenges and open questions. At room temperature and pressure, the collective molecular interactions and dynamics of water molecules may form local structural arrangements that are non-trivial to classify. Here we employ a data-driven approach built on Smooth Overlap of Atomic Position (SOAP) that allows us to compare and classify how widely used classical models represent liquid water. Macroscopically, the obtained results are rationalized based on water thermodynamic observables. Microscopically, we directly observed how transient ice-like ordered environments may dynamically/statistically form in liquid water, even above the freezing temperature, by comparing the SOAP spectra for different ice structures with those of the simulated liquid systems. This confirms recent *ab initio*-based calculations, but also reveals how the emergence of ephemeral local ice-like environments in liquid water at room conditions can be captured by classical water models.

[a] Electronic mail: riccardo.capelli@polito.it
[b] Electronic mail: giovanni.pavan@polito.it

AIP Publishing

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

The Journal of Chemical Physics

AIP Publishing

I. INTRODUCTION

Water is one of the most ubiquitous substances on the Earth's surface. Yet, a complete understanding of its structural properties, fundamental for its role in Life[1,2], remains out of reach. In fact, the molecular arrangements observed in different phases of water are extraordinarily various and complex, resulting in a plethora of unusual and often not yet fully understood properties.[3–5] From a microscopic point of view, such a complexity arises from a concerted action of directional and strong hydrogen bonds and non-specific Van der Waals interactions, which are in dynamic competition with the collective motion of the water molecules at finite temperature regimes. Traditionally, the minimal picture to interpret the 1st solvation shell shows each water molecule in condensed phase engaged in 4 hydrogen bonds with its immediate neighbors, yet in the last decades even this basic understanding has been challenged.[6–8] New interpretations have been proposed to explain water's less-common phases[9–11], showing how its detailed phase diagram is still subject to intense debate, both from an experimental[12,13] and computational perspective.[14,15]

Indeed, achieving a satisfactory accuracy in the modeling of liquid water is a major challenge. Simulation approaches at a high level of theory, *e.g.*, those treating the electronic structure at the quantum-mechanical level (usually in its density-functional theory approximation) can accurately describe the covalent interactions[16–18] and hydrogen-bond dynamics[19,20] of water molecules. Experimental properties such as the O...H radial distribution function and other observed behaviors[21–24] also require further corrections to account for nuclear quantum effects to be perfectly matched. In general, water has the tendency to display large local fluctuations[25,26], a fact that has consequences in phenomena such as supercooling[27–29] or in the emergence of many metastable "phases".[9,30] Nonetheless, the high computational cost associated to these accurate quantum approaches to describe such phenomena typically hinders

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

The Journal of Chemical Physics

AIP Publishing

the possibility to achieve the ergodic limit for large samples of liquid water systems.

The need of a reasonable compromise between computational cost and accuracy of the results (especially for biological applications) paved the way to the design of many classical force field potentials. Differently from the *ab initio* approach, standard classical potentials lack polarization and treat molecules as rigid bodies to allow for longer time integration steps. With this kind of limitations in the system representation, the model performances are assessed on the comparison with known specific properties (obtained from experiments or from higher level of theory calculations). In general, classical models for water would be expected to be accurate in the treatment of the liquid phase at room conditions, but sensible differences between models can still be observed and their development/optimization is still a very active research field.

Historically, from the first simpler and still widely used 3-sites rigid water models (*i.e.*, one site per atomic nucleus) developed by Berendsen *et al.*[31] and by Jorgensen *et al.*[32], the water models evolved to more complex representations, including one, two or even more virtual sites (see Figure 1 and Table I) to have an accurate accounting of effects that cannot be obtained *via* classical potentials acting on only 3 nuclei. Indeed, the necessity of at least one virtual site in classical description of water for a qualitative reproduction of the phase diagram has been extensively demonstrated in previous works.[33,34]

Differences in the features of the water models usually may lead to large discrepancies in both the collective behavior and in the measured properties of water. But a direct correlation between microscopic parameters and macroscopic observables is not always straightforward in the simulation of systems containing a large number of explicit molecules. For this reason, the detection of ordered arrangements in liquid water models has been attempted using, *e.g.*, order parameters[35], graph theory[26], or environment descriptors[36]. In an interesting recent study, using a machine learning-

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT
The Journal of Chemical Physics
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599
AIP Publishing

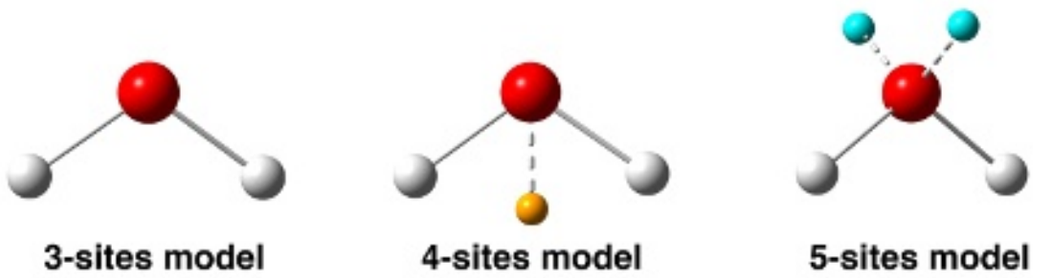


FIG. 1. Rigid classical models for the water molecule, including 3 point-charges (left), 4 point-charges (center), and 5 point-charges (right). Charges not coinciding with atomic positions are displayed in orange (center) and cyan (right).

based potential of water trained on *ab initio* calculations, Monserrat *et al.* employed a local environment analysis which highlighted the presence of structures described as "building blocks of ice phases" in liquid water.[36] A relevant question is whether similar ice-like precursor motifs can be encountered in liquid water also by classical water models, especially considered the importance of this in understanding the behavior of water and the central role played by classical models in nowadays molecular simulations.

Here we apply a data-driven approach based on Smooth Overlap of Atomic Position (SOAP) distances, which allow us to compare molecular systems and how different force fields treat them: some of us previously adopted it to classify lipid systems and classical lipid force fields[37] and to track defects and dynamic molecular arrangements in complex supramolecular systems[38,39]. In the present paper, we employ this approach to compare the local fluctuating molecular environments that are formed in molecular dynamics simulations of liquid water, and to classify them based on their SOAP similarity with known ice environments. Macroscopically, we compared 16 of the most widely used rigid water models in liquid state and we clas-

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

The Journal of Chemical Physics

AIP Publishing

sified how these represent the liquid water systems, correlating the obtained results with average thermodynamic observables of water. Microscopically, by comparing the SOAP spectra of geometrically-derived ice structures with those detected in the liquid systems during the MD simulations, we obtained evidence of the presence of local ice structures that may form and disappear during the MD even at temperature above the freezing one. In particular, we found presence of transient ordered ice-like structures in all the water models studied in this paper, which suggests that, at the very least, this is a common feature in the modeling itself of liquid water.

II. RESULTS AND DISCUSSION

We started our analysis performing a series of 10 ns-long classical molecular dynamics (MD) simulations in the NpT ensemble (with the three constants NpT representing the number of particles, the pressure, and the temperature during the MD run, respectively) of identical cubic boxes filled with 1,024 water molecules for all the 16 rigid water models listed in Table I. For the best comparability, the simulation conditions were kept the same for all studied models (see the Methods section for complete computational details).

A. Macroscopic comparison with the experimental data

From the equilibrium MD trajectories of all simulated systems, we computed a series of observables (displayed in Figure 2): (1) the liquid water density, (2) the static dielectric constant, and (3) the pair radial distribution function of the oxygen atoms. Both the water density (1) and the dielectric constant (2) were calculated reported as functions of the temperature (Figure 2: left and center), while the oxygen atoms' radial distribution functions were calculated at 298 K (Figure 2: right).

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT
The Journal of Chemical Physics
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599
AIP Publishing

TABLE I. Summary of the classical rigid water models compared herein.

Model name	Refs.	Year	Number of sites
SPC	ref.[31]	1981	3 Sites
SPC/E	ref.[40]	1987	
SPC/E_b	ref.[41]	2012	
TIP3P	ref.[32]	1983	
TIP3P-FB	ref.[42]	2014	
OPC3	ref.[43]	2016	
TIP4P	ref.[32]	1983	4 Sites
TIP4P-Ew	ref.[44]	2004	
TIP4P-FB	ref.[42]	2014	
TIP4P-ICE	ref.[45]	2005	
TIP4P/2005	ref.[46]	2005	
TIP4P/ε	ref.[47]	2014	
OPC	ref.[48]	2014	
TIP5P	ref.[49]	2000	5 Sites
TIP5P-E	ref.[50]	2004	
TIP5P/2018	ref.[51]	2018	

From such data, we can observe that an experimentally-consistent representation of the water density is out of reach for 3-sites models of water (Figure 2: top-left panel). This is true also for the most recent models like TIP3P-FB and OPC3 (as it was also anticipated in the original publications)[42,43]. On the other hand, we observe an overall qualitative agreement with experiments for the density and dielectric constant trends for some 4-sites models (*e.g.*, TIP4P/ε and OPC), while the other

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

The Journal of Chemical Physics

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

AIP Publishing

4-sites and, in general, all 5-sites models display poor agreement with the experimental observables. All of them (with the exception of TIP4P and TIP4P/ICE) show a maximum around ∼277 K (Figure 2: middle- and bottom-left panels), as expected from experiments.

Regarding the static dielectric constant (ε), we observe that a higher/lower number of sites is not crucial to have a satisfactory representation of it. Shown in Figure 2 (center), two models with 3-sites and three models with 4-sites representation display a good agreement with the experimentally observed ε. While this is somewhat expected for TIP4P/ε, which indeed has been parametrized with this main goal, this is a characteristic that is also well reproduced in recent parametrizations obtained by Wang *et al.* (TIP3P-FB and TIP4P-FB),[42] and Izadi *et al.* (OPC3 and OPC).[43,48]

Regarding oxygen-oxygen radial distribution function (Figure 2: right panels), we observe:

- a general discrepancy between all models and the experiments in terms of height of the 1st radial distribution function peak, which has been generally imputed to the fact that quantum effects are not taken into account in such classical representations[21];

- overall, a good representation of the subsequent solvation shells (up to at least the 3rd) by all water models, with the notable exception of SPC and TIP3P models.

From such first macroscopic analysis, the considered observables show that all the models and representations analyzed here are more or less comparable in terms of accuracy in the treatment of the bulk properties of liquid water. The largest deviation from the experimental data, in this sense, is observed for the oldest, yet widely employed, SPC and TIP3P models.

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT
The Journal of Chemical Physics
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599
AIP Publishing

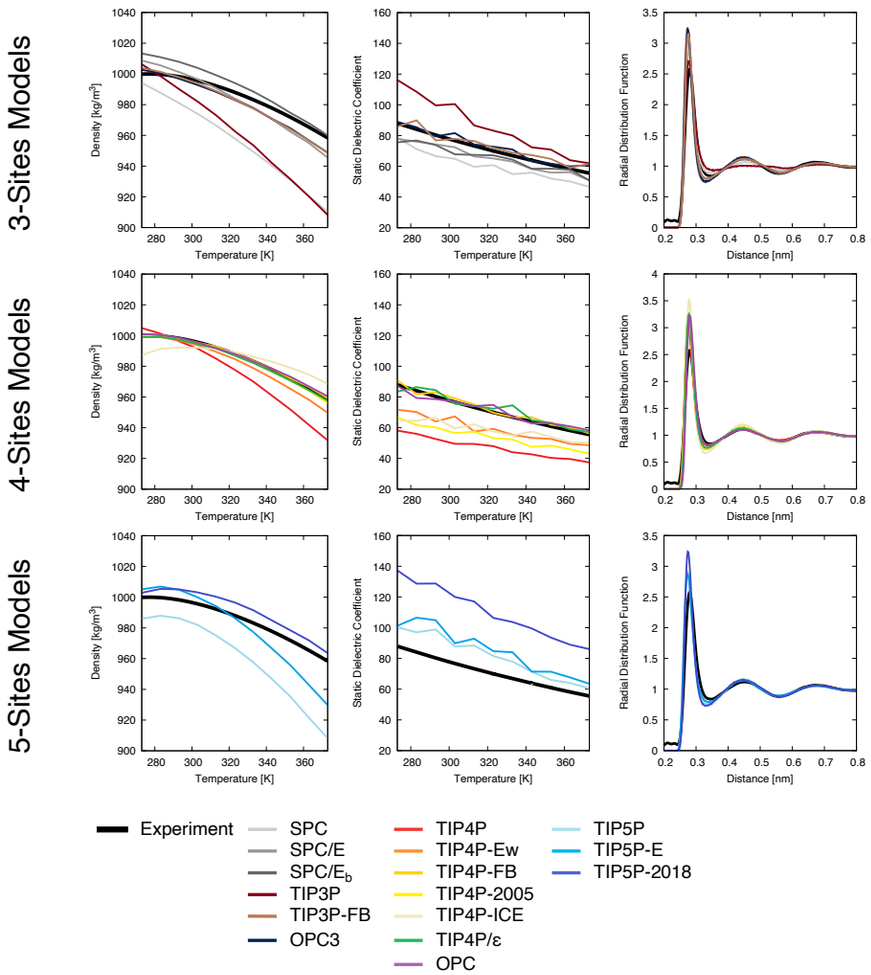


FIG. 2. Observables obtained from our simulations compared with experimental data. From left to right, liquid water density in function of temperature (experimental data from ref.[52]), static dielectric constant in function of temperature (experimental data from ref.[53]), and the radial distribution function at 298 K computed on oxygen atoms (experimental data from ref.[54]). To have a fair comparison between all the models, the first row contains the models that have a 3-sites representation, then in the middle row the 4-sites representation, and in the last row we have the 5-sites models.

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT
The Journal of Chemical Physics
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599
AIP Publishing

B. Microscopic SOAP analysis of liquid water models

We opted for a data-driven analysis approach in order to gain a deeper insight into the differences/similarities between the various considered models and their representation of the bulk properties of water. Inspired by a recent work on the comparability of lipid force fields,[37] our approach uses a local environment descriptor, *i.e.*, the Smooth Overlap of Atomic Position (SOAP),[55,56] and SOAP-based distances that allow us to compare and classify the various modeled systems. This SOAP-based analysis allows us to characterize the systems based on the molecular environments that are formed during the equilibrium MD simulation. In particular, centering one SOAP vector on the oxygen atom of every water molecule, this analysis substantially accounts for the geometrical arrangements of the molecules (*i.e.*, both hydrogen and oxygen atoms) that surround each water molecule during MD simulations. In this way, we can compare the simulated systems based on the molecular environments that emerge in the systems themselves during the equilibrium MD simulations.

For this SOAP analysis, we considered all the simulations performed at T=298 K: the underlying assumption is that models parametrized to represent liquid water should work at their best at room conditions. As a first step, we computed the SOAP power spectra centered on the oxygen atoms of each water molecule in each system (1,024 water molecules in the simulation boxes). This has been done for each frame during the equilibrium MD trajectories (for a total of 1,001 MD frames analysed for each simulated model). The SOAP power spectra were computed using a cutoff of 1 nm and $l_{\max} = n_{\max} = 8$ (see Methods section for complete details). The same parameters were used for the SOAP calculations of all systems studied in this work. Such a cutoff distance guarantees that the SOAP analysis takes into account at least up to the 3rd solvation shell around each water molecule. Also, it has been recently shown that in liquid and liquid-like environments a larger cutoff

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT
The Journal of Chemical Physics
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599
AIP Publishing

is preferable rather than shorter ones.[37,38] In this way, we first obtain a dataset containing a total of 1,025,024 spectra (1,024 for each water molecule in the system, for each of the 1,001 sampled MD frames), that we then average, obtaining a final characteristic average SOAP spectrum for each simulated system. This analysis captures with high sensitivity all local environments present in the simulated systems, and collect them into a global high-dimensional SOAP spectrum that contains a unique *multi-dimensional fingerprint*. Similar to what recently done to classify lipid assemblies,[37] from the average SOAP spectra associated to each modelled system, we then computed a pairwise distance matrix, based on the metrics d^{SOAP}. Such d^{SOAP} quantifies the difference between two spectra, measuring the distance between the simulated systems in the high-dimensional SOAP space (see Method section for additional details). This distance represents then their similarity/dissimilarity in terms of the local molecular environments formed in the liquid systems at equilibrium conditions. The matrix we obtained is shown in Figure 3, left panel. To provide a more informative representation of the data, we also projected the distance matrix on a 2D plane making use of the Multidimensional Scaling Algorithm (MDS)[57], reported in Figure 3, right panel.

From the scale in Figure 3, it appears evident that the SOAP distances involved are very small, with magnitudes between the $\sim 10^{-4} - 10^{-3}$ range. For sake of clarity, we underline here that d^{SOAP} is by definition in the $[0, \sqrt{2}]$ interval, and the range of distances obtained between different force fields in a previous work on the comparison between lipid bilayers[37] gave d^{SOAP} values in the order of $\sim 10^{-1}$. Furthermore, we analyzed the distribution of d^{SOAP} among different environments in liquid water (Figure SI3), confirming that they spread in the SOAP space (as shown in Montserrat et al.[36]). This observation is reassuring, because it means that, even on a more microscopic level, the many classical models analyzed here yield globally similar bulk water average structures. Yet, it is worth noting that the d^{SOAP} values,

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT
The Journal of Chemical Physics
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599
AIP Publishing

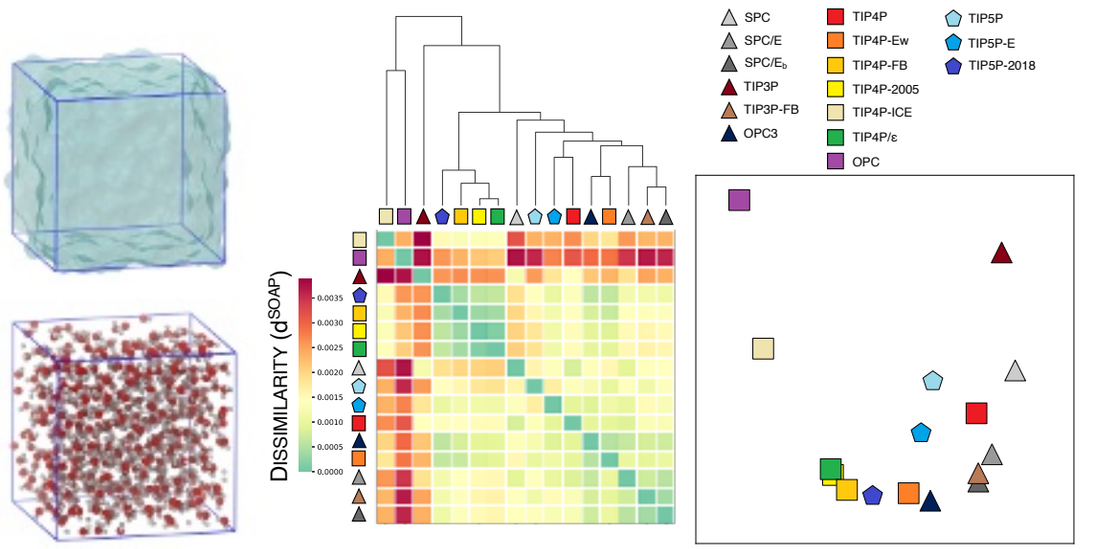


FIG. 3. SOAP distance analysis of different rigid water models. On the left we represented the cubic box at atomic detail (bottom right) and in "liquid" representation (top right). In the central panel we have a reciprocal SOAP distance matrix for all the water models analyzed; in the right panel, we have the MDS projection on a 2D plane of the distance matrix. The point shape is related to the number of sites of the model: triangle for 3-sites models, square for 4-sites, and pentagon for 5-sites representations.

albeit small, are non-negligible, and indeed some models are more/less distant that other ones.

Looking in more detail at the data of Figure 3, Ewald-corrected 3-sites models, *i.e.* SPC models (SPC/E and SPC/E$_b$) and TIP3P-FB, are rather close to each other, as expected from similar parametrization protocols and representation used. Noteworthy, we obtain a similar situation for TIP4P-FB, TIP4P-2005, TIP5P-2018, TIP4P-Ew, and OPC3. This confirms that the number of sites is not the unique critical parameter for obtaining a different representation of water.

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT
The Journal of Chemical Physics
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599
AIP Publishing

Some interesting observations can be made regarding the 3 more distant outliers of this analysis:

- *TIP3P*. Despite being one of the most widely used rigid water models, this appears different from the large cluster of models observed. This result can be interpreted also in light of the limited ability of TIP3P to capture the water solvation shell (Figure 2, right: brown radial distribution function) and its deviation from the experiments in terms of density values (Figure 2, left).

- *TIP4P-ICE*. Here the main discrepancy that can partially explain the d^{SOAP} from most of the other water models is in the density, which appears to be deviated from the experimental data (Figure 2, left: in light yellow).

- *OPC*. This case is interesting because this model is definitely in agreement with all the three observables evaluated in Figure 2. The large d^{SOAP} distance from the other models presented here could be due to differences in the O-H equilibrium distance ($\sim$ 0.87 Å for OPC) with respect to the others (*e.g.*, $\sim$ 0.96 Å for TIP3P, TIP4P-2005, and TIP4P-Ew, and 1.0 Å for SPC/E). This point is corroborated by the comparison of the $g_{OH}(r)$ radial distribution functions, indicating some differences also in the distances of the 1st solvation shell (see Figure S1 in the Supporting Information).

Connecting the d^{SOAP} data of Figure 3 to the low dimensional data of Figure 2 is nonetheless non-trivial: in fact, the amount of information contained in SOAP data is rich, and strongly related to the local environments that are formed along the MD simulations.

To further confirm the difficulty in this comparison, and the richness of the information contained in SOAP representation, we also computed two scatterplots that

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT
The Journal of Chemical Physics
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599
AIP Publishing

compare the d^{SOAP} between a reference model (here TIP3P) and the relative density and dielectric constant difference (see figure 4).

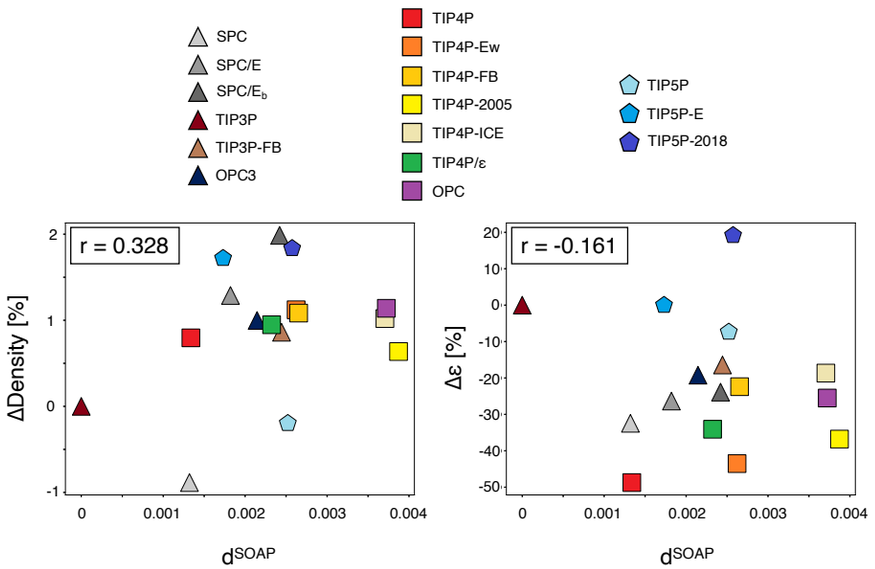


FIG. 4. Comparison of SOAP distance with density (left) and static dielectric constant (right). We calculated these distances with respect to TIP3P water model; we can observe that the information contained in SOAP power spectra does not correlate neither with density (Pearson correlation coefficient of 0.328), nor with static dielectric constant (Pearson correlation coefficient of -0.161).

Figure 4 shows that while the density component is an important contributor to the SOAP analysis (and thus SOAP distance), nevertheless it is not enough to completely define and interpret SOAP results. Similarly, also the static dielectric constant does not represent a main contributor to interpret the SOAP distance. As a consequence of this, we underline that SOAP is a tool that provides information that could not be obtained via a combination of other macroscopic observables.

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

The Journal of Chemical Physics

AIP Publishing

C. SOAP analysis of ordered ice phases

In a recent work, Monserrat *et al.* computed the SOAP spectra of a collection of 1,000 snapshots taken from *ab initio* liquid water simulations.[36] Similarly, Engel *et al.*[58] studied diverse ice structures employing a combination of SOAP spectra and a dimensionality reduction technique (sketchmap), applied to over 70,000 ice geometries. Calculating *a posteriori* the SOAP spectra of a set of relevant ice structures, the authors could observe presence of ordered ice-like constructs (*i.e.,* domains in which the arrangement of water molecules is similar to those present in geometrically-optimized ice structures), which can be formed also in liquid water. Such a comparison was performed by using PCA as a dimensionality reduction technique, and by comparing the PCAs of the SOAP spectra obtained for the liquid *vs.* ice aqueous systems.

Building on such idea, here *we revert the approach*, designing a SOAP-based classification strategy that allows us to automatically recognize ice-like motifs in liquid water MD simulations. As a first step, (1) we obtain *ab initio*-optimized reference structures of different ice types; for all these, (2) we compute their characteristic SOAP spectra, and eventually (3) we use these spectra as references for our classification analysis.

Among the known ice types, in this study we consider a subset of them (investigated also in[36]) as the references ice II, ice VIII, ice IX, ice X, ice XIII, ice XIV, ice XV, ice XVII, clathrates sII and sH, ice Ic, ice Ih, and we discard other ones. In particular:

- *Ice Ih*: this is the hydrogen disordered form of hexagonal ice. It accounts for basically all ice naturally found on Earth and thus we include it in our study.;
- *Ice III*: this is the hydrogen-bond disordered form of ice IX, and is not consid-

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT
The Journal of Chemical Physics
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599
AIP Publishing

ered, while we retain the latter;

- *Ice IV*: not considered here, nor in Ref.[36] (this ice type displayed issues in the *ab initio* optimization)[36];
- *Ice V*: this is the hydrogen-bond disordered form of ice XIII, and it is not considered here, while we retain only the latter;
- *Ice VI*: this is the hydrogen-bond disordered form of ice XV (reported in ref.[59]) – this is not considered, while we retain only the latter;
- *Ice VII*: this is the hydrogen-bond disordered form of ice VIII, and we retained only the latter;
- *Ice X*: displayed issues in the *ab initio* optimization performed in ref.[36], and was not included in that study. However, this ice type is a well-studied ice type, and we include it in our analysis;
- *Ice XI*: this is the low-temperature equilibrium structure of hexagonal ice Ih. Since ice Ih is the most common form of ice, here we retain this one, and we discard ice XI;
- *Ice XII*: this is the hydrogen-bond disordered form of ice XIV, and we retained only the latter;
- *Ice XVI*: this is a low-density clathrate cage topologically equivalent to ice sII. For this reason, we decided to retain only sII and to discard ice XVI in this analysis.

In a similar way with respect to the previous section, we obtained the mutual SOAP distance matrix for all these reference ice structures (see Figure 5).

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT
The Journal of Chemical Physics
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599
AIP Publishing

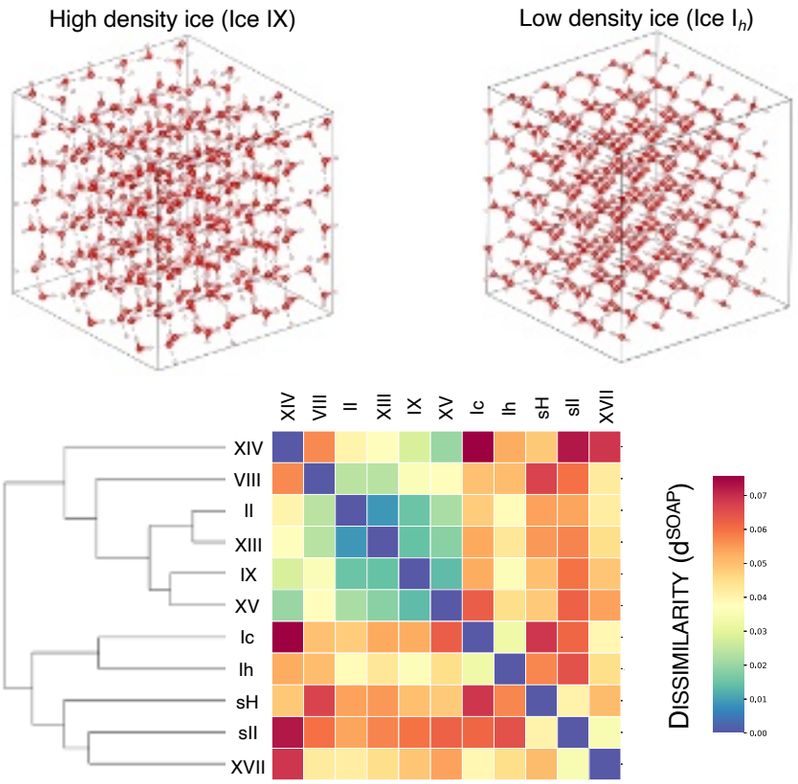


FIG. 5. SOAP distance matrix between reference ice configurations. Ice X has not been inserted in this matrix for sake of clarity because is far away from all the other ice phases (the complete matrix is available in the Supporting Information, Figure S2).

The most straightforward interpretation of the differences captured in this analysis by the d^{SOAP} is related to the different ice densities. The densest analyzed ice phase, ice X, is found very far from all others (see Figure S2 in the Supporting Information). Regarding the other ice types, from this analysis we can identify two main sub-groups (referring to the dendrogram in Figure 3): (i) low-densities ice phases (clathrates — sII and sH—, Ic, Ih, and XVII) and (ii) high-density phases (II, VIII, IX, XIII, XIV, and XV).

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

The Journal of Chemical Physics

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

AIP Publishing

D. Identification of ice-like environments in liquid water

In the work of Monserrat *et al.*[36], the authors detected the building blocks of different ice phases in MD simulations of liquid water using a machine-learning (ML) potential. This ML potential was trained on quantum calculations at the rev-PBE0[60] level of theory with semiclassical D3 dispersion correction[61] performed on some hundreds of different liquid water conformations.

In particular, the ML potential used by Monserrat et al. allowed to observe in a liquid water system, within a range of 8 Å around the oxygen atoms of the water molecules, local ice-like environments compatible with those of ice structures.

Here we aim to determine to what extent widely used classical rigid water models can capture the same phenomena, using our approach based on the SOAP distance d^{SOAP} as defined in eq. (6) in the Methods section. In particular, we want to label every spectrum (1,024 water molecules for 1,001 MD snapshots, for a total of 1,025,024 SOAP spectra for every simulated model) as closer to a certain reference of either ice or liquid water.

To achieve this goal, we want to generate reference structure for every state of interest. Operatively (i) we consider the SOAP spectra computed from the geometrically optimized structures of ice considered in the previous section (static ice SOAP references). (ii) Then, to obtain the same reference for liquid water, from every simulation carried out with a specific water model, we extract its average SOAP power spectrum. This is considered as the fingerprint of how each force field describes liquid water (liquid water SOAP reference). Finally, for every spectrum obtained from the simulations, (iii) we compute its distance from all the reference SOAP spectra (*i.e.*, average water model *vs.* all the ices), assigning it to the closest one: in other words, all the local environments around oxygen atoms in liquid water simulations are attributed to either (average) liquid water or some ice phases, based on the smallest

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT
The Journal of Chemical Physics
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599
AIP Publishing

d^{SOAP} value. The results obtained with such analysis are available in Figure 6.

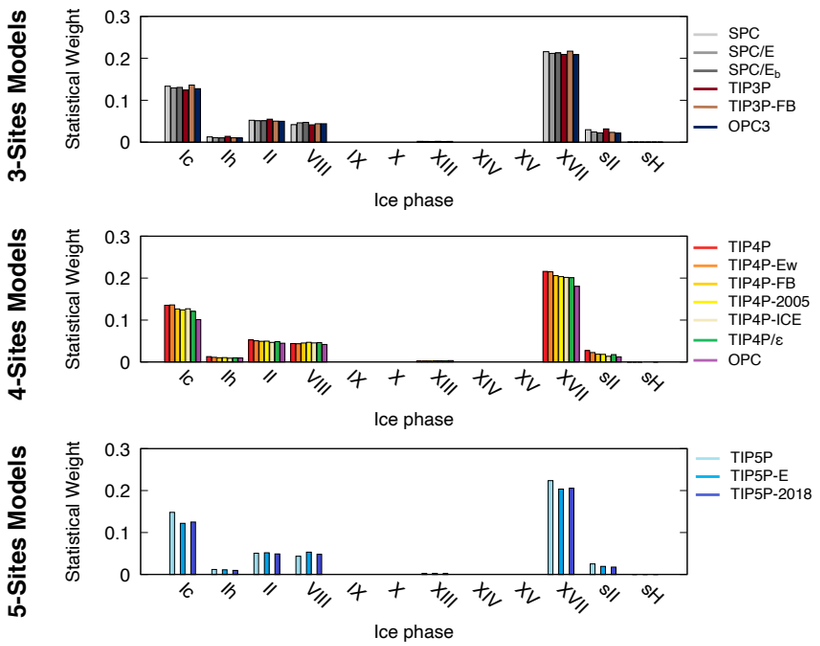


FIG. 6. Statistics related to ice-like molecular environments in classical MD simulation of liquid water at 298 K. The probability to be in liquid water state was removed for sake of clarity.

We obtain a statistical weight of ∼20% for water molecules to visit local molecular motifs resembling those proper of ice XVII. This is observed consistently in most of the considered classical water models. Also Ic, VIII, and II ice-like environments are visited, with a statistically significance of ∼10-12%, ∼5%, and ∼4%, respectively (Figure 5). In this analysis, we consider water molecules becoming part of an ordered domain when these are in contact (*i.e.*, when they form hydrogen bonding) with the neighbor ones, forming a network. It is worth noting that we never observe large ice domains (crystal nuclei) emerging during the MD runs, as it is expected at room temperature. In fact, our analysis pinpoints the general evidence that the

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

The Journal of Chemical Physics

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

AIP Publishing

probabilities to find ordered ice-like structures are non-null, and that these can be compared with each other to understand the ordered motifs that emerge along the MD with some statistical relevance.

While sparsely detected along the MD, the formation of other ice-like phases (*e.g.*, ice Ih and sII) is found not significant with respect to the molecular environments reminiscent of ice XVII, Ic, VIII, and II, and of liquid water. As expected for a classical water model,[62] standard ice (*i.e.,* ice Ih), conventionally formed at 1 bar pressure and $T < 0$C, is found less favored with respect to ice Ic in liquid water conditions.

In such a population analysis, we can observe a substantial agreement between all the classical water models analyzed herein (irrespectively to the number of sites, or their parametrization protocols). The only (non-substantial) difference detected is for OPC, where the population of Ic and XVII ice-likes visited environments is found smaller than for the other models, while the liquid water population is found larger. This can be imputed to the short O-H intra-molecular distance already discussed above, which favors on average liquid water and reduces the probability to observe the statistical formation of local ordered water motifs having a SOAP spectra similar to that of ices. Interestingly, aside from low-density ice-like molecular environments (sII, sH, Ic, Ih, and XVII ice-like domains), we observe that ordered water environments closer to ice VIII, a high-density form of solid water, may also statistically emerge during the MD simulations. Summarizing, for all compared classical water models, we observe along the MD runs a non-negligible probability for the emergence of local arrangements of water molecules resembling ices in liquid water at 298 K.

The presence of local ordered domains that may form in liquid water even far from the freezing temperature is also supported by the experimental evidence. In particular, X-ray spectroscopy revealed a bimodal structural distribution in liquid water[8]. Furthermore, two distinct peaks in the Raman spectra of liquid water have

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

The Journal of Chemical Physics

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

AIP Publishing

been also experimentally assigned to molecules involved in ordered (e.g., tetrahedral) H-bond environments vs. a disordered/liquid H-bond network[63]. IR spectra also confirmed similar heterogeneity in liquid water.[64] These experimental observations find consistency with the SOAP-based results we report herein, providing robustness to our analysis.

To verify how the most representative ice-labeled water molecules are distributed during the MD simulation in the water box, we dissected in more detail a pair of conformations (MD snapshots) obtained from the MD run of TIP4P-2005 water model (Figure 7); for sake of completeness, we specify that in this analysis the only three reference structures considered are the average liquid water simulated with the TIP4P-2005 force field, Ice XVII, and Ice Ic.

Looking more in detail at the individual SOAP obtained for each water molecule in the system at all sampled MD snapshots, we clearly observe that molecules which are labeled within an ice-like environment at a given MD snapshot easily undergo transition in water, or are even observed to become part of a different ice-like local ordered domain at a successive MD snapshot. This demonstrates the ephemeral and statistical nature of such ordered local domains, that are formed and destroyed along the MD, as it is reasonable for a water system of liquid water at room temperature.

We underline the fact that the identification of an "isolated" single water molecule classified as ice-like does not mean that it is the only ordered molecule (no order can exist with just one molecule). In fact, in our SOAP analysis, the cutoff threshold r_{cut} is equal to 1 nm, which accounts for more than the first 3 solvation shells around each water molecule. This means that even in the case of clusters composed of one single molecule, we have a water molecule in the simulation box around which, within a spherical volume of radius 1 nm, the neighbor water molecules are found distributed in a way that is closer to a specific (ordered) ice type rather than to the average (disordered) distribution typical of liquid water. The analysis of the lifetime of such

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT
The Journal of Chemical Physics
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599
AIP Publishing

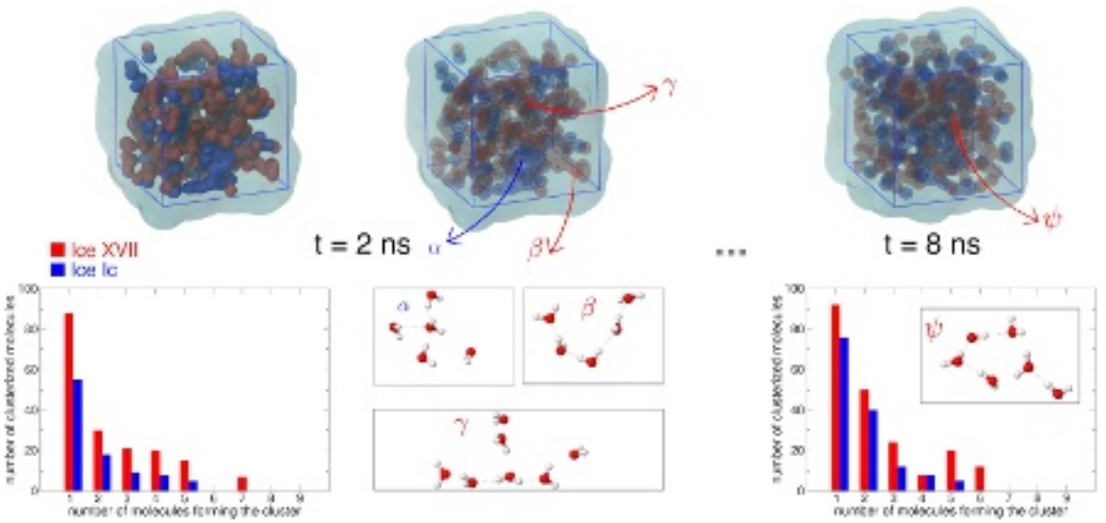


FIG. 7. Ephemeral ice-like water arrangements visited in liquid water MD simulation. Taken from the same MD trajectory of a liquid TIP4P-2005 water box at room temperature, we show 2 MD frames – at 2 ns (left) and 8 ns (right) – separated by a set of dots (...). As it can be observed from the rendered structures (top) and the population histograms (bottom), we do not observe very large ice like-structures along the simulation (we consider 2 molecules in contact if they share an H-bond). This is a further hint that the structures observed here at 298 K are in fact metastable and transient, while at lower temperatures they might possibly represent pre-critical ice nuclei.

ice-like water environments indicated that this is on average very short, namely, in the order of ~ 10 ps or below. This is reasonable, as we are at room condition, where the formation of ordered ice-like clusters is highly unfavored, and it underlines the ephemeral nature of these pseudo-ordered environments in such conditions.

Summarizing, the strength of this analysis relies in the fact that is fully data-driven, and it is based on a general and abstract high-dimensional descriptor of the molecular environments that surround each water molecule in the system. The

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

The Journal of Chemical Physics

AIP Publishing

SOAP environments that are sampled during the MD simulation are then compared with ice-environments/configurations contained in a reference dataset. It is worth noting that this analysis is quite flexible and abstract, and rather unbiased - namely, it is intrinsically limited only by the set of ice structures used as reference points in the dataset, and by the time and space resolution used for the calculation of the SOAP spectra (i.e., the cutoff and the timestep between the sampled snapshots). Furthermore, it can be in principle applied to assess similarities/differences in the local molecular configurations of any system (even different from water) for which well-characterized reference configurations are available.

III. CONCLUSIONS

In this work, we compare how classical water models represent liquid water macroscopically and microscopically, using a metrics that is built on a high-dimensional SOAP framework. The general behavior of rigid water models analyzed here is found overall similar, but some differences in the studied observables can still be detected (as also shown in an extensive way by the recent work of Kadaoluwa Pathirannahalage *et al.*[65]). With our approach we easily detect models with considerable differences in the monitored observables (mainly TIP3P and TIP4P-ICE), but also another state-of-the-art model (OPC) that displays a very peculiar shorter H-O intramolecular bond, and thus is correctly identified as an outlier as well.

We compare by means of our SOAP distance a collection of ice structures obtained from *ab initio* geometrical refinement of experimental ones. We confirm that our SOAP analysis mainly separates ice structures on the basis of their packing (and thus density), like already shown by Monserrat *et al.*[36]. Comparing the SOAP distances obtained in this analysis, we can qualitatively observe that the differences between ices are significantly larger (1 order of magnitude) than those which can be observed

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

The Journal of Chemical Physics

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

AIP Publishing

between the various studied liquid water models. This is reasonable because in all the simulations of the liquid phase transient domains resembling ices appear. Thus, said models share a large number of possible structural realizations and their close d^{SOAP} between averaged spectra reflects this. On the contrary, SOAP distances on the static structures of ices display the inherent differences among their geometries.

Combining our SOAP distance, with a classification of disordered model systems (liquid water) based on ordered reference systems (ice structures), we could clearly identify the formation of ice-like local molecular environment. In other words, water molecules arrange with respect to each other in an ordered configuration reminiscent of that proper of molecules in different ice types, even in liquid water at room conditions. Computing the SOAP distances between every water molecule in the system at each MD simulation step *vs.* reference ice structures and average water displacement, we could quantify the relative probabilities to observe the emergence of different types of local ordered domains in the liquid systems along the equilibrium MD. We find a non-negligible probability for having water molecules becoming part of ordered local domains, more similar to ice-like elementary domains than to liquid water, using all the classical water force fields analyzed. This observation is in agreement with recent investigations performed on *ab initio* or *ab initio*-trained Machine Learning potentials[36,66], as well as in high-dimensional analysis of classical water[67]. On this last point, this observation is valid in *both* quantum and classical simulations of water, and thus it is reasonable to wonder if this is an *artefact* due to the very way that water is currently modeled by computational means, or if it is an actual phenomenon which is correctly modeled and is still unnoticed experimentally. We also observe the characteristic ephemeral and statistical nature of such ordered water environments, from which water molecules can exchange in an out, and which may dynamically emerge and disappear along the MD simulations, as it can be expected for a liquid water system. The high-dimensional analysis reported herein can

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

The Journal of Chemical Physics

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

be extremely useful to characterize complex molecular systems where ordered and disordered states coexist in dynamical exchange and equilibrium.

Such a classification approach holds a considerable potential for the development of more accurate and reliable force fields at an affordable computational cost, as well as to study and classify a variety of self-organizing systems, not necessarily restricted to water models.

As a last point, the fact that (ephemeral) local ordered domains arise even in liquid water, and quite far from the freezing temperature, finds consistency with the available experimental evidence (e.g., X-ray, Raman, IR spectroscopy)[8,63,64,68,69]. On the one hand, this supports the results reported herein. On the other hand, this underlines the importance of identifying such ephemeral ordered domains, and studying them, as these may play a central role in global thermodynamic properties of water, as well as phase transitions and ice nucleation processes upon temperature decrease.

Summarizing, the outcomes of this work are interesting from multiple point of view. First, our results demonstrate that liquid-like ordered water environments may emerge also in relatively simple rigid water models. Such models are widely used to simulate a variety of systems in water. This makes these observations relevant in multiple directions, considered, for example, that it has been demonstrated recently how liquid water can play a key role in controlling host-guest interactions[70], and how such ephemeral ordered domains may play an important role in, e.g., the processes of ice nucleation in liquid water during freezing.

AIP Publishing

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

The Journal of Chemical Physics

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

AIP Publishing

IV. METHODS

A. MD Simulations

We created a cubic box containing 1,024 water molecules, arranged in initial random configurations, using Packmol[71]. All simulations were conducted using GROMACS version 2021.2[72]. All model systems were first minimized using a steepest descent algorithm for $2 \cdot 10^3$ steps, and then equilibrated *via* 1 ns of MD simulation conducted in NVT conditions (with the constants NVT being the number of particles, the simulation volume, and the simulation temperature, respectively). As a last preparation step, we equilibrated the volume of the simulation box *via* 1 ns of MD simulation in NpT conditions (with the constants NpT being the number of particles, the pressure, and the simulation temperature, respectively). All MD runs used a 2 fs timestep. After these equilibration phases, we ran 10 ns long production MD runs with the same conditions as the last equilibration phase (NPT). For all those steps, we kept temperature with velocity rescale thermostat[73] (with coupling constant of 0.2 ps) and the pressure was maintained constant (1 bar) with an isotropic scheme using the cell rescale barostat[74] (with a coupling constant of 2 ps, and a compressibility of $4.5 \cdot 10^{-5}$ bar^{-1}). A cutoff distance of 1 nm was used for short-range electrostatic and van der Waals interactions, and the long-range interactions were computed with the particle-mesh Ewald summation method[75]. Long-range dispersion corrections to the pressure and potential energy were considered[76]. We pinpoint that this choice of PME for long-range interactions can cause some deviation from the original values in the oldest water models in their original simulation conditions. This choice has been made nevertheless to have the best common ground for a fair comparison between different classical water models used herein (and to employ simulation protocols as similar as possible to the ones routinely used nowadays).

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT
The Journal of Chemical Physics
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599
AIP Publishing

For all the water models, we performed a simulation from 273 K to 373 K (with a temperature-step of 10 K), and a last block of simulations performed at 298 K (25 C) from which we computed radial distribution functions for O-O (see Figure 3) and for O-H (see Figure S1 in Supporting Information).

B. Observables

During our analysis we computed a series of observables, namely liquid water density (ρ_L) as a function of temperature, static dielectric constant (ε) as a function of temperature, and the radial distribution function between oxygen atoms (g_{OO}) at 298 K.

Liquid water density. We computed the mass density of the water models by calculating the size of the simulation box, namely as:

$$\rho_L = \frac{N \cdot m_{H_2O}}{\mathcal{N}_A \cdot V_{\text{box}}} \ , \tag{1}$$

where N is the number of water molecules (1024 in our case), m_{H_2O} is the mass of the water molecules in a.u., $\mathcal{N}_A$ is the Avogadro's number and V_{box} is the simulation box volume. This analysis was performed with PLUMED plugin version 2.7.[77,78]

Static dielectric constant. We computed the static dielectric constant from the fluctuations of the total dipole moment M of the simulation box, namely as:

$$\varepsilon = 1 + \frac{\langle M^2 \rangle - \langle M \rangle^2}{3\varepsilon_0 V k_B T} \ , \tag{2}$$

where ε_0 is the vacuum permittivity, V the volume of the simulation box, k_B the Boltzmann constant, T the system temperature, and $\langle \cdot \rangle$ represents the thermodynamic average. We computed this observable by using the `gmx dipoles` routine in the GROMACS[72] suite.

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

The Journal of Chemical Physics

AIP Publishing

C. SOAP

Smooth Overlap of Atomic Position (SOAP)[55,56] is an atomic environment descriptor that aims to translate the 3D configuration of a set of atoms around a center (that can be an atom, or more in general a point in space) to a high-dimensional representation. The result of such analysis is a vector called "power spectrum".

Given a system conformation Γ in the 3D space, the SOAP power spectrum calculation is done by expanding the local atomic/particle density $\rho_i(\Gamma, \vec{r})$ (defined in the neighborhood of every SOAP center within a spatial cutoff, r_{cut}) projecting it onto a basis of orthogonal radial functions $g_n(r)$ and spherical harmonics $Y_{lm}(\theta, \phi)$, which for the i-th site can be expressed as:

$$\rho_i(\Gamma, \vec{r}) = \sum_{j \in r_{cut}} \sum_{nlm} c^j_{nlm}(\Gamma) g_n(r) Y_{lm}(\theta, \phi) \ , \tag{3}$$

where the j index runs over all the sites in the cutoff. In our case, one SOAP is centered on the oxygen atom of each water molecule in the simulated systems. At each sampled MD snapshot, we thus obtain 1,024 SOAP spectra, characteristic for the environment that surrounds each individual water molecule in the system. The analysis is then repeated for 1,001 frames obtained from the equilibrium MD simulations, obtaining a total of 1,025,024 individual SOAP spectra for all simulated water models.

It is worth noting that $\rho_i(\Gamma, \vec{r})$ is multi component (it has one component for each chemical species). From the eq. (3) we can obtain the SOAP power spectrum vector defined as

$$\vec{p}(\Gamma)_{nn'l} = \pi \sqrt{\frac{8}{2l+1}} \sum_{m=-l}^{l} c^*_{nlm}(\Gamma) c_{n'lm}(\Gamma) \ , \tag{4}$$

which encodes all the information of the atomic environment. Eq. (4) represents also the computational output obtained from the SOAP calculation using the DScribe[79] package.

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

The Journal of Chemical Physics

AIP Publishing

A similarity measure between two environments centered in two sites can be formally defined by building a linear kernel of their density representations. Such kernel can be reduced to the dot product of the two sites' SOAP power spectra as:[55]

$$K^{SOAP}(i,j) = \rho_i(\Gamma,\vec{r}) \cdot \rho_j(\Gamma,\vec{r}) \propto \vec{p}_i \cdot \vec{p}_j \ . \tag{5}$$

Eq. (5) can be interpreted as a measure of how much the two local environments (surrounding a water molecule) are superimposed to each other (*i.e.*, how similar they are). The value of K^{SOAP} ranges from 0 for completely different to 1 for exactly matching environments.

From eq.(5), we can further define a metrics (that we call "SOAP distance") between two environments, namely as:

$$d^{\mathrm{SOAP}}(i,j) = \sqrt{2 - 2 \cdot K^{SOAP}(i,j)} \propto \sqrt{2 - 2\vec{p}_i \cdot \vec{p}_j} \ , \tag{6}$$

where $\vec{p}_i$ is the i-th center's power spectrum. Both SOAP kernel and distance representations give a bounded measure of how similar two local sites/environments are (*i.e.*, how their local densities are superimposable in the power spectra space).

ACKNOWLEDGMENTS

We thank Andrea Gardin and Carlo Camilloni for useful discussion, and Daniele Rapetti for reviewing the code used in this work. GMP acknowledges the funding received by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 818776 - DYNAPOL).

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

The Journal of Chemical Physics

AIP Publishing

DATA AVAILABILITY

All the input files and relevant analysis script are available on GitHub at the address `https://github.com/GMPavanLab/Water_SOAP`.

SUPPORTING INFORMATION

Additional figures on our analyses (O-H radial distribution function, full SOAP distance matrix for ice and distribution of d^{SOAP}) are available in the Supporting Information PDF file.

REFERENCES

[1]F. Frappart, "Water and life," Nature Geoscience **6** (2013).

[2]G. Vladilo and A. Hassanali, "Hydrogen bonds and life in the universe," Life **8** (2018), 10.3390/life8010001.

