Dissecting the Molecular Structure of the Air/Water Interface from Quantum Simulations of the Sum-Frequency Generation Spectrum

Gregory R. Medders and Francesco Paesani[*]

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093

E-mail: fpaesani@ucsd.edu

1 Classical and quantum infrared spectra of liquid water

In Figure S1, the infrared (IR) spectra of liquid water calculated from classical and quantum many-body molecular dynamics (MB-MD) simulations with the many-body MB-pol potential energy[1–3] and MB-μ dipole moment[4,5] surfaