Bathochromic Shift in the UV-Visible Absorption Spectra of Phenols at Ice Surfaces: Insights from First-Principles Calculations

Fernanda C. Bononi,[†] Zekun Chen,[†] Dario Rocca,[‡] Oliviero Andreussi,[¶] Ted Hullar,[§] Cort Anastasio,[§] and Davide Donadio[*,†]

[†]Department of Chemistry, University of California Davis

[‡]Université de Lorraine, CNRS, LPTC, F-54000, Nancy, France

[¶]Department of Physics, University of North Texas

[§]Department of Land, Air and Water Resources, University of California Davis

E-mail: ddonadio@ucdavis.edu

Abstract

Some organic pollutants in snowpack undergo faster photodegradation than in solution. One possible explanation for such effect is that their UV-visible absorption spectra are shifted toward lower energy when the molecules are adsorbed at the air-ice interface. However, such bathochromic shift is difficult to measure experimentally. Here we employ a multiscale/multimodel approach that combines classical and first-principles molecular dynamics, quantum chemical methods and statistical learning to compute the light absorption spectra of two phenolic molecules in different solvation environments at the relevant thermodynamic conditions. Our calculations provide an accurate estimate of the bathochromic shift of the lowest-energy UV-visible absorption band when these molecules are adsorbed at the air-ice interface, and they shed light into its molecular origin.

Introduction

Although primarily composed of ice, snowpacks contain small regions of disordered water molecules, where most solutes reside.[1–3] They are either bulk liquid-like regions (LLR) located within the ice matrix or quasi-liquid layers (QLL) at the surface of ice crystals. These characteristics make snow and ice unique environments for environmentally relevant chemical reactions such as the photodegradation of pollutants, which may be transformed into more volatile molecules that can then be released into the atmosphere.[4,5] While extremely important, to date only a few direct photochemical reactions have been studied on ice and snow. Some works show an increase in the rate of photodegradation at the air-ice interface compared to in solution (e.g., for naphthalene, toluene, ethylbenzene and xylene),[6,7] while others show that photodegradation proceeds at similar rates at the air-ice interface and in solution (e.g., for nitrate, nitrite, hydrogen peroxide, anthracene, pyrene and fluoranthene).[8,11–13] A possible reason for the observed photodegradation enhancement for molecules solvated at the air-ice interface is that their molar absorptivities are shifted to lower energy (bathochromic

shift), where there is a higher sunlight photon flux. This is a hypothesis that has been previously examined for a variety of chemicals.[7,14–18] Since the summer polar actinic flux at the surface of the Earth increases by a factor of approximately 1 million between 290 and 310 nm,[19] compounds that absorb light below this range of wavelengths would normally not be readily photolyzed by sunlight in solution, but could undergo degradation in snow if there is a bathochromic shift in their absorbance toward wavelengths with higher photon flux.

Due to the high concentrations required to measure light absorbance of chemicals, these measurements are extremely challenging for molecules at the air-ice interface and require concentrations much higher than those of pollutants in the natural environment. Increasing the concentration of molecules at surfaces to the measurable threshold would lead to high surface density and possibly molecular aggregation, which may affect the features of UV-visible spectra.[14,20] For this reason, molecular modeling becomes an essential tool to understand the solvation of small molecules and ions at water and ice surfaces,[21–23] and to predict their optical spectra.[24] A number of former studies supplemented experimental observations with quantum chemical calculations of the excited states of molecules optimized in the gas phase.[14,15,17,18] Whereas these calculations provide an accurate estimate of electronic transition energies, they fail to account for the solvation effects in complex environments at the relevant thermodynamic conditions. Accurate electronic excitations for gas-phase molecules may be computed by many-body quantum chemical approaches, such as coupled-clusters[25] or the Bethe-Salpeter equation,[26] but methods based on time dependent density functional theory (TDDFT) also provide a reasonably good approximation of UV-visible absorption spectra at lower computational cost,[27–29] thus enabling ensemble averaging.[30,31]

The absorption spectra of solvated molecules, just like many other physical properties, depend critically on solute-solvent interactions. The interaction with the solvent affects the molecular configuration of the solute and its electron density, both by means of local interactions – local electric fields – and through the long-range dielectric environment. The effect of embedding the solute in a solvent medium is customarily approximated using continuum

solvation models (CSM).[32,33] In turn, for explicit solvent schemes, it has been proven that in most cases it is necessary to include multiple solvation shells to converge excitation energies, thus leading to very large first-principles calculations.[34,35] A viable alternative to treat both long-range electrostatic effects and local interactions, e.g. hydrogen-bonding, at an accessible computational cost consists of supplementing the continuum models with a shell of explicit solvent molecules.[36,37] Furthermore, the positions of the absorption peaks and their line shapes are determined by the temperature-driven fluctuations of the solute molecule, and possibly of its solvation shell.[38–41] The line broadening produced by molecular vibrations should be, in principle, treated at the quantum mechanical level, i.e., computing correlation functions of quantum operators. However, a more viable approach for systems near and above room temperature consists of acquiring absorption spectra as thermodynamic averages over a sufficiently large number of frames along molecular dynamics (MD) trajectories, obtained using either classical empirical potentials[40,41] or first-principles MD (FPMD) based on density functional theory (DFT).[39] FPMD overcomes the transferability issues of classical models, and it has been employed together with TDDFT and a self-consistent continuum solvation model (SCCS),[42,43] to compose a multimodel approach that has proven successful to predict the color and the optical spectra of natural dyes in solution.[30,44]

A model to establish the subtle differences in the UV-visible absorption spectra as a function of the solvation environment, should consider both electrostatic embedding and temperature fluctuations, and also provide an accurate description of the solvation environment. For the case of molecules adsorbed at the air-ice interface, it is then necessary to start from a reliable model of the quasi-liquid layer (QLL) at the ice surface[45,46] and to probe how the solutes are solvated in this complex environment. Hence the need to supplement the state-of-the-art computational approach discussed above with classical MD simulations able to probe suitably larger size and time scales.[47]

In this work we employ molecular modeling to assess the UV-visible absorption spectra of two common organic pollutants – phenol and 2-methoxyphenol (guaiacol) – in solution and at

the air-ice interface. Our goal is to examine whether these molecules exhibit a spectral shift at the ice surface, which would lead to higher photon uptake and enhanced photodegradation in snowpacks. To achieve this goal we developed and tested a multiscale/multimodel approach, in which we first investigate the structural and thermodynamic features of the solvation of phenol and guaiacol in bulk water and at the air-ice interface by classical MD, using a realistic well-characterized ice surface molecular model.[47] We eventually refine the details of the solvation shells of these molecules in water and on ice using smaller models treated by Born-Oppenheimer FPMD. We then compute the UV-visible absorption spectra by TDDFT, averaging over hundreds of frames from the FPMD trajectories and using different types of solvation models, including a new SCCS setup designed to correctly represent dielectric screening at interfaces. The least absolute shrinkage and selection operator (LASSO)[48] algorithm is finally employed to refine the lowest energy band of the absorption spectra and to gain insight into the physical origin of the observed spectral shift.

Using this approach, which explicitly takes into account solvent environment and thermodynamic effects, we find that the absorption band corresponding to the lowest energy electronic transition undergoes a significant solvatochromic red shift when the molecules are adsorbed at the air-ice interface, thus enhancing the absorption rate of solar light and resulting photochemistry. The statistical learning analysis suggests that the main reason for the observed solvatochromic shift lies in the differences in the atomic configurations determined by different solvation environments.

Methods

Our sequential multiscale approach comprises the following steps: (i) classical MD simulations to characterize the structure and dynamics of phenol and guaiacol at the air-ice interface and to compute solvation free energies; (ii) FPMD simulations to refine the structure of the solvation shell and to obtain molecular configurations at finite temperature for the following

step; (iii) TDDFT calculations of the absorption spectra using the SCCS implicit solvent model with an increasing number of explicit water molecules; (iv) fitting of a LASSO regression model, trained on TDDFT calculations, to compute the lowest transition energy as a function of the molecular configuration. Each step is described in the following subsections.

Classical Molecular Dynamics

Classical MD is used to probe the structure and dynamics of phenol and guaiacol at the surface of ice slabs and in bulk water, and to calculate adsorption and solvation free energies. Simulation models consist of ~ 1000 water molecules and MD trajectories extend to hundreds of nanoseconds. The equations of motion are integrated with the velocity Verlet algorithm with a time step of 0.5 fs. Simulations are carried out in the constant volume canonical ensemble (NVT), where the temperature is controlled by stochastic velocity rescaling with a relaxation time of 1 ps,[49] and long-range electrostatics are computed using the particle-particle particle-mesh solver (PPPM).[50] MD simulations are performed using the LAMMPS package.[51] Water molecules are modeled using the fixed-charge TIP4P/Ice force field, which was parameterized to reproduce the experimental melting temperature of ice at ambient conditions, and it accurately reproduces the equations of state of water and ice near the melting temperature and at mild supercooling.[52] Phenol and guaiacol are modeled using the generalized Amber force field (GAFF),[53] using the AM1-BCC charge model[54] in order to calculate the atomic point charges, which we later reparameterized so to reproduce the experimental solvation free energy in TIP4P/Ice water at room temperature. These calculations are carried out by free energy perturbation (FEP),[55,56] and details are provided in SI (Figures S1, S2 and S3).

For the solution model, a molecule of phenol or guaiacol was added to a cubic box containing 1000 water molecules and allowed to equilibrate before running free energy calculations followed by production runs of 100 ns at 300 K. For the ice model, a proton-disordered hexagonal ice model containing 1536 water molecules was generated using a Monte Carlo

6

procedure to minimize the total electric dipole,[57] and it was cut into a slab 45 Å thick along the basal plane. The orthorhombic simulation box measures 36.393×31.517×144 Å^3 and periodic boundary conditions (PBC) are applied in all three directions. The ∼ 100 Å vacuum separating periodic images in the z direction is sufficient to avoid finite-size artifacts.[47] The system was equilibrated for 100 ns at 263 K, so to obtain a stable QLL at the surface. The QLL is about two ice bilayers thick, in agreement with previous experimental observations and MD simulations.[45,47] After equilibration, we added one phenol or guaiacol molecule on each surface of the slab, positioning the OH group within hydrogen bonding distance from the water molecules of the ice slab, and allowed the system to equilibrate for 2 ns. Production runs of 100 ns were then obtained at 263 K for both phenol and guaiacol.

First-Principles Molecular Dynamics

FPMD simulations of several tenths of picoseconds are performed for molecules in both solution and at the air-ice interface in smaller simulation boxes with explicit solvent. The purpose of these simulations is to attain a more accurate description of the solvation environment and to sample the configuration of the solvated molecule with a computational framework consistent with that used for the excited state calculations. DFT-based FPMD simulations were carried out using the Quickstep approach implemented in the CP2K package,[58,59] employing the Perdew-Burke-Ernzerhof (PBE) generalized gradient approximation (GGA) for the exchange and correlation functional.[60] Valence Kohn-Sham orbitals are expanded on a double-ζ localized basis set[61] in real space, and core states are treated implicitly using Geodecker-Teter-Hutter pseudopotentials.[62] Plane waves up to a cutoff energy of 300 Ry are used to represent the density in reciprocal space.

FPMD simulations of phenol and guaiacol in aqueous solution were carried out in a cubic simulation box (12.8 Å) containing 64 water molecules with PBC. Simulations of each organic molecule adsorbed on the ice surface are performed putting one molecule on the surface of an ice slab made of 192 water molecules in a periodic orthorhombic cell with dimensions 18,

15.589 and 80 Å. Both aqueous solutions and ice slabs were prepared using classical MD with the empirical force fields described above. Eventually, the systems were equilibrated at the target temperature using CP2K, and 50 ps long production runs in the constant volume canonical ensemble (NVT)[49] were carried out. Aqueous solutions are equilibrated at room temperature (300 K), whereas ice slabs are kept at 263 K. In these simulations hydrogen atoms are replaced with deuterium, thus allowing a relatively large timestep of 0.5 fs to integrate the equations of motion. Up to 200 statistically independent molecular configurations obtained from each one of these trajectories are used to compute UV-visible absorption spectra with the ensemble method.[30,44]

Excited State Calculations by Time-Dependent Density Functional Theory

Finite temperature spectra are calculated as time averages of single spectra of individual frames over a 50 ps FPMD trajectory. For each molecular configuration selected from the FPMD runs in solution and on ice, we computed the absorption spectra using TDDFT in the time-dependent density functional perturbation theory implementation (TDDFPT)[63] in the Quantum-Espresso package.[64,65] Absorption spectra are computed in a two-step process that consists of a single-point self-consistent ground state total energy calculation followed by the calculation of excitation energies and oscillator strengths solving TDDFPT equations with the recursive Lanczos algorithm.[63] Since we are interested in the relative shifts between the spectra of a molecule in aqueous solution and at the air-ice interface, we adopt the PBE functional[60] for the excitation energy calculations, rather than more accurate, but much more computationally expensive, hybrid functionals. It was indeed shown that the main effect of using hybrid functionals is a nearly rigid shift of the lower energy excitation, and it is possible to rescale the GGA spectra to those calculated with hybrids with a simple morphing procedure.[44] These calculations are performed in the plane-wave pseudopotential formalism: nuclei and core states are represented by ultrasoft pseudopotentials,[66] and a

plane-wave basis set with an energy cutoff of 30 Ry for the Kohn-Sham orbitals and a cutoff of 240 Ry for the electron density. Calculations are performed for isolated systems in tetragonal simulation cells, with dimension $25 \times 25 \times 50$ Å^3, with long-range electrostatic corrections for non-periodic systems.[67]

Solvation Environment

To compute the absorption spectra of molecules either in solution or at the air-ice interface, we need to include the electrostatic screening effect of the solvation medium. This may be achieved in two different ways: either using an explicit embedding scheme or using an implicit polarizable continuum solvent.[44,68] The latter presents the advantage of providing the correct long-range limit.[30,36,69] Furthermore, local effects may be refined by supplementing the implicit solvent model with one or more explicit solvation shells. The SCCS implicit solvent model[42,43,70–74] can capture the dominant electrostatic effects in water solvation, while relying on a limited number of parameters. The extension of SCCS to model the solvatochromic shifts in electronic excitations within the linear response TDDFT formalism produces optical spectra in perfect agreement with similar approaches in the quantum-chemistry literature.[30] In addition, SCCS can handle both isolated and periodic, or partially periodic systems. Moreover, it relies on a smoothly varying definition of the boundary between the quantum-mechanical solute and the embedding continuum medium. For these reasons, it is ideally suited to model the smooth transition in non-homogeneous environments, such as water-vacuum, water-ice, or ice-vacuum interfaces. The SCCS model is implemented in the ENVIRON plugin[75] to Quantum Espresso.[76]

For calculations in bulk, we modeled the environment as a homogeneous dielectric medium characterized by the dielectric constant of water at 300 K, which is equal to 78.5.[77] For calculations on the ice surface, we developed an extension of SCCS, where dielectric regions of arbitrary shape and dimensionality are included in the simulation cell. In the original SCCS definition,[42] the dielectric function varies from a value of 1 inside the quantum-mechanical

region of space, to a constant environment permittivity ϵ_0, where the electronic density of the solute becomes negligible. Here, in addition to the continuum interface due to the presence of the solute, a position-dependent permittivity of the environment, $\epsilon_0(\mathbf{r})$, is introduced. Starting from a uniform homogeneous background $\epsilon_0^0(\mathbf{r}) \equiv \epsilon_0^0$, a step-wise procedure is designed to include a user-defined number of heterogeneous media. For each dielectric region i, centered at position $\mathbf{R}_i$, the environment permittivity is modified as follows

$$\epsilon_0^i(\mathbf{r}) = \epsilon_0^{i-1}(\mathbf{r}) + \left(\epsilon_0^i - \epsilon_0^{i-1}(\mathbf{r})\right) h_i\left(\mathbf{r} - \mathbf{R}_i\right), \tag{1}$$

where ϵ_0^i is the target dielectric permittivity of region i, while h_i is a smooth function that varies from one to zero as we move outside of the region. In particular, a definition in terms of the complementary error function is adopted

$$h_i(\mathbf{r}) = \frac{1}{2}\text{erfc}\left(\frac{|\mathbf{r}|_d - W_i}{\Delta_i}\right), \tag{2}$$

where W_i represents the width of the region, Δ_i is the spread of the interface, while d controls the dimensionality of the interface. For simulating the ice surface, a two-dimensional $(d = 2)$ region is added to the simulation, with coordinates and width based on our FPMD models, an interface smoothing thickness $\Delta_i = 2$ Å, a dielectric permittivity of $\epsilon_0 = 107$ and an optical permittivity of $\epsilon_{opt} = 1.716$ (Figure 1).

Statistical Learning

To obtain an accurate absorption line shape from the lowest electronic excitation, we would need to perform several thousand TDDFT calculations. To abate the inherent computational cost, we exploit a machine learning (ML) model, trained on the TDDFT calculations described above, to predict first excitation energies from molecular configurations. As the training set is made of a few hundred calculated frames and the optimization space is highly dimensional, we opted for the least absolute shrinkage and selection operator (LASSO), a
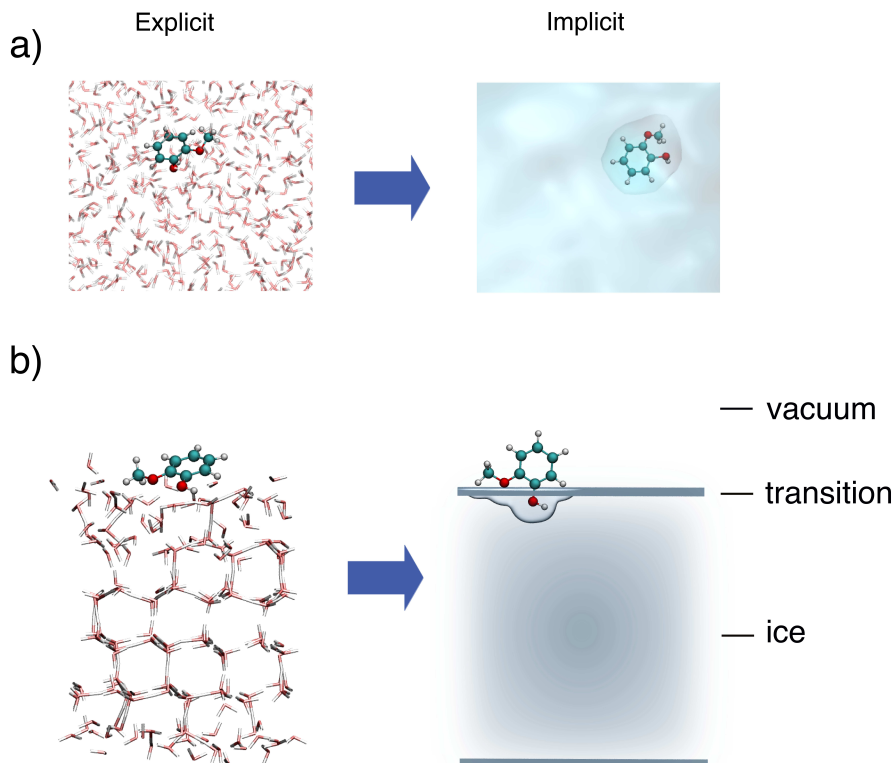
Figure 1: Scheme of the transition from explicit to implicit solvation models for a bulk aqueous solution (a) and for the air-ice interface (b). In both the three-dimensional representations of the implicit solvent models the guaiacol molecule occupies a cavity with the dielectric constant of vacuum embedded in a continuum with the dielectric constant of the solvent medium. In the case of adsorption on ice the molecule sits at the interface between a region with the dielectric constant of ice and one with dielectric constant of vacuum. The two regions are connected by a smooth transition region with a parameterized thickness (Δ_i), which is represented by the grey area.

linear regression model with norm one regularization.[48,78]

The dimensionality of the configurational space corresponds to the number of coordinates that define phenol (13 atoms, 39 coordinates) or guaiacol (17 atoms, 51 coordinates). Dimensionality is reduced by aligning each molecule with respect to a reference and by ignoring the coordinates of the hydrogen atoms, except the one on the hydroxyl group for both molecules. Hence the optimization space for phenol is 24-dimensional and for guaiacol 30-dimensional. We found that fitting directly excitation wavelengths using inverse coordinates as descriptors is optimal to minimize the error. The regularization parameter was set to 10^{-8}, and fitting

and testing were carried out using an 8-fold cross validation scheme. The statistical metrics of R^2 and mean absolute error (MAE) were used to assess the performance of the ML model. Figure S4 in SI shows that the ML model has an average testing R^2 of 0.781 and MAE of 1.94 nm for guaiacol and 0.844 and 1.44 nm for phenol.

Since the model fitted on bulk aqueous solution is able to reproduce the excitation wavelength of the same molecule adsorbed at the air-ice interface, the ML model was further developed by using the TDDFT calculations obtained for the molecules in both environments. This combined model was trained on a total of 325 TDDFT calculations for guaiacol and 312 calculations for phenol. The regularization parameter remained at 10^{-8} for both molecules, with fitting and testing carried out using an 5-fold cross validation for guaiacol and 4-fold validation for phenol. Figures S5 and S6 in SI shows that combined model for guaiacol has R^2 of 0.838 (solution) and 0.778 (ice), along with MAE of 1.74 nm (solution) and 2.05 nm (ice), while the combined model for phenol entails R^2 of 0.892 (solution) and 0.863 (ice), with MAE of 1.29 nm (solution) and 1.50 nm (ice).

Results and Discussion

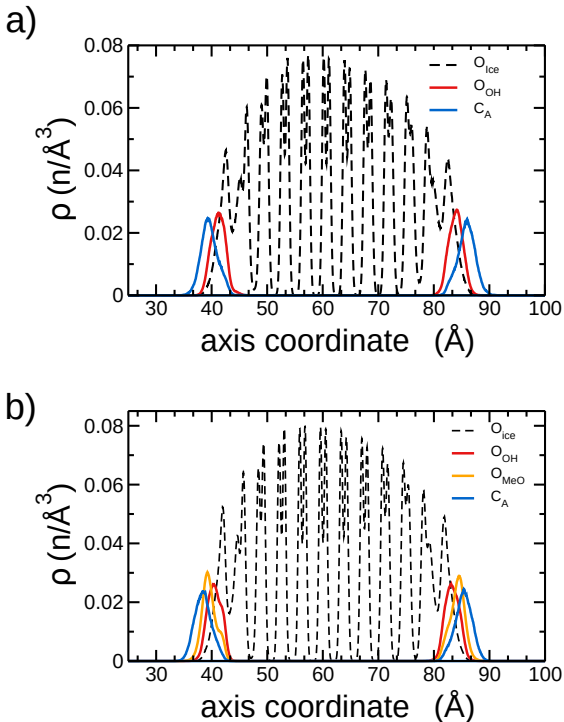## Adsorption of Phenol and Guaiacol at the Air-Ice Interface

a)

b)

Figure 2: Probability distribution for the oxygen atoms in the hydroxy (O_{OH}) and, in the case of guaiacol, the methoxy group (O_{MeO}) and for the carbon opposite to the hydroxy carbon in the aromatic ring (C_A) with respect to the density profile of the ice slabs, obtained from classical MD simulations of phenol (panel a) and guaiacol (panel b) at the air-ice interface.

Experiments on the photodegradation of organic pollutants in snow are often carried out at temperatures between 253 and 263 K.[9,10,13,79] The first step in this study consisted of studying the structure, dynamics and free energy of adsorption of the molecule at the air-ice interface at 263 K, so as to compare to the corresponding properties in aqueous solutions at room temperature. At this temperature the basal plane of ice exhibits a disordered and diffusive QLL about two bi-layers thick.[45,47,80,81] The probability distributions of the positions of the oxygen atom of the hydroxyl group of both phenol and guaiacol over the surface, overlapped to the density profile of the ice slabs (Figure 2), show that both molecules adsorb at the ice surface forming hydrogen bonds between the OH group and water molecules

13

in the outermost layer of the QLL. The three-dimensional solvation structure of the hydroxyl groups (Figure S7ab) shows that OH accepts two hydrogen bonds. As opposed to the case of solvation in bulk water (Figure S7cd), the phenyl ring does not get fully solvated and floats above the surface. Similarly, the CH_3 end of the methoxy group of guaiacol exhibits hydrophobic behavior, but the bridging oxygen interacts with water as it can accept a weak hydrogen bond. The hydrophilic interaction of the polar OH group and the hydrophobic behavior of the phenyl ring determine the orientation of the molecules with respect to the ice surface. The analysis of the angles that define such orientation shows that both molecules lie with the phenyl ring at $\sim 45°$ with respect to the ice surface, with differences in the tilt angle, dictated by the presence of the methoxy group in guaiacol (a detailed description of the angle distributions is provided in the caption of Figure S8 in SI).

Table 1: Calculated solvation free energies in kcal/mol for phenol and guaiacol 300K (ΔG_{300K}) and 263 K (ΔG_{263K}) and adsorption free energy at the air-ice interface at 263 K (ΔG_{ice}). Adsorption free energies at the air-ice interface are not corrected for the change in translational entropy.

Molecule	ΔG_{300K}	ΔG_{263K}	ΔG_{ice}
Phenol	-5.30±0.30	-6.12±0.29	-9.90±0.18
Guaiacol	-5.62±0.46	-8.05±0.65	-13.12±0.30

We compute the adsorption free energy of phenol and guaiacol at the air-ice interface at 263 K, and we compare it to the solvation free energies in supercooled and room temperature water. The details of these FEP calculations are reported in Figures S2 and S3. Adsorption free energies at the air-ice interface turn out lower than the solvation free energies in water at 263 K and even lower when compared to solution at room temperature, (Table 1). The adsorption free energy obtained by FEP is a lower boundary, as it may not fully take into account the removal of translational entropy from gas phase molecules. However, we estimated that such correction amounts to about 2.5 kcal/mol, so that adsorption on ice remains thermodynamically favorable with respect to solvation in LLR at the same temperature. As FEP calculations are carried out in two stages, it is possible to sort out the origin of the

observed free energy differences, distinguishing between the contribution from electrostatic interactions on one hand and short-range repulsion and dispersion forces on the other hand. The difference between solvation in water and adsorption at ice surfaces mostly comes from the latter, and it stems from the cost of creating a cavity for the solute in the bulk liquid, which is higher than the equivalent term for the same molecule adsorbed at the air-ice interface. This result is very important because it shows that it is more thermodynamically favorable for both molecules to adsorb at the air-ice interface than to be solvated in bulk liquid regions.
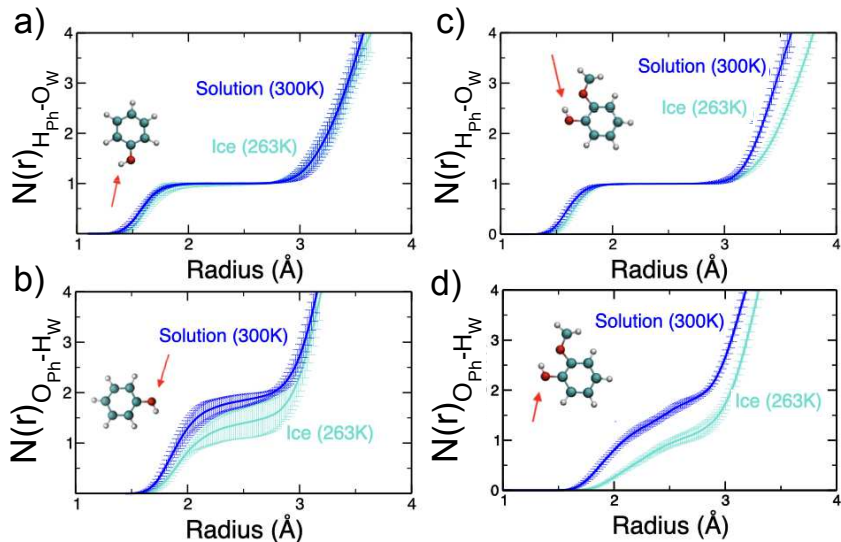

Figure 3: Coordination number comparison showing the variance over 5 trajectories between solution and ice surface obtained from FPMD calculations, focusing on the hydrogen bonding region (2–3 Å) for a) phenol and b) guaiacol. Plots on the left show the bonding pattern of the hydrogen of the hydroxyl group (H_{Ph}) with the water oxygen (O_W), while plots on the right show the bonding pattern of the oxygen of the hydroxyl group (O_{Ph}) with the water hydrogen atoms (H_W), as indicated by the red arrows.

Whereas classical MD simulations provide a coarse picture of the thermodynamics of solvation, FPMD gives more accurate insight into the local structure of the hydration shell of each molecule in either aqueous solution or at the air-ice interface. In particular, we focus on the hydration shell of the hydroxyl group, which forms hydrogen bonds with water.

Figure S9 shows that there are significant differences in the radial distribution functions and their integral between classical and first-principles MD simulations. As shown in Figure 3, the hydroxyl groups donate one hydrogen bond regardless of the solvation environment. Conversely, differences in the accepted hydrogen bonds are more substantial when comparing the two molecules in solution at room temperature or at the air-ice interface at 263 K. For both molecules the hydroxyl group accepts between one and two hydrogen bonds from water. The nearest neighbor hydration structure of the hydroxyl group of phenol does not vary beyond the statistical uncertainty of the FPMD simulations from bulk water to ice (Figure 3c). In contrast, the coordination of the hydroxyl group of guaiacol drops significantly when the molecule is adsorbed on ice rather than in solution (Figure 3d), as a result of the interaction of the methoxy group with the surface. These results show that there are subtle differences in the hydration environment of phenol and guaiacol between solvation in bulk water and at the air-ice interface. In the following section, we assess whether these differences affect the light absorption spectra, making it necessary to consider explicitly the water molecules in the first solvation shells in the computational model.

UV-Visible Absorption Spectra

To compare UV-visible absorption spectra of phenol and guaiacol in aqueous solution and at the air-ice interface, and compute solvatochromic shifts, we need to take into account three main factors: (i) the sampling of the configurational space of the molecules, driven by temperature fluctuations and affected by different solvation environments, (ii) long-range dielectric screening, and (iii) local interactions with the hydration shell.

Previous works suggest that it is necessary to consider the thermal fluctuations of the solvation shell in the shape and width of absorption spectra.[30,31,38–41] These works highlight the importance of calculating absorption spectra for a distribution of configurations that sample the relevant thermodynamic ensemble. To this aim, we have calculated the absorption spectra for a number of statistically independent frames from the FPMD trajectories described

in the previous section. Convergence tests show that within this ensemble approach, the line-shape of the lowest-energy excitation band converges averaging over at least 100 frames (Figure S10). To assess the effect of configurational sampling, we compare the absorption spectra of geometry optimized phenol and guaiacol in aqueous solution, using molecular configurations extracted from the corresponding FPMD runs at 300 K (Figure 4). Both spectra are compared to experimental molar absoptivities measured in solution.[82,83] TDDFT calculations of the UV-visible absorption spectra are carried out for phenol and guaiacol in the SCCS model, which was used also for the geometry optimization of the zero-temperature configurations. In time-dependent density functional perturbation theory absorption spectra are obtained from the resolvent of the TDDFT operator, which is computed through the Lanczos iterative algorithm. The poles of the resolvent, which correspond to the excitation energies of the system, are smoothed by adding a small imaginary part to the frequency; this implies that peaks in the spectra are broadened by Lorentzian functions, with a width σ set by the user. We used $\sigma = 0.027$ eV in all our calculations.

Averaging over a finite temperature ensemble has a major effect on peak positions and amplitudes for both molecules. Temperature fluctuations result in a red shift with respect to the excitation energy of zero-temperature optimized frames (8 nm for guaiacol and 6 nm for phenol) and broadening of the peaks. The lineshapes obtained from the ensemble method resemble fairly closely the experimental measurements and, in the case of phenol, also the peak positions correspond. As expected, we observe a red shift of the guaiacol solution spectrum with respect to experiment. The energy difference between the maximum of the theoretical and the experimental lowest-energy absorption bands of guaiacol in Figure 4 is about 0.15 eV. This discrepancy is not surprising, as TDDFT calculations using semi-local GGA functionals are known to lead to a systematic underestimation of the excitation energies up to ~ 0.4 eV for benzene-like molecules.[84] The temperature induced broadening is by far the most prominent effect on the line shape of the absorption spectrum of aromatic molecules. Experimental estimates of the radiative and non-radiative lifetime of the low

17

energy excited states of benzene, both of the order of 10^{-7} s, suggest that the broadening induced by this additional effect is of the order of 10^{-8} eV, so it can be neglected in our calculations.
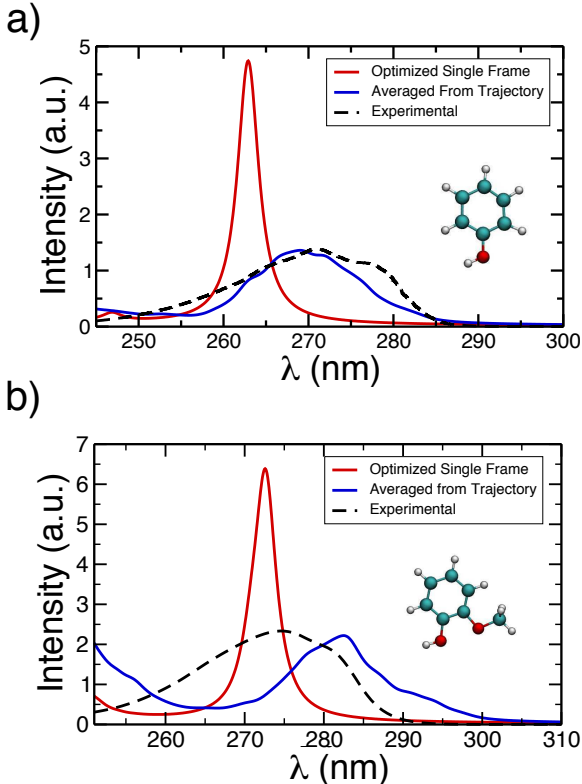
a)

b)

Figure 4: Absorption spectra of a) phenol and b) guaiacol in aqueous solution obtained from a single geometry-optimized configuration (red line) and an ensemble average over the first-principles molecular dynamics trajectory at 300 K (blue line) calculated via TDDFT. The dotted lines on both plots show the measured absorptivities of phenol and guaiacol in solution.[82,83]

The different solvation environments are treated with the standard SCCS model for bulk water, and with the newly implemented extension of the SCCS model for the air-ice interface,[75] in which phenol and guaiacol are embedded at the interface between regions with the dielectric constants of ice and air (vacuum). In Figure S11 we assess the influence of the different implicit solvent models on the absorption spectra of phenol and guaiacol optimized in the respective solvation environment, comparing to the gas-phase spectra. While the differences are substantially smaller than those observed by ensemble averaging, the different

electrostatic embeddings produce changes in both peak positions and oscillator strength. Yet, the sign of these changes is different for the two molecules.

In order to account for the effect of the local solvation environment, we have computed the absorption spectra including a systematically increasing number of explicit water molecules in the hydration shell of both phenol and guaiacol. We have considered molecules at an increasing distance from the oxygen atom of the hydroxyl group, as it is the part of the molecule that interacts more strongly with the solvent. Similar to implicit solvent calculations, about 100 frames are necessary to converge the line shape of spectra. Whereas this hybrid approach accounts for more physics, as it includes both short-range and long-range effects, the number of explicit water molecules needs to be converged to avoid artifacts. Provorse et al.[69] showed that in a mixed implicit/explicit solvent model, a small number of explicit solvent molecules is sufficient to achieve well converged spectra for neutral solutes. In a similar fashion we tested a mixed explicit/implicit solvent framework with both phenol and guaiacol both in bulk water and at the air-ice interface. We observe that for both molecules either in solution or at the air-ice interface the wavelength of the maximum of the lowest energy absorption peak is not sensitive to the number of explicit water molecules in the hybrid model, and with five explicit water molecules it does not vary with respect to that computed with the fully implicit solvation model (Figure S12), indicating that local solvation has only secondary effects on the absorption spectra of these molecules. Hence, hereafter we solely rely on the SCCS implicit solvent model to represent either the aqueous solution or the air-ice interface solvation environment.

Figure 5 shows the theoretical (TDDFT) absorption spectra of guaiacol and phenol at the air-ice interface and in water, each computed by averaging over about 200 statistically independent frames picked from FPMD trajectories at random intervals. The figure also shows the modeled solar photon flux for Summit, Greenland on the summer solstice from the tropospheric ultraviolet and visible (TUV) model.[19] The main result is that our model predicts a solvatochromic (bathochromic) shift of the lowest-energy absorption band for both
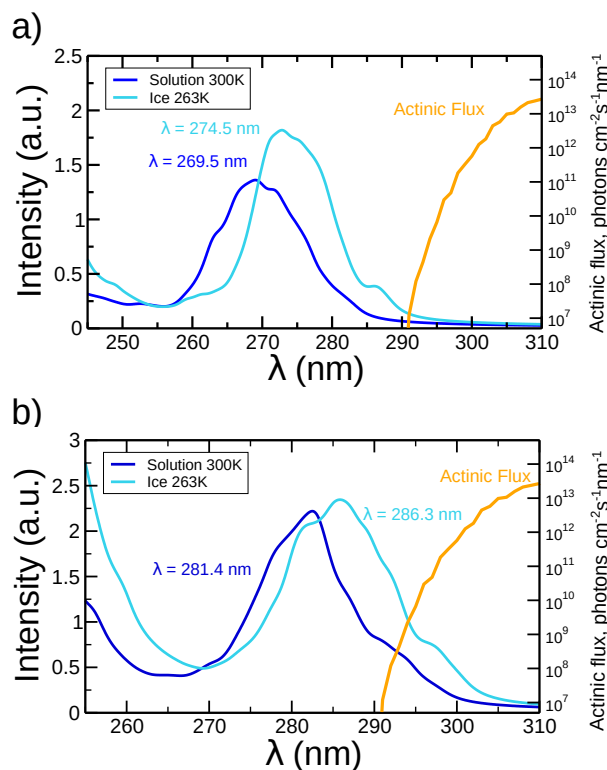
Figure 5: Absorption spectra for a) phenol and b) guaiacol comparing theoretical results obtained in solution and on the ice surface. A shift of approximately 5 nm to the right can be seen for both molecules at the air-ice interface compared to solution. Values represent wavelengths of peak absorbance in the long-wavelength band. The orange line represents the modeled solar spectrum for Summit, Greenland at midday on the summer solstice from the tropospheric ultraviolet and visible (TUV) model.[19]

molecules adsorbed at the air-ice interface, with the maximum of the molar absorptivity peak red-shifted by about 5 nm relative to bulk solution. Although small, such shift modifies significantly the overlap between the molar absorptivity of the molecules and the actinic flux, thus leading to a significant increase of the photon absorption when the two molecules are adsorbed at the air-ice interface. This is especially true for guaiacol, which has the low-energy tail of UV-visible absorption spectrum in a region where the actinic flux increases by several orders of magnitude. Specifically, there is a million-fold increase between 290 and 310 nm,[19,83] which is the wavelength range in which the spectra tail off. Conversely, we are not aware of experiments on phenol photodegradation at the air-ice interface, but, based on the small degree of overlap with the solar flux, our computed spectra suggest that the

photodecay of phenol is very slow both in solution and at the ice surface, but it would be enhanced in the latter.

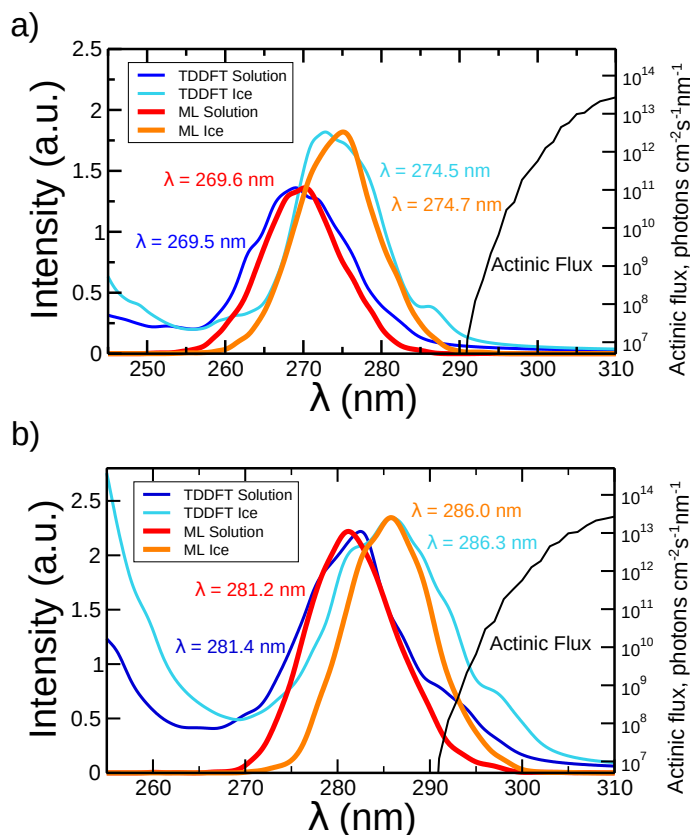## Spectral Line Shapes from Machine Learning

Figure 6: Absorption spectra for (a) phenol and (b) guaiacol in solution and on the ice surface, comparing results obtained from TDDFT calculations (blue/light blue) and through the LASSO regression model (red/orange). The amplitudes of the machine-learning spectra are normalized to those of the TDDFT spectra.

A quantitative estimate of the increased photon absorption of molecules at the air-ice interface can be obtained by computing the convolution of the molar absorptivity with the photon flux, which, however, would require a precise evaluation of the shape of the lowest energy band of the UV-visible absorption spectra. Yet, in our calculations these are affected by the Lorentzian broadening of the TDDFT spectra and by insufficient statistical sampling, which yield slowly decaying tails and features like the accentuated shoulders at 287 nm and

298 nm in the spectra of phenol and guaiacol at the air-ice interface (Figure 5). To remove these artifacts we have designed an efficient statistical learning approach, outlined in the Methods section, which fits the excitation energy on a reduced set of molecular coordinates using the LASSO algorithm, thus allowing us to compute the spectra of thousands of frames at nearly no computational cost. The outcome of this ML model is the lowest energy bands of the spectra, obtained by averaging 5,000 frames from the 50 ps FPMD trajectories, in which each frame contributes to the total spectrum as a Gaussian with 0.014 nm width. The comparison between the UV-visible absorption bands computed by ML and by TDDFT (Figure 6) shows that the ML approach provides smoother line-shapes without requiring a fictitiously large line broadening and eliminates unphysical long-wavelength tails, while peak positions and line width remain similar to those computed by the TDDFT ensemble method. In spite of the solvatochromic shift the spectrum of phenol at the air-ice interface does not overlap with the solar flux, thus we do not expect any enhancement of photodegradation rates due to an increased absorption of photons. In contrast, the solvatochromic shift of guaiacol can produce a substantial enhancement of photon absorption, the extent of which depends on the detailed shape of the actinic flux. For example, we calculated that a 5 nm red shifts increases the rate of sunlight absorption by a factor of 10 under polar summer sunlight conditions.[83]

Furthermore, our approach also sheds light into the physical origin of the observed bathochromic shift. While fitting and testing the ML model, we realized that a model fitted on bulk aqueous solution data is capable of reproducing the lowest excitation energies of the same molecule adsorbed at the air-ice interface, with similar accuracy, i.e., MAE$\lesssim 2$ nm, and without systematic deviations. This means that the molecular configuration is the main parameter that determines the lowest excitation energy, and therefore the position and shape of the longest wavelength absorption band. The screening provided by the solvation environment turns out to be a higher order correction. The weights of the LASSO fitting parameters associated with the coordinates of each atom give the relative influence of such
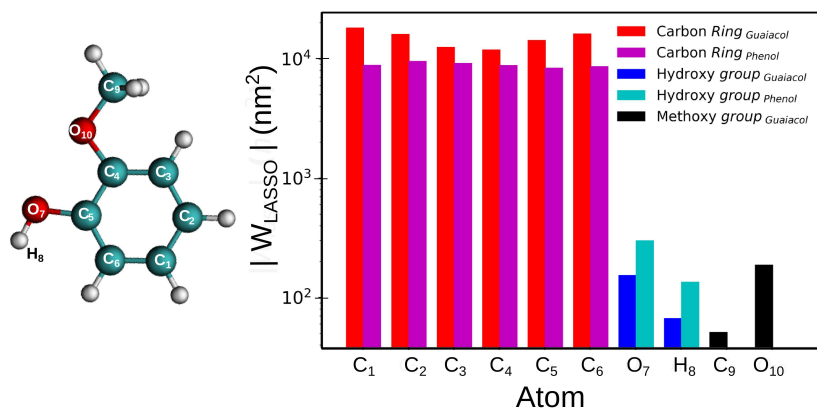
Figure 7: Histogram representing the cumulative weights to the LASSO regression models from coordinates of the atoms of phenol/guaiacol as labeled in the molecular model (left). Atoms are labeled consistently for both molecules. Data from guaiacol adapted from Hullar et al.[83]

coordinates to determine the excitation energy. Figure 7 shows that for both phenol and guaiacol the coordinates of the carbon atoms in the phenyl ring have the largest weights, thus indicating that the observed bathochromic shift is caused by configurational changes in the phenyl ring, dictated by different temperatures and solvation environments. In fact, we observed that for both molecules the distribution of the C-C bond lengths changes at the different conditions, with slightly larger bond lengths when the molecules are adsorbed at the air-ice interface. The centers of the bond length distributions are shifted by 0.011 Å for phenol and 0.012 Å for guaiacol (Figure S13). It is worth noting that for phenol the coordinates of the OH group provide a more significant contribution to the ML model than for guaiacol. To further verify the hypothesis that the solvatochromic shift is essentially dictated by the different configurations of the molecules on ice and in water, we recalculated the absorption spectrum for the ice frames by TDDFT using the bulk water SCCS parameters: the two spectra show minimal differences (Figure S14).

Conclusions

We have developed and validated a multimodel approach for the calculation of absorption spectra of organic molecules at the air-ice interface, that combines classical and quantum simulations and statistical learning. Our approach takes into account the relevant thermodynamic conditions and the effects of the interactions between solvent and solute. For the latter we have extended the SCCS implicit solvent model to represent regions with different dielectric constants.

Classical MD simulations of well equilibrated ice surface models show that the adsorption of phenolic molecules at the air-ice interface is thermodynamically more favorable than the solvation in bulk liquid regions. Moreover they show that these molecules remain confined to the surface of the QLL on ice, and their dynamics are dictated by the hydrogen bonds between the hydroxy group and water.

TDDFT calculations over an ensemble of configurations, extracted from FPMD simulations, show that these solvation conditions determine a solvatochromic red shift of the lowest energy band of the UV-visible absorption spectrum of about 5 nm. The origin of this bathochromic shift for both molecules lies in slight configurational changes of the aromatic ring, dictated by the different solvation environments. This allows us to fit the lowest excitation energies on a simple statistical learning model from which we can calculate statistically converged line shapes for the lowest energy absorption band. These calculations can be exploited to estimate the enhanced photon absorption for molecules adsorbed at the air-ice interface with respect to bulk solutions and, therefore, interpret the origin of enhanced photodegradation rates of organic pollutants in snowpacks.

Our efforts represent the most sophisticated modeling of light absorption spectra for organic molecules at the air-ice interface to date. The resulting red shifts are probably not only significant for phenols, which are important pollutants in Arctic regions,[85–87] but also for a wide range of organic contaminants. While these red shifts increase the photodegration rates and shorten pollutant lifetimes, in order to fully understand photochemistry of molecules at

the air-ice interface, we also need to determine the impact of this environment on quantum yields.[83]

Supporting Information Available

The Supporting Information is available and free of charge. Force field parameterization methods, further classical MD analysis (3D density maps, average positioning, coordination number comparison with AIMD), single frame absorption spectra calculations using different implicit solvation models, absorption spectra convergence plots, changes in absorption spectra max with the addition of explicit waters, machine learning parity plots for the two described methods.

Acknowledgement

References

(1) Barret, M.; Domine, F.; Houdier, S.; Gallet, J.-C.; Weibring, P.; Walega, J.; Fried, A.; Richter, D. Formaldehyde in the Alaskan Arctic Snowpack: Partitioning and Physical Processes Involved in Air-Snow Exchanges. *J. Geophys. Res. Atmos.* **2011**, *116*.

(2) Bartels-Rausch, T.; Jacobi, H.-W.; Kahan, T. F.; Thomas, J. L.; Thomson, E. S.; Abbatt, J. P. D.; Ammann, M.; Blackford, J. R.; Bluhm, H.; Boxe, C. et al. A Review of

Air-Ice Chemical and Physical Interactions (AICI): Liquids, Quasi-Liquids, and Solids in Snow. *Atmos. Chem. Phys.* **2014**, *14*, 1587–1633.

(3) Jacobi, H.-W.; Bales, R. C.; Honrath, R. E.; Peterson, M. C.; Dibb, J. E.; Swanson, A. L.; Albert, M. R. Reactive Trace Gases Measured in the Interstitial Air of Surface Snow at Summit, Greenland. *Atmos. Environ.* **2004**, *38*, 1687 – 1697.

(4) Dibb, J. E.; Arsenault, M. Shouldn't Snowpacks be Sources of Monocarboxylic Acids? *Atmos. Environ.* **2002**, *36*, 2513 – 2522.

(5) Sumner, A.; Shepson, P. Snowpack Production of Formaldehyde and its Effect on the Arctic Troposphere. *Nature* **1999**, *398*, 230–233.

(6) Stathis, A. A.; Hendrickson-Stives, A. K.; Kahan, T. F. Photolysis Kinetics of Toluene, Ethylbenzene, and Xylenes at Ice Surfaces. *J. Phys. Chem. A* **2016**, *120*, 6693–6697.

(7) Kahan, T. F.; Donaldson, D. J. Photolysis of Polycyclic Aromatic Hydrocarbons on Water and Ice Surfaces. *J. Phys. Chem. A* **2007**, *111*, 1277–1285.

(8) Chu, L.; Anastasio, C. Quantum Yields of Hydroxyl Radical and Nitrogen Dioxide from the Photolysis of Nitrate on Ice. *J. Phys. Chem. A* **2003**, *107*, 9594–9602.

(9) Kahan, T. F.; Zhao, R.; Jumaa, K. B.; Donaldson, D. J. Anthracene Photolysis in Aqueous Solution and Ice: Photon Flux Dependence and Comparison of Kinetics in Bulk Ice and at the Air-Ice Interface. *Environ. Sci. Technol.* **2010**, *44*, 1302–1306.

(10) Hullar, T.; Magadia, D.; Anastasio, C. Photodegradation Rate Constants for Anthracene and Pyrene Are Similar in/on Ice and in Aqueous Solution. *Environ. Sci. Technol.* **2018**, *52*, 12225–12234.

(11) Chu, L.; Anastasio, C. Formation of Hydroxyl Radical from the Photolysis of Frozen Hydrogen Peroxide. *J. Phys. Chem. A* **2005**, *109*, 6264–6271.

(12) Chu, L.; Anastasio, C. Temperature and Wavelength Dependence of Nitrite Photolysis in Frozen and Aqueous Solutions. *Environ. Sci. Technol.* **2007**, *41*, 3626–3632.

(13) Ram, K.; Anastasio, C. Photochemistry of Phenanthrene, Pyrene, and Fluoranthene in Ice and Snow. *Atmos. Environ.* **2009**, *43*, 2252 – 2259.

(14) Heger, D.; Jirkovsk, J.; Klán, P. Aggregation of Methylene Blue in Frozen Aqueous Solutions Studied by Absorption Spectroscopy. *J. Phys. Chem. A* **2005**, *109*, 6702–6709.

(15) Kania, R.; Malongwe, J. K.; Nachtigallová, D.; Krausko, J.; Gladich, I.; Roeselová, M.; Heger, D.; Klán, P. Spectroscopic Properties of Benzene at the Air-Ice Interface: A Combined Experimental-Computational Approach. *J. Phys. Chem. A* **2014**, *118*, 7535–7547.

(16) Matykiewiczová, N.; Kurkova, R.; Klanova, J.; Klán, P. Photochemically Induced Nitration and Hydroxylation of Organic Aromatic Compounds in the Presence of Nitrate or Nitrite in Ice. *J. Photochem. Photobiol. A* **2007**, *187*, 24 – 32.

(17) Malongwe, J. K.; Nachtigallová, D.; Corrochano, P.; Klán, P. Spectroscopic Properties of Anisole at the Air-Ice Interface: A Combined Experimental-Computational Approach. *Langmuir* **2016**, *32*, 5755–5764.

(18) Krausko, J.; Malongwe, J. K.; Bičanová, G.; Klán, P.; Nachtigallová, D.; Heger, D. Spectroscopic Properties of Naphthalene on the Surface of Ice Grains Revisited: A Combined Experimental-Computational Approach. *J. Phys. Chem. A* **2015**, *119*, 8565–8578.

(19) Madronich, S.; Flocke, S. In *Environmental Photochemistry*; Boule, P., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 1999; pp 1–26.

(20) Heger, D.; Nachtigallová, D.; Surman, F.; Krausko, J.; Magyarová, B.; Brumovský, M.; Rubes, M.; Gladich, I.; Klán, P. Self-Organization of 1-Methylnaphthalene on the Surface of Artificial Snow Grains: A Combined Experimental-Computational Approach. *J. Phys. Chem. A* **2011**, *115*, 11412–11422.

(21) Vácha, R.; Cwiklik, L.; Řezáč, J.; Hobza, P.; Jungwirth, P.; Valsaraj, K.; Bahr, S.; Kempter, V. Adsorption of Aromatic Hydrocarbons and Ozone at Environmental Aqueous Surfaces. *J. Phys. Chem. A* **2008**, *112*, 4942–4950.

(22) Gopalakrishnan, S.; Jungwirth, P.; Tobias, D. J.; Allen, H. C. Air-Liquid Interfaces of Aqueous Solutions Containing Ammonium and Sulfate: Spectroscopic and Molecular Dynamics Studies. *J. Phys. Chem. B* **2005**, *109*, 8861–8872.

(23) Jungwirth, P.; Tobias, D. J. Specific Ion Effects at the Air/Water Interface. *Chem. Rev.* **2006**, *106*, 1259–1281.

(24) Epstein, S. A.; Shemesh, D.; Tran, V. T.; Nizkorodov, S. A.; Gerber, R. B. Absorption Spectra and Photolysis of Methyl Peroxide in Liquid and Frozen Water. *J. Phys. Chem. A* **2012**, *116*, 6068–6077.

(25) Dreuw, A.; Head-Gordon, M. Single-Reference Ab Initio Methods for the Calculation of Excited States of Large Molecules. *Chem. Rev.* **2005**, *105*, 4009–4037.

(26) Blase, X.; Duchemin, I.; Jacquemin, D. The Bethe–Salpeter Equation in Chemistry: Relations With TD-DFT, Applications and Challenges. *Chem. Soc. Rev.* **2018**, *47*, 1022–1043.

(27) Casida, M. E. Time-Dependent Density-Functional Theory for Molecules and Molecular Solids. *J. Mol. Struct.* **2009**, *914*, 3–18.

(28) Gross, E. K. U.; Kohn, W. In *Density Functional Theory of Many-Fermion Systems*;

Lowdin, P.-O., Ed.; Advances in Quantum Chemistry; Academic Press: Boston, 1990; Vol. 21; pp 255 – 291.

(29) Burke, K.; Werschnik, J.; Gross, E. K. U. Time-Dependent Density-Functional Theory: Past, Present, and Future. *J. Chem. Phys.* **2005**, *123*, 062206–10.

(30) Timrov, I.; Micciarelli, M.; Rosa, M.; Calzolari, A.; Baroni, S. Multimodel Approach to the Optical Properties of Molecular Dyes in Solution. *J. Chem. Theory Comput.* **2016**, *12*, 4423–4429.

(31) Zuehlsdorff, T. J.; Montoya-Castillo, A.; Napoli, J. A.; Markland, T. E.; Isborn, C. M. Optical Spectra in the Condensed Phase: Capturing Anharmonic and Vibronic Features Using Dynamic and Static Approaches. *J. Chem. Phys.* **2019**, 074111.

(32) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999–3094.

(33) Mennucci, B.; Cappelli, C.; Guido, C. A.; Cammi, R.; Tomasi, J. Structures and Properties of Electronically Excited Chromophores in Solution from the Polarizable Continuum Model Coupled to the Time-Dependent Density Functional Theory. *J. Phys. Chem. A* **2009**, *113*, 3009–3020.

(34) Isborn, C. M.; Götz, A. W.; Clark, M. A.; Walker, R. C.; Martínez, T. J. Electronic Absorption Spectra from MM and ab InitioQM/MM Molecular Dynamics: Environmental Effects on the Absorption Spectrum of Photoactive Yellow Protein. *J. Chem. Theory Comput.* **2012**, *8*, 5092–5106.

(35) Milanese, J. M.; Provorse, M. R.; Alameda Jr., E.; Isborn, C. M. Convergence of Computed Aqueous Absorption Spectra with Explicit Quantum Mechanical Solvent. *J. Chem. Theory Comput.* **2017**, *13*, 2159–2171.

(36) Provorse Long, M. R.; Isborn, C. M. Combining Explicit Quantum Solvent with a Polarizable Continuum Model. *J. Phys. Chem. B* **2017**, *121*, 10105–10117.

(37) Loco, D.; Gelfand, N.; Jurinovich, S.; Protti, S.; Mezzetti, A.; Mennucci, B. Polarizable QM/Classical Approaches for the Modeling of Solvation Effects on UV-Vis and Fluorescence Spectra: An Integrated Strategy. *J. Phys. Chem. A* **2018**, *122*, 390–397.

(38) Barone, V.; Bloino, J.; Monti, S.; Pedone, A.; Prampolini, G. Theoretical Multilevel Approach for Studying the Photophysical Properties of Organic Dyes in Solution. *Phys. Chem. Chem. Phys.* **2010**, *12*, 10550–10561.

(39) Malcıoğlu, O. B.; Calzolari, A.; Gebauer, R.; Varsano, D.; Baroni, S. Dielectric and Thermal Effects on the Optical Properties of Natural Dyes: A Case Study on Solvated Cyanin. *J. Am. Chem. Soc.* **2011**, *133*, 15425–15433.

(40) De Mitri, N.; Monti, S.; Prampolini, G.; Barone, V. Absorption and Emission Spectra of a Flexible Dye in Solution: a Computational Time-Dependent Approach. *J. Chem. Theory Comput.* **2013**, *9*, 4507–4516.

(41) Cacelli, I.; Cinacchi, G.; Prampolini, G.; Tani, A. Computer Simulation of Solid and Liquid Benzene with an Atomistic Interaction Potential Derived from ab Initio Calculations. *J. Am. Chem. Soc.* **2004**, *126*, 14278–14286.

(42) Andreussi, O.; Dabo, I.; Marzari, N. Revised Self-Consistent Continuum Solvation in Electronic-Structure Calculations. *J. Chem. Phys.* **2012**, *136*, 064102.

(43) Andreussi, O.; Marzari, N. Electrostatics of Solvated Systems in Periodic Boundary Conditions. *Phys. Rev. B* **2014**, *90*, 245101.

(44) Ge, X.; Timrov, I.; Binnie, S.; Biancardi, A.; Calzolari, A.; Baroni, S. Accurate and Inexpensive Prediction of the Color Optical Properties of Anthocyanins in Solution. *J. Phys. Chem. A* **2015**, *119*, 3816–3822.

(45) Sánchez, M. A.; Kling, T.; Ishiyama, T.; van Zadel, M.-J.; Bisson, P. J.; Mezger, M.; Jochum, M. N.; Cyran, J. D.; Smith, W. J.; Bakker, H. J. et al. Experimental and Theoretical Evidence for Bilayer-by-Bilayer Surface Melting of Crystalline Ice. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 227–232.

(46) Benet, J.; Llombart, P.; Sanz, E.; MacDowell, L. G. Premelting-Induced Smoothening of the Ice-Vapor Interface. *Phys. Rev. Lett.* **2016**, *117*, 096101.

(47) Kling, T.; Kling, F.; Donadio, D. Structure and Dynamics of the Quasi-Liquid Layer at the Surface of Ice from Molecular Simulations. *J. Phys. Chem. C* **2018**, *122*, 24780–24787.

(48) Tibshirani, R. Regression Shrinkage and Selection Via the Lasso: a Retrospective. *J. Royal Stat. Soc. B* **2011**, *73*, 273–282.

(49) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling Through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(50) Hockney, R. W.; Eastwood, J. W. *Computer Simulation using Particles*; IOP Publishing Ltd.: Bristol, England, 1988.

(51) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comp. Phys.* **1995**, *117*, 1–19.

(52) Abascal, J.; Sanz, E.; Garcia Fernandez, R.; Vega, C. A Potential Model for the Study of Ices and Amorphous Water: TIP4P/Ice. *J. Chem. Phys.* **2005**, *122*, 234511.

(53) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(54) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *Journal of Computational Chemistry* **2000**, *21*, 132–146.

(55) Zwanzig, R. W.; Kirkwood, J. G.; Oppenheim, I.; Alder, B. J. Statistical Mechanical Theory of Transport Processes. VII. The Coefficient of Thermal Conductivity of Monatomic Liquids. *J. Chem. Phys.* **1954**, *22*, 783–790.

(56) Zhang, J.; Tuguldur, B.; van der Spoel, D. Force Field Benchmark of Organic Liquids. 2. Gibbs Energy of solvation. *J. Chem. Inf. Model.* **2015**, *55*, 1192–1201.

(57) Buch, V.; Sandler, P.; Sadlej, J. Simulations of H2O Solid, Liquid, and Clusters, with an Emphasis on Ferroelectric Ordering Transition in Hexagonal Ice. *J. Phys. Chem. B* **1998**, *102*, 8641–8653.

(58) Vandevondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and Accurate Density Functional Calculations Using a Mixed Gaussian and Plane Waves Approach. *Comput. Phys. Commun.* **2005**, *167*, 103–128.

(59) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. CP2K: Atomistic Simulations of Condensed Matter Systems. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 15–25.

(60) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(61) Feller, D. The Role of Databases in Support of Computational Chemistry Calculations. *J. Comput. Chem.* **1996**, *17*.

(62) Goedecker, S.; Teter, M.; Hutter, J. Separable Dual-Space Gaussian Pseudopotentials. *Phys. Rev. B* **1996**, *54*, 1703–1710.

(63) Rocca, D.; Gebauer, R.; Saad, Y.; Baroni, S. Turbo Charging Time-Dependent Density-Functional Theory with Lanczos Chains. *J. Chem. Phys.* **2008**, *128*, 154105.

(64) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I. et al. QUANTUM ESPRESSO: a Modular

and Open-Source Software Project for Quantum Simulations of Materials. *J. Phys. Condens. Matter* **2009**, *21*, 395502.

(65) Malcıoğlu, O. B.; Gebauer, R.; Rocca, D.; Baroni, S. TurboTDDFT - A Code for the Simulation of Molecular Spectra Using the Liouville-Lanczos Approach to Time-Dependent Density-Functional Perturbation Theory. *Comput. Phys. Commun.* **2011**, *182*, 1744 – 1754.

(66) Vanderbilt, D. Soft Self-Consistent Pseudopotentials in a Generalized Eigenvalue Formalism. *Phys. Rev. B* **1990**, *41*, 7892–7895.

(67) Martyna, G. J.; Tuckerman, M. E. A Reciprocal Space Based Method for Treating Long Range Interactions in Ab Initio and Force-Field-Based Calculations in Clusters. *J. Chem. Phys.* **1999**, *110*, 2810–2821.

(68) Zuehlsdorff, T. J.; Haynes, P. D.; Payne, M. C.; Hine, N. D. M. Predicting Solvatochromic Shifts and Colours of a Solvated Organic Dye: The Example of Nile Red. *J. Chem. Phys.* **2017**, *146*, 124504.

(69) Provorse, M. R.; Peev, T.; Xiong, C.; Isborn, C. M. Convergence of Excitation Energies in Mixed Quantum and Classical Solvent: Comparison of Continuum and Point Charge Models. *J. Phys. Chem. B* **2016**, *120*, 12148–12159.

(70) Fattebert, J.-L.; Gygi, F. Density Functional Theory for Efficient Ab Initio Molecular Dynamics Simulations in Solution. *J. Comput. Chem* **2002**, *23*, 662–666.

(71) Fattebert, J.-L.; Gygi, F. First-Principles Molecular Dynamics Simulations in a Continuum Solvent. *Int. J. Quantum Chem.* **2003**, *93*, 139.

(72) Dupont, C.; Andreussi, O.; Marzari, N. Self-Consistent Continuum Solvation (SCCS): The Case of Charged Systems. *J. Chem. Phys.* **2013**, *139*, 214110.

(73) Scherlis, D. A.; Fattebert, J. L.; Gygi, F.; Cococcioni, M.; Marzari, N. A Unified Electrostatic and Cavitation Model for First-Principles Molecular Dynamics in Solution. *J. Chem. Phys.* **2006**, *124*, 074103.

(74) Dziedzic, J.; Helal, H. H.; Skylaris, C. K.; Mostofi, A. A.; Payne, M. C. Minimal Parameter Implicit Solvent Model for Ab Initio Electronic-Structure Calculations. *Europhys. Lett.* **2011**, *95*, 43001.

(75) Andreussi, O.; Dabo, I.; Marzari, N. Revised Self-Consistent Continuum Solvation in Electronic-Structure Calculations. *J. Chem. Phys.* **2012**, *136*, 064102.

(76) Giannozzi, P.; Andreussi, O.; Brumme, T.; Bunau, O.; Nardelli, M. B.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Cococcioni, M. et al. Advanced Capabilities for Materials Modelling with QUANTUM ESPRESSO. *J. Phys. Condens. Matter* **2017**, *29*, 465901.

(77) Owen, B. B.; Miller, R. C.; Milner, C. E.; Cogan, H. L. The Dielectric Constant of Water as a Function of Temperature and Pressure. *J. Phys. Chem.* **1961**, *65*, 2065–2070.

(78) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

(79) Phillips, G. J.; Simpson, W. R. Verification of Snowpack Radiation Transfer Models Using Actinometry. *J. Geophys. Res. Atmos.* **2005**, *110*.

(80) Llombart, P.; Noya, E. G.; MacDowell, L. G. Surface Phase Transitions and Crystal Habits of Ice in the Atmosphere. *Sci. Adv.* **2020**, *6*, eaay9322.

(81) Llombart, P.; Noya, E. G.; Sibley, D. N.; Archer, A. J.; MacDowell, L. G. Rounded Layering Transitions on the Surface of Ice. *Phys. Rev. Lett.* **2020**, *124*, 065702.

(82) Smith, J. D.; Kinney, H.; Anastasio, C. Phenolic Carbonyls Undergo Rapid Aqueous Photodegradation to Form Low-Volatility, Light-Absorbing Products. *Atmos. Environ.* **2016**, *126*, 36 – 44.

(83) Hullar, T.; Bononi, F. C.; Chen, Z.; Magadia, D.; Palmer, O.; Tran, T.; Rocca, D.; Andreussi, O.; Donadio, D.; Anastasio, C. Photodecay of Guaiacol is Faster in Ice, and Even More Rapid on Ice, Than in Aqueous Solution. *Environ. Sci.: Processes Impacts* **2020**, *22*, 1666–1677.

(84) Miura, M.; Aoki, Y.; Champagne, B. Assessment of Time-Dependent Density-Functional Schemes for Computing the Oscillator Strengths of Benzene, Phenol, Aniline, and Fluorobenzene. *J. Chem. Phys.* **2007**, *127*, 084103–17.

(85) Dubowski, Y.; Hoffmann, M. R. Photochemical Transformations in Ice: Implications for the Fate of Chemical Species. *Geophys. Res. Lett.* **2000**, *27*, 3321–3324.

(86) Pokhrel, A.; Kawamura, K.; Kunwar, B.; Ono, K.; Tsushima, A.; Seki, O.; Matoba, S.; Shiraiwa, T. Ice Core Records of Levoglucosan and Dehydroabietic and Vanillic Acids from Aurora Peak in Alaska Since the 1660s: a Proxy Signal of Biomass-Burning Activities in the North Pacific Rim. *Atmos. Chem. Phys.* **2020**, *20*, 597–612.

(87) Shi, G.; Wang, X.-C.; Li, Y.; Trengove, R.; Hu, Z.; Mi, M.; Li, X.; Yu, J.; Hunter, B.; He, T. Organic Tracers from Biomass Burning in Snow from the Coast to the Ice Sheet Summit of East Antarctica. *Atmos. Environ.* **2019**, *201*, 231 – 241.

(88) Towns, J.; Cockerill, T.; Dahan, M.; Foster, I.; Gaither, K.; Grimshaw, A.; Hazlewood, V.; Lathrop, S.; Lifka, D.; Peterson, G. et al. XSEDE: Accelerating Scientific Discovery. *Comput Sci. Eng.* **2014**, *16*, 62–74.

TOC Graphic