[3]P. Ball, "Water — an enduring mystery," Nature **452**, 291–292 (2008).

[4]P. Gallo, K. Amann-Winkel, C. A. Angell, M. A. Anisimov, F. Caupin, C. Chakravarty, E. Lascaris, T. Loerting, A. Z. Panagiotopoulos, J. Russo, *et al.*, "Water: A tale of two liquids," Chemical reviews **116**, 7463–7500 (2016).

[5]M. Chaplin, "Water structure and science, URL: https://water.lsbu.ac.uk/water/," (retrieved April 2022).

[6]P. Wernet, D. Nordlund, U. Bergmann, M. Cavalleri, M. Odelius, H. Ogasawara, L. Å. Näslund, T. K. Hirsch, L. Ojamäe, P. Glatzel, L. G. M. Pettersson, and A. Nilsson, "The structure of the first coordination shell in liquid water," Science **304**, 995–999 (2004), https://science.sciencemag.org/content/304/5673/995.full.pdf.

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

The Journal of Chemical Physics

AIP Publishing

[7]C. Huang, K. T. Wikfeldt, T. Tokushima, D. Nordlund, Y. Harada, U. Bergmann, M. Niebuhr, T. Weiss, Y. Horikawa, M. Leetmaa, *et al.*, "The inhomogeneous structure of water at ambient conditions," Proceedings of the National Academy of Sciences **106**, 15214–15218 (2009).

[8]A. Nilsson and L. G. M. Pettersson, "The structural origin of anomalous properties of liquid water," Nature Communications **6**, 8998 (2015).

[9]P. H. Poole, F. Sciortino, U. Essmann, and H. E. Stanley, "Phase behaviour of metastable water," Nature **360**, 324–328 (1992).

[10]S. Fanetti, A. Lapini, M. Pagliai, M. Citroni, M. Di Donato, S. Scandolo, R. Righini, and R. Bini, "Structure and dynamics of low-density and high-density liquid water at high pressure," The Journal of Physical Chemistry Letters **5**, 235–240 (2014), pMID: 26276206, https://doi.org/10.1021/jz402302z.

[11]N. Ansari, R. Dandekar, S. Caravati, G. Sosso, and A. Hassanali, "High and low density patches in simulated liquid water," The Journal of Chemical Physics **149**, 204507 (2018), https://doi.org/10.1063/1.5053559.

[12]J. Niskanen, M. Fondell, C. J. Sahle, S. Eckert, R. M. Jay, K. Gilmore, A. Pietzsch, M. Dantz, X. Lu, D. E. McNally, *et al.*, "Compatibility of quantitative x-ray spectroscopy with continuous distribution models of water at ambient conditions," Proceedings of the National Academy of Sciences **116**, 4058–4063 (2019).

[13]C. G. Salzmann, "Advances in the experimental exploration of water's phase diagram," The Journal of Chemical Physics **150**, 060901 (2019), https://doi.org/10.1063/1.5085163.

[14]S. Pipolo, M. Salanne, G. Ferlat, S. Klotz, A. M. Saitta, and F. Pietrucci, "Navigating at will on the water phase diagram," Phys. Rev. Lett. **119**, 245701 (2017).

[15]A. Reinhardt and B. Cheng, "Quantum-mechanical exploration of the phase diagram of water," Nature Communications **12** (2021), https://doi.org/10.1038/s41467-020-20821-w.

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

The Journal of Chemical Physics

AIP Publishing

[16]A. Hassanali, M. K. Prakash, H. Eshet, and M. Parrinello, "On the recombination of hydronium and hydroxide ions in water," Proceedings of the National Academy of Sciences **108**, 20410–20415 (2011), https://www.pnas.org/content/108/51/20410.full.pdf.

[17]A. Hassanali, F. Giberti, J. Cuny, T. D. Kühne, and M. Parrinello, "Proton transfer through the water gossamer," Proceedings of the National Academy of Sciences **110**, 13723–13728 (2013), https://www.pnas.org/content/110/34/13723.full.pdf.

[18]A. A. Hassanali, J. Cuny, V. Verdolino, and M. Parrinello, "Aqueous solutions: state of the art in ab initio molecular dynamics," Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **372**, 20120482 (2014), https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2012.0482.

[19]F. Muniz-Miranda, M. Pagliai, G. Cardini, and R. Righini, "Bifurcated hydrogen bond in lithium nitrate trihydrate probed by ab initio molecular dynamics," The Journal of Physical Chemistry A **116**, 2147–2153 (2012), pMID: 22309150, https://doi.org/10.1021/jp2120115.

[20]P. Gasparotto, A. A. Hassanali, and M. Ceriotti, "Probing defects and correlations in the hydrogen-bond network of ab initio water," Journal of Chemical Theory and Computation **12**, 1953–1964 (2016), pMID: 26881726, https://doi.org/10.1021/acs.jctc.5b01138.

[21]M. Ceriotti, J. Cuny, M. Parrinello, and D. E. Manolopoulos, "Nuclear quantum effects and hydrogen bond fluctuations in water," Proceedings of the National Academy of Sciences **110**, 15591–15596 (2013), https://www.pnas.org/content/110/39/15591.full.pdf.

[22]M. Machida, K. Kato, and M. Shiga, "Nuclear quantum effects of light and heavy water studied by all-electron first principles path integral simulations," The Journal of Chemical Physics **148**, 102324 (2018), https://doi.org/10.1063/1.5000091.

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

The Journal of Chemical Physics

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

[23]F. Giberti, A. A. Hassanali, M. Ceriotti, and M. Parrinello, "The role of quantum effects on structural and electronic fluctuations in neat and charged water," The Journal of Physical Chemistry B **118**, 13226–13235 (2014), pMID: 25286281, https://doi.org/10.1021/jp507752e.

[24]J. Kessler, H. Elgabarty, T. Spura, K. Karhan, P. Partovi-Azar, A. A. Hassanali, and T. D. Kühne, "Structure and dynamics of the instantaneous water/vapor interface revisited by path-integral and ab initio molecular dynamics simulations," The Journal of Physical Chemistry B **119**, 10079–10086 (2015), pMID: 26174102, https://doi.org/10.1021/acs.jpcb.5b04185.

[25]G. C. Sosso, S. Caravati, G. Rotskoff, S. Vaikuntanathan, and A. Hassanali, "On the role of nonspherical cavities in short length-scale density fluctuations in water," The Journal of Physical Chemistry A **121**, 370–380 (2017), pMID: 27935707, https://doi.org/10.1021/acs.jpca.6b11168.

[26]R. Dandekar and A. A. Hassanali, "Hierarchical lattice models of hydrogen-bond networks in water," Phys. Rev. E **97**, 062113 (2018).

[27]M. Fitzner, G. C. Sosso, S. J. Cox, and A. Michaelides, "Ice is born in low-mobility regions of supercooled liquid water," Proceedings of the National Academy of Sciences **116**, 2009–2014 (2019).

[28]N. Ansari, B. Onat, G. C. Sosso, and A. Hassanali, "Insights into the emerging networks of voids in simulated supercooled water," The Journal of Physical Chemistry B **124**, 2180–2190 (2020), pMID: 32032486, https://doi.org/10.1021/acs.jpcb.9b10144.

[29]R. Hu, C. Zhang, X. Zhang, and L. Yang, "Research status of supercooled water ice making: A review," Journal of Molecular Liquids **347**, 118334 (2022).

[30]P. H. Handle, T. Loerting, and F. Sciortino, "Supercooled and glassy water: Metastable liquid(s), amorphous solid(s), and a no-man's land," Proceedings of the National Academy of Sciences **114**, 13336–13344 (2017),

AIP Publishing

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

The Journal of Chemical Physics

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

https://www.pnas.org/content/114/51/13336.full.pdf.

[31]H. J. Berendsen, J. P. Postma, W. F. van Gunsteren, and J. Hermans, "Interaction models for water in relation to protein hydration," in *Intermolecular forces* (Springer, 1981) pp. 331–342.

[32]W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," The Journal of Chemical Physics **79**, 926–935 (1983).

[33]C. Vega, J. Abascal, E. Sanz, L. MacDowell, and C. McBride, "Can simple models describe the phase diagram of water?" Journal of Physics: Condensed Matter **17**, S3283 (2005).

[34]J. L. Abascal and C. Vega, "Dipole-quadrupole force ratios determine the ability of potential models to describe the phase diagram of water," Physical review letters **98**, 237801 (2007).

[35]J. R. Errington and P. G. Debenedetti, "Relationship between structural order and the anomalies of liquid water," Nature **409**, 318–321 (2001).

[36]B. Monserrat, J. G. Brandenburg, E. A. Engel, and B. Cheng, "Liquid water contains the building blocks of diverse ice phases," Nature Communications **11**, 5757 (2020).

[37]R. Capelli, A. Gardin, C. Empereur-mot, G. Doni, and G. M. Pavan, "A data-driven dimensionality reduction approach to compare and classify lipid force fields," *The Journal of Physical Chemistry B*, The Journal of Physical Chemistry B **125**, 7785–7796 (2021).

[38]P. Gasparotto, D. Bochicchio, M. Ceriotti, and G. M. Pavan, "Identifying and tracking defects in dynamic supramolecular polymers," The Journal of Physical Chemistry B **124**, 589–599 (2020).

[39]T. Bian, A. Gardin, J. Gemen, L. Houben, C. Perego, B. Lee, N. Elad, Z. Chu, G. M. Pavan, and R. Klajn, "Electrostatic co-assembly of nanoparticles with op-

AIP Publishing

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

The Journal of Chemical Physics

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

AIP Publishing

positely charged small molecules into static and dynamic superstructures," Nature Chemistry **13**, 940–949 (2021).

[40]H. Berendsen, J. Grigera, and T. Straatsma, "The missing term in effective pair potentials," Journal of Physical Chemistry **91**, 6269–6271 (1987).

[41]K. Takemura and A. Kitao, "Water model tuning for improved reproduction of rotational diffusion and nmr spectral density," The Journal of Physical Chemistry B **116**, 6279–6287 (2012), pMID: 22587421.

[42]L.-P. Wang, T. J. Martinez, and V. S. Pande, "Building force fields: An automatic, systematic, and reproducible approach," The Journal of Physical Chemistry Letters **5**, 1885–1891 (2014), pMID: 26273869.

[43]S. Izadi and A. V. Onufriev, "Accuracy limit of rigid 3-point water models," The Journal of Chemical Physics **145**, 074501 (2016).

[44]H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, and T. Head-Gordon, "Development of an improved four-site water model for biomolecular simulations: Tip4p-ew," The Journal of Chemical Physics **120**, 9665–9678 (2004).

[45]J. Abascal, E. Sanz, R. García Fernández, and C. Vega, "A potential model for the study of ices and amorphous water: Tip4p/ice," The Journal of Chemical Physics **122**, 234511 (2005).

[46]J. L. Abascal and C. Vega, "A general purpose model for the condensed phases of water: Tip4p/2005," The Journal of Chemical Physics **123**, 234505 (2005).

[47]R. Fuentes-Azcatl and J. Alejandre, "Non-polarizable force field of water based on the dielectric constant: Tip4p/ε," The Journal of Physical Chemistry B **118**, 1263–1272 (2014).

[48]S. Izadi, R. Anandakrishnan, and A. V. Onufriev, "Building water models: A different approach," The Journal of Physical Chemistry Letters **5**, 3863–3871 (2014), pMID: 25400877.

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

The Journal of Chemical Physics

AIP Publishing

[49]M. W. Mahoney and W. L. Jorgensen, "A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions," The Journal of Chemical Physics **112**, 8910–8922 (2000).

[50]S. W. Rick, "A reoptimization of the five-site water potential (tip5p) for use with ewald sums," The Journal of Chemical Physics **120**, 6085–6093 (2004).

[51]Y. Khalak, B. Baumeier, and M. Karttunen, "Improved general-purpose five-point model for water: Tip5p/2018," The Journal of Chemical Physics **149**, 224507 (2018).

[52]W. M. Haynes, *CRC handbook of chemistry and physics* (CRC press, 2014).

[53]D. P. Fernández, A. R. Goodwin, E. W. Lemmon, J. M. Levelt Sengers, and R. C. Williams, "A formulation for the static permittivity of water and steam at temperatures from 238 K to 873 K at pressures up to 1200 MPa, including derivatives and Debye-Hückel coefficients," Journal of Physical and Chemical Reference Data **26**, 1125–1166 (1997).

[54]L. B. Skinner, C. Huang, D. Schlesinger, L. G. Pettersson, A. Nilsson, and C. J. Benmore, "Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide q-range," The Journal of chemical physics **138**, 074506 (2013).

[55]A. P. Bartók, R. Kondor, and G. Csányi, "On representing chemical environments," Phys. Rev. B **87**, 184115 (2013).

[56]S. De, A. P. Bartók, G. Csányi, and M. Ceriotti, "Comparing molecules and solids across structural and alchemical space," Phys. Chem. Chem. Phys. **18**, 13754–13769 (2016).

[57]J. B. Kruskal, "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis," Psychometrika **29**, 1–27 (1964).

[58]E. A. Engel, A. Anelli, M. Ceriotti, C. J. Pickard, and R. J. Needs, "Mapping uncharted territory in ice from zeolite networks to ice structures," Nature Commu-

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

The Journal of Chemical Physics

nications **9**, 2173 (2018).

[59]C. G. Salzmann, P. G. Radaelli, E. Mayer, and J. L. Finney, "Ice xv: A new thermodynamically stable phase of ice," Phys. Rev. Lett. **103**, 105701 (2009).

[60]C. Adamo and V. Barone, "Toward reliable density functional methods without adjustable parameters: The pbe0 model," The Journal of Chemical Physics **110**, 6158–6170 (1999), https://doi.org/10.1063/1.478522.

[61]S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu," The Journal of Chemical Physics **132**, 154104 (2010), https://doi.org/10.1063/1.3382344.

[62]B. Cheng, E. A. Engel, J. Behler, C. Dellago, and M. Ceriotti, "Ab initio thermodynamics of liquid and solid water," Proceedings of the National Academy of Sciences **116**, 1110–1115 (2019), https://www.pnas.org/content/116/4/1110.full.pdf.

[63]J. R. Scherer, M. K. Go, and S. Kint, "Raman spectra and structure of water from-10 to 90. deg." The Journal of Physical Chemistry **78**, 1304–1313 (1974).

[64]Y. Maréchal, "The molecular structure of liquid water delivered by absorption spectroscopy in the whole ir region completed with thermodynamics data," Journal of Molecular Structure **1004**, 146–155 (2011).

[65]S. P. K. Pathirannahalage, N. Meftahi, A. Elbourne, A. C. Weiss, C. F. McConville, A. Padua, D. A. Winkler, M. C. Gomes, T. L. Greaves, Q. A. Besford, *et al.*, "A systematic comparison of the structural and dynamic properties of commonly used water models for molecular dynamics simulations," Chemrxiv (2021).

[66]B. Cheng, R.-R. Griffiths, S. Wengert, C. Kunkel, T. Stenczel, B. Zhu, V. L. Deringer, N. Bernstein, J. T. Margraf, K. Reuter, *et al.*, "Mapping materials and molecules," Accounts of Chemical Research **53**, 1981–1991 (2020).

[67]A. Offei-Danso, A. Hassanali, and A. Rodriguez, "High dimensional fluctuations in liquid water: Combining chemical intuition with unsupervised learning," arXiv

AIP Publishing

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

The Journal of Chemical Physics

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

preprint arXiv:2112.11894 (2021).

[68]G. Camisasca, D. Schlesinger, I. Zhovtobriukh, G. Pitsevich, and L. G. Pettersson, "A proposal for the structure of high-and low-density fluctuations in liquid water," The Journal of Chemical Physics **151**, 034508 (2019).

[69]G. E. Walrafen, W. H. Yang, and Y. C. Chu, "Raman evidence for the clathrate-like structure of highly supercooled water," in *Supercooled Liquids*, ACS Symposium Series, Vol. 676 (American Chemical Society, 1997) Doi:10.1021/bk-1997-0676.ch021 21, pp. 287–308.

[70]V. Rizzi, L. Bonati, N. Ansari, and M. Parrinello, "The role of water in host-guest interaction," Nature Communications **12**, 93 (2021).

[71]L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez, "Packmol: a package for building initial configurations for molecular dynamics simulations," Journal of computational chemistry **30**, 2157–2164 (2009).

[72]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**, 19–25 (2015).

[73]G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," The Journal of Chemical Physics **126**, 014101 (2007).

[74]M. Bernetti and G. Bussi, "Pressure control using stochastic cell rescaling," The Journal of Chemical Physics **153**, 114107 (2020).

[75]U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, "A smooth particle mesh ewald method," The Journal of Chemical Physics **103**, 8577–8593 (1995).

[76]M. R. Shirts, D. L. Mobley, J. D. Chodera, and V. S. Pande, "Accurate and efficient corrections for missing dispersion interactions in molecular simulations," The journal of physical chemistry B **111**, 13052–13063 (2007).

AIP Publishing

Accepted to J. Chem. Phys. 10.1063/5.0088599

ACCEPTED MANUSCRIPT

The Journal of Chemical Physics

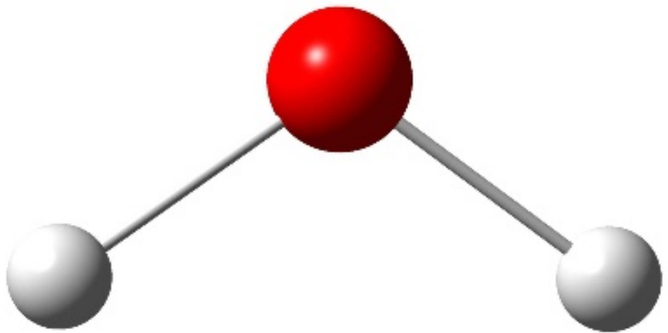
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI:10.1063/5.0088599

[77]G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, "Plumed 2: New feathers for an old bird," Computer Physics Communications **185**, 604–613 (2014).

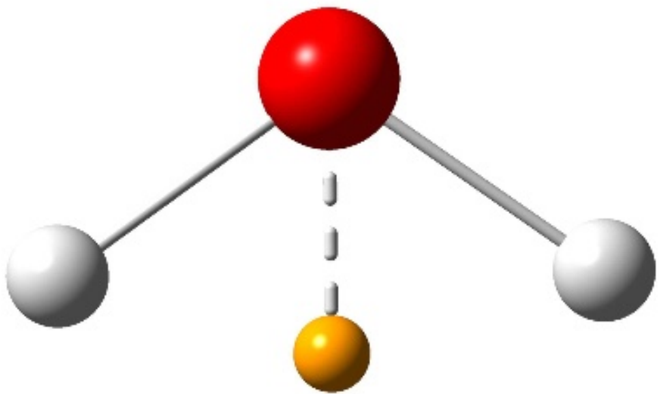
[78]M. Bonomi, G. Bussi, C. Camilloni, G. A. Tribello, P. Banáš, A. Barducci, M. Bernetti, P. G. Bolhuis, S. Bottaro, D. Branduardi, R. Capelli, P. Carloni, M. Ceriotti, A. Cesari, H. Chen, W. Chen, F. Colizzi, S. De, M. De La Pierre, D. Donadio, V. Drobot, B. Ensing, A. L. Ferguson, M. Filizola, J. S. Fraser, H. Fu, P. Gasparotto, F. L. Gervasio, F. Giberti, A. Gil-Ley, T. Giorgino, G. T. Heller, G. M. Hocky, M. Iannuzzi, M. Invernizzi, K. E. Jelfs, A. Jussupow, E. Kirilin, A. Laio, V. Limongelli, K. Lindorff-Larsen, T. Löhr, F. Marinelli, L. Martin-Samos, M. Masetti, R. Meyer, A. Michaelides, C. Molteni, T. Morishita, M. Nava, C. Paissoni, E. Papaleo, M. Parrinello, J. Pfaendtner, P. Piaggi, G. Piccini, A. Pietropaolo, F. Pietrucci, S. Pipolo, D. Provasi, D. Quigley, P. Raiteri, S. Raniolo, J. Rydzewski, M. Salvalaglio, G. C. Sosso, V. Spiwok, J. Šponer, D. W. H. Swenson, P. Tiwary, O. Valsson, M. Vendruscolo, G. A. Voth, and A. White, "Promoting transparency and reproducibility in enhanced molecular simulations," Nature Methods **16**, 670–673 (2019).

[79]L. Himanen, M. O. J. Jäger, E. V. Morooka, F. Federici Canova, Y. S. Ranawat, D. Z. Gao, P. Rinke, and A. S. Foster, "DScribe: Library of descriptors for machine learning in materials science," Comput. Phys. Commun. **247**, 106949 (2020).

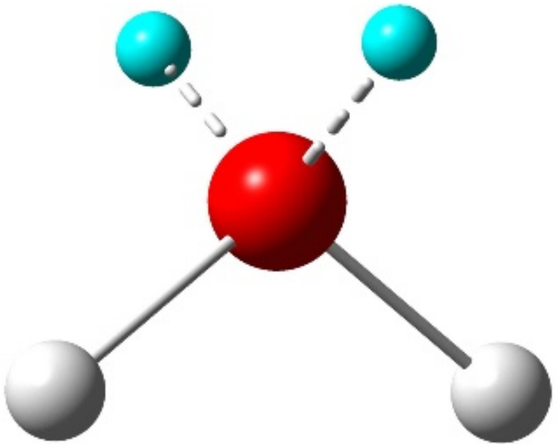
AIP Publishing



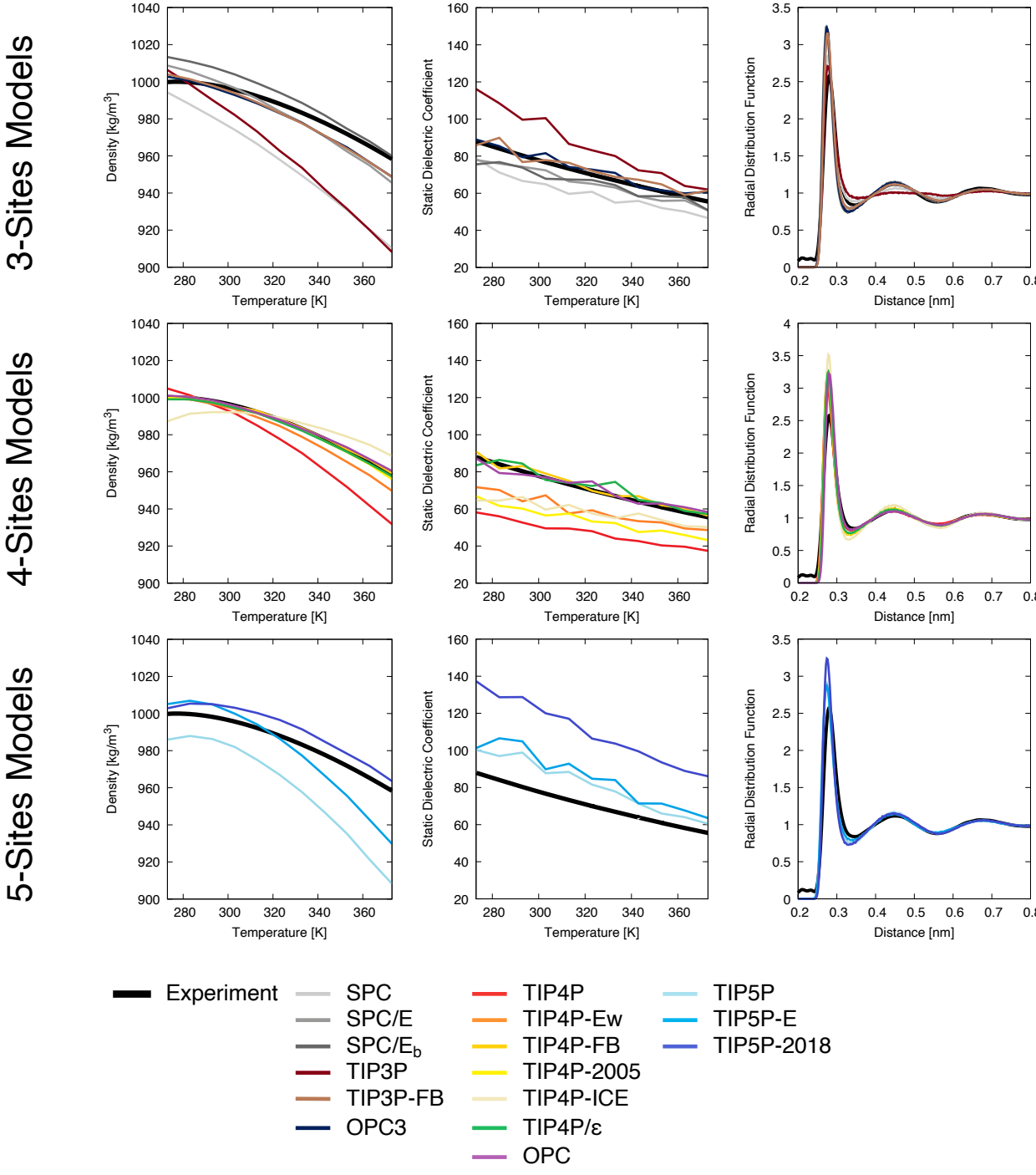
3-sites model

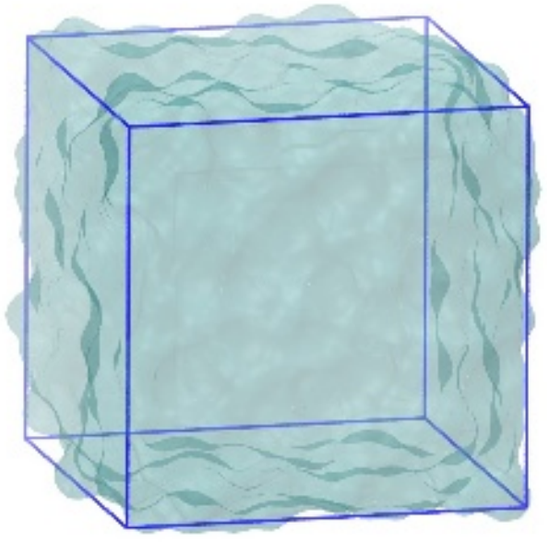


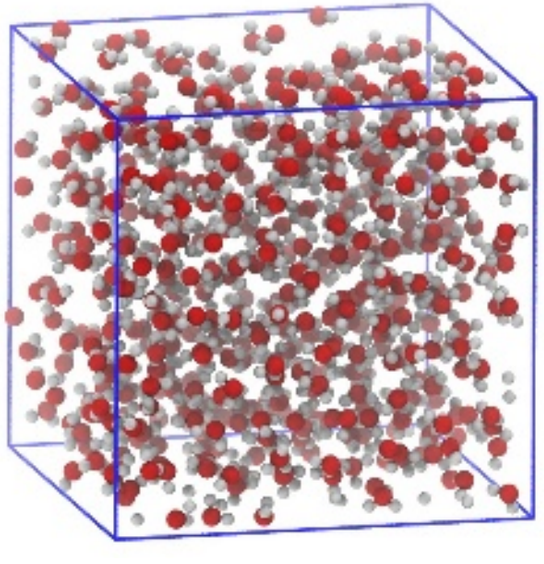
4-sites model

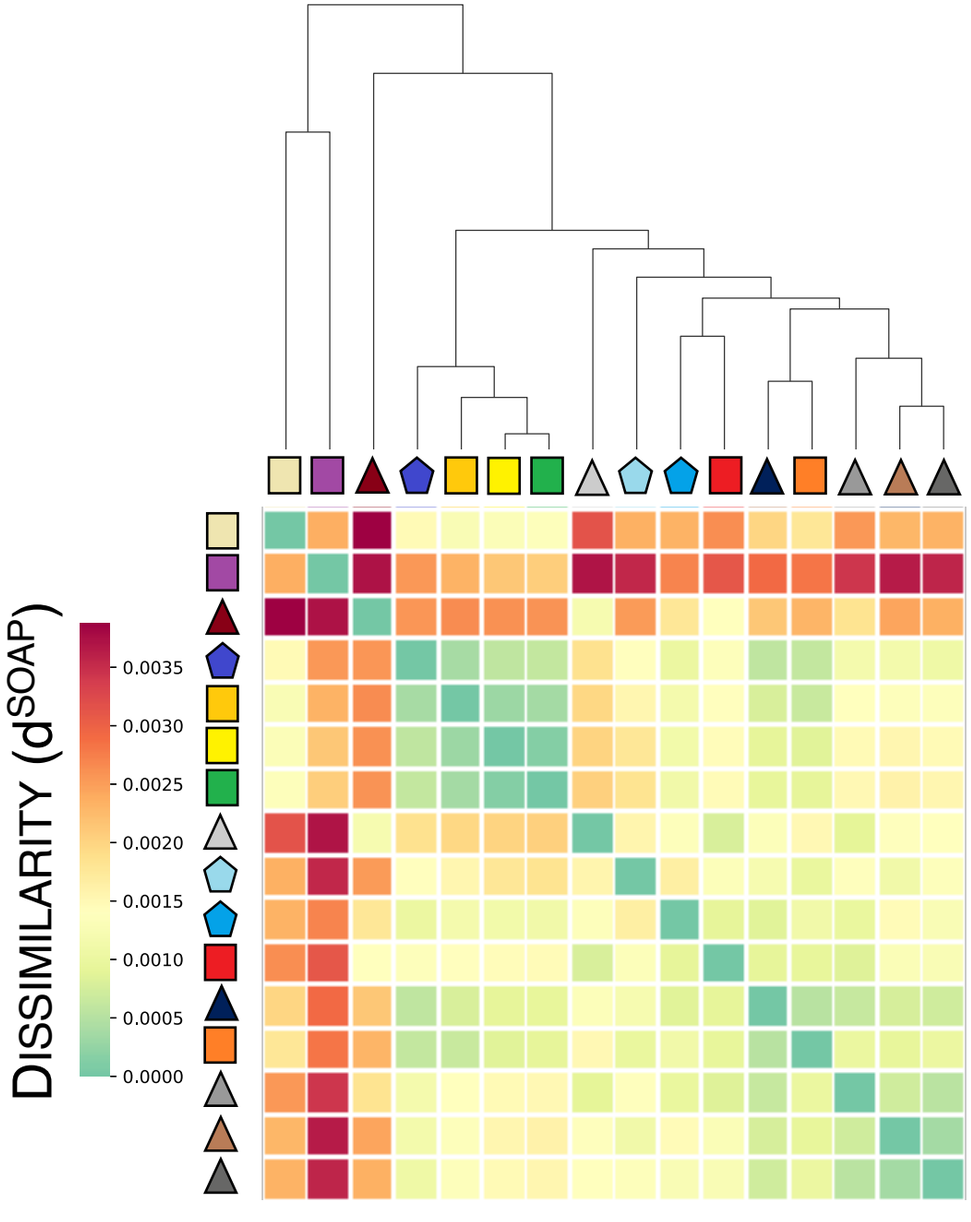


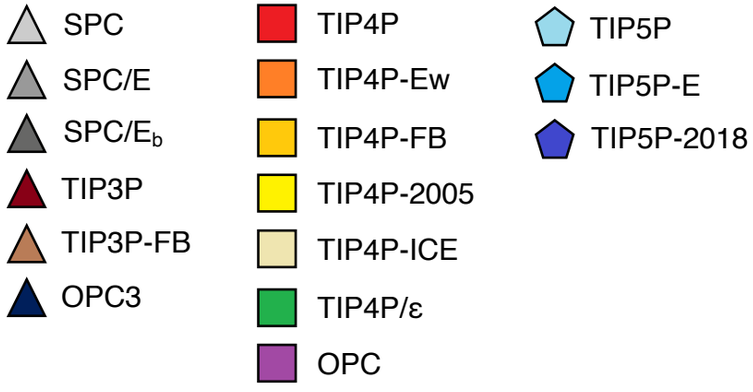
5-sites model

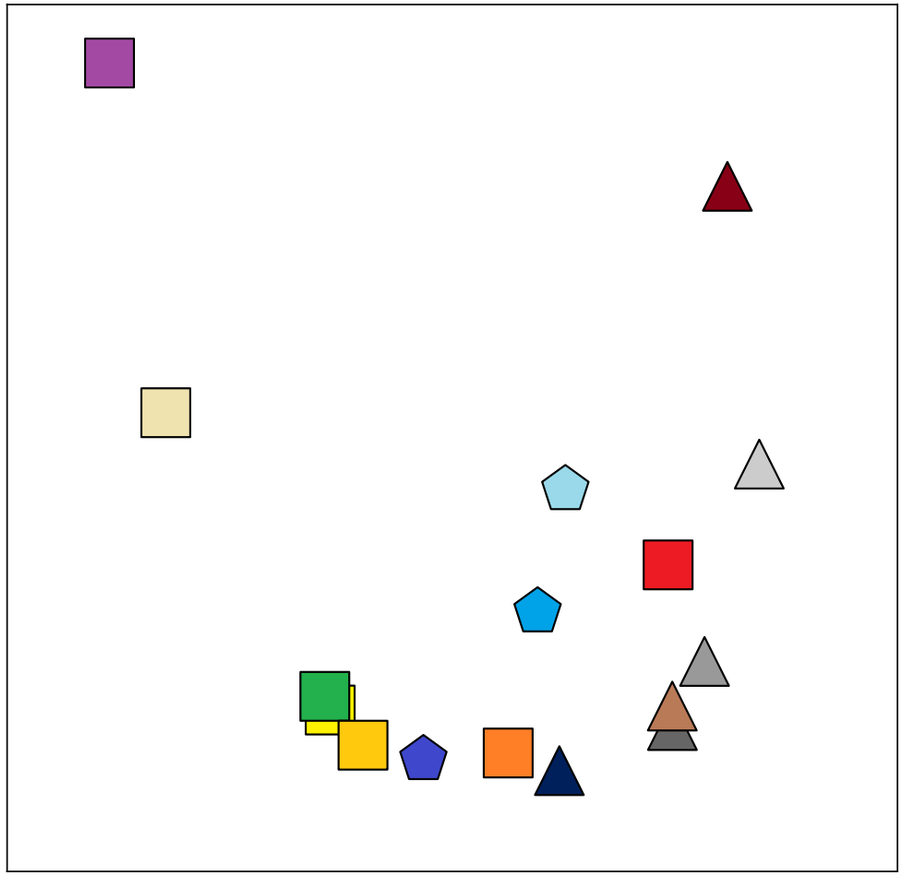

3-Sites Models
4-Sites Models
5-Sites Models
Density [kg/m3]
Static Dielectric Coefficient
Radial Distribution Function
Temperature [K]
Distance [nm]
Experiment
SPC
SPC/E
SPC/Eb
TIP3P
TIP3P-FB
OPC3
TIP4P
TIP4P-Ew
TIP4P-FB
TIP4P-2005
TIP4P-ICE
TIP4P/ε
OPC
TIP5P
TIP5P-E
TIP5P-2018

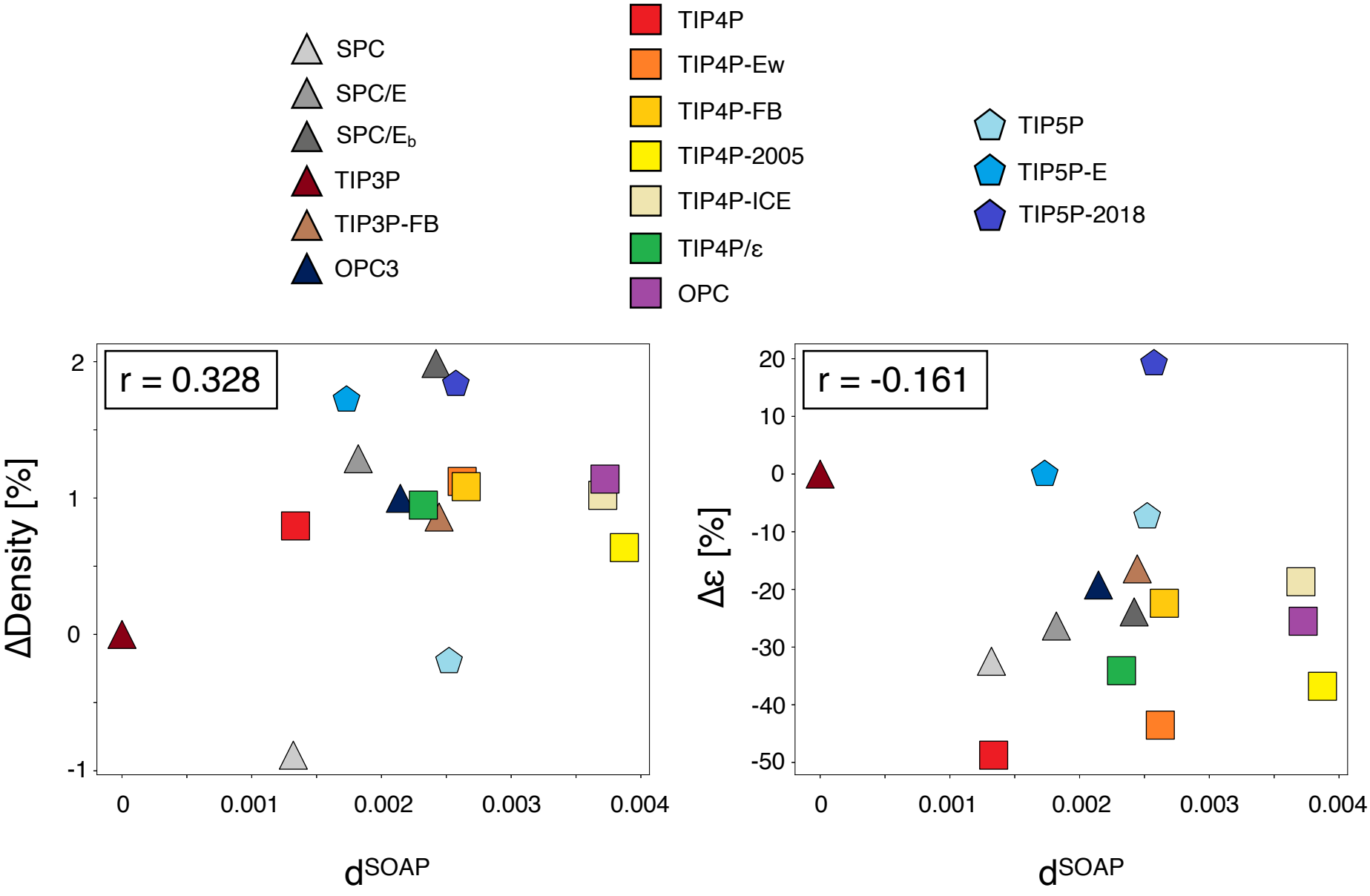


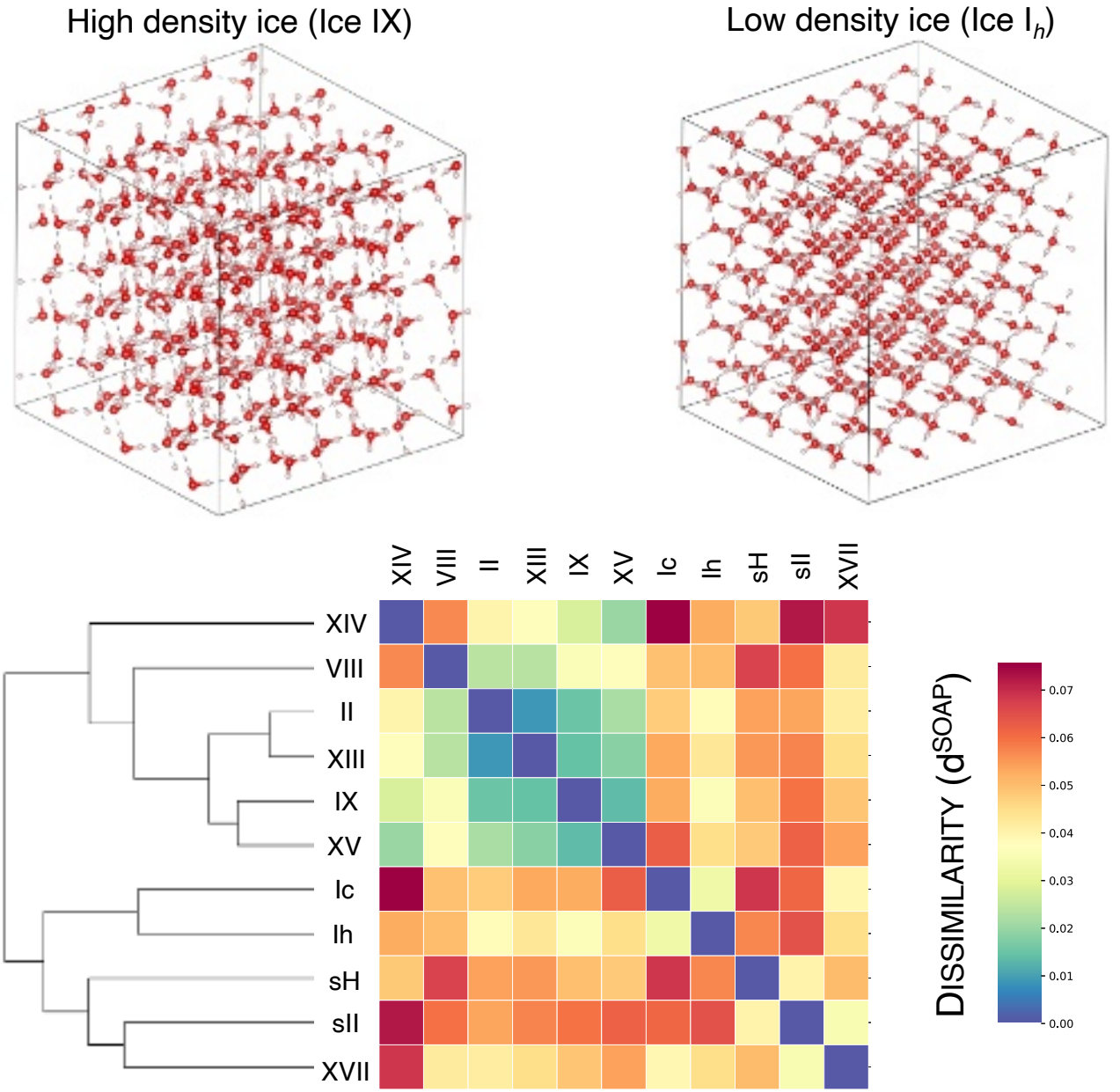


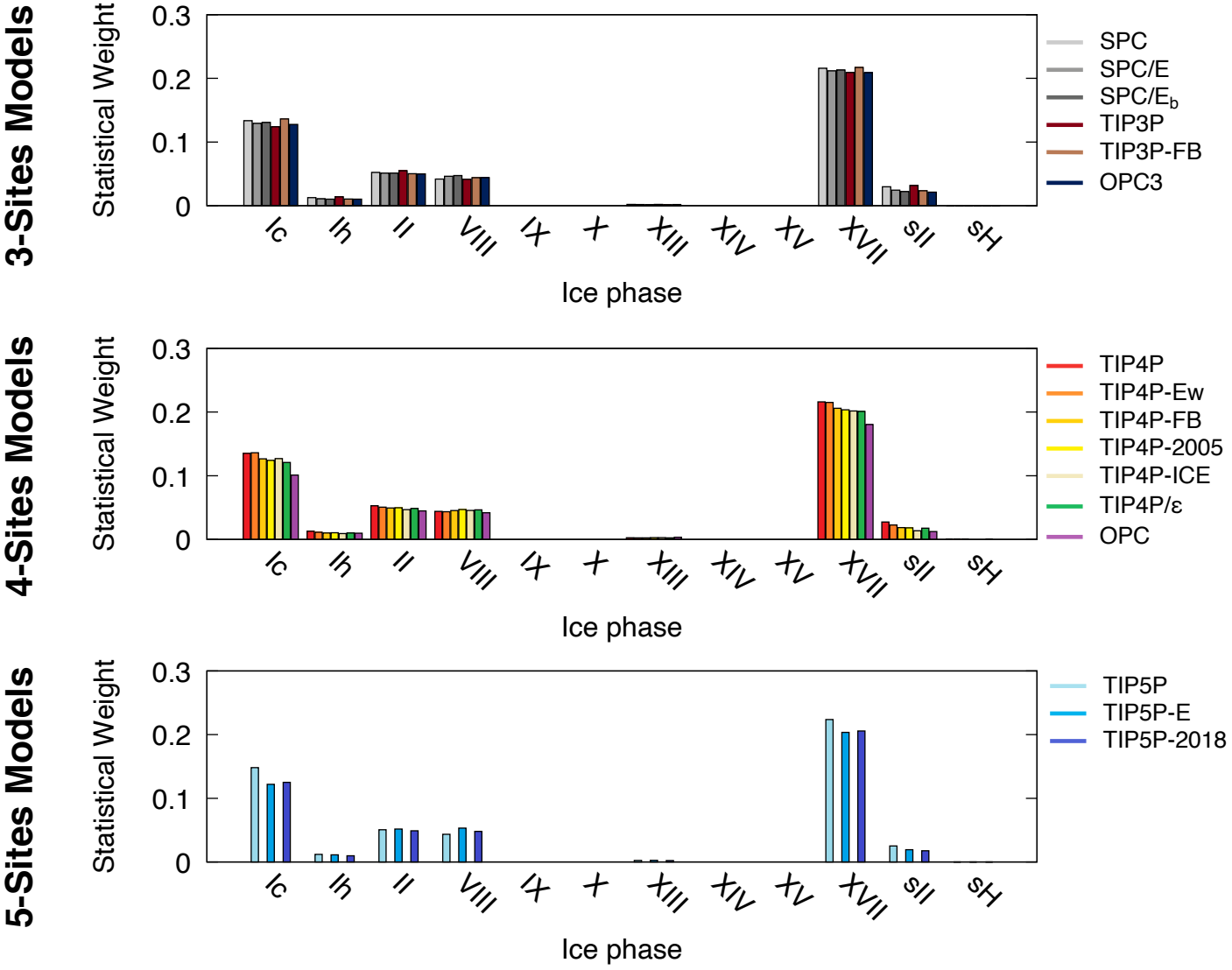

DISSIMILARITY (d^{SOAP})
0.0035
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000

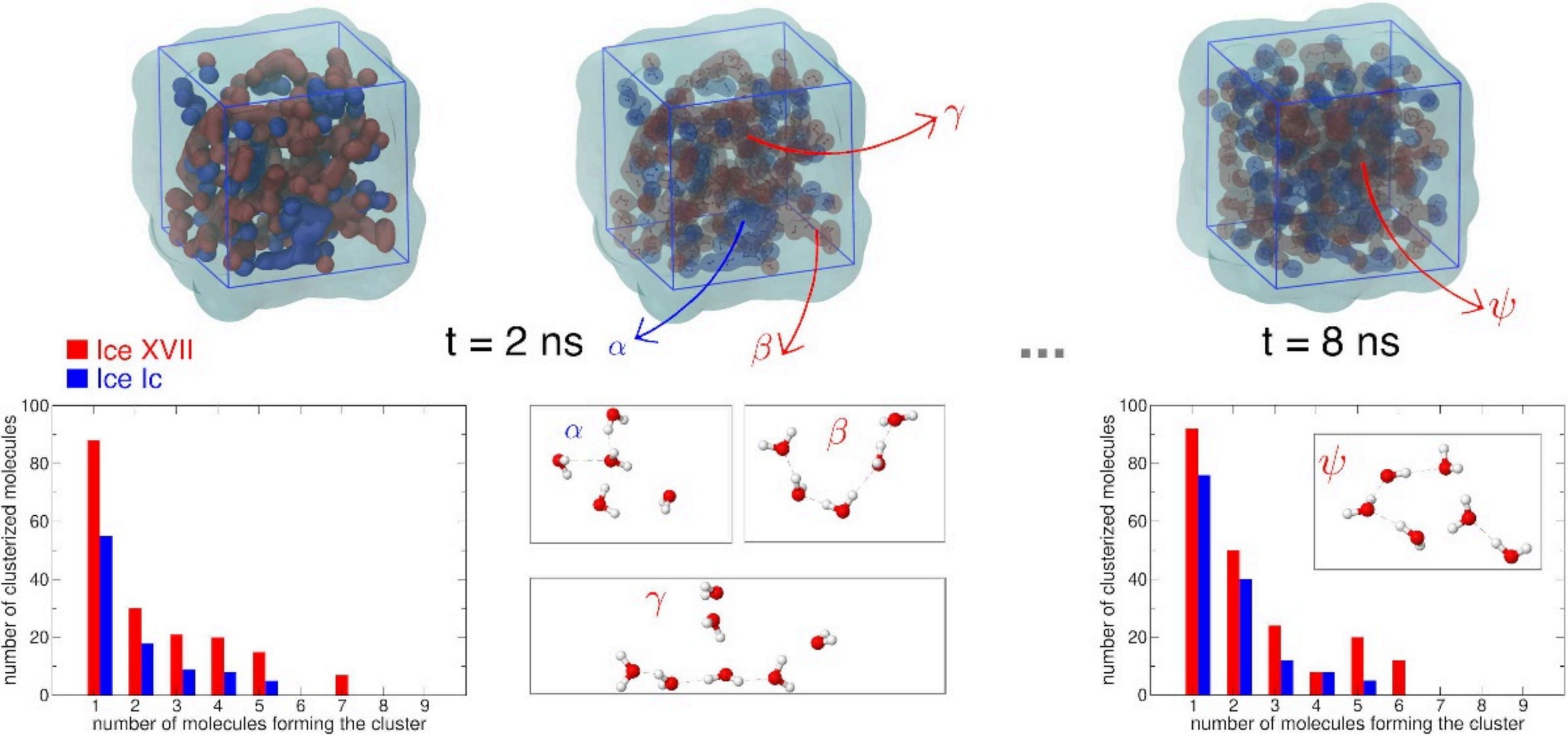

SPC
SPC/E
SPC/E_b
TIP3P
TIP3P-FB
OPC3
TIP4P
TIP4P-Ew
TIP4P-FB
TIP4P-2005
TIP4P-ICE
TIP4P/ε
OPC
TIP5P
TIP5P-E
TIP5P-2018




SPC
SPC/E
SPC/E$_b$
TIP3P
TIP3P-FB
OPC3
TIP4P
TIP4P-Ew
TIP4P-FB
TIP4P-2005
TIP4P-ICE
TIP4P/ε
OPC
TIP5P
TIP5P-E
TIP5P-2018
r = 0.328
ΔDensity [%]
2
1
0
-1
0
0.001
0.002
0.003
0.004
d^{SOAP}
r = -0.161
Δε [%]
20
10
0
-10
-20
-30
-40
-50
0
0.001
0.002
0.003
0.004
d^{SOAP}


High density ice (Ice IX)
Low density ice (Ice I_h)
XIV
VIII
II
XIII
IX
XV
Ic
Ih
sH
sII
XVII
DISSIMILARITY (d^{SOAP})
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00


3-Sites Models
Statistical Weight
0.3
0.2
0.1
0
Ic
Ih
II
VIII
IX
X
XIII
XIV
XV
XVII
sII
sH
Ice phase
SPC
SPC/E
SPC/Eb
TIP3P
TIP3P-FB
OPC3
4-Sites Models
Statistical Weight
0.3
0.2
0.1
0
Ic
Ih
II
VIII
IX
X
XIII
XIV
XV
XVII
sII
sH
Ice phase
TIP4P
TIP4P-Ew
TIP4P-FB
TIP4P-2005
TIP4P-ICE
TIP4P/ε
OPC
5-Sites Models
Statistical Weight
0.3
0.2
0.1
0
Ic
Ih
II
VIII
IX
X
XIII
XIV
XV
XVII
sII
sH
Ice phase
TIP5P
TIP5P-E
TIP5P-2018


Ice XVII
Ice Ic
t = 2 ns
α
β
γ
...
t = 8 ns
ψ
number of clusterized molecules
number of molecules forming the cluster
0
20
40
60
80
100
1
2
3
4
5
6
7
8
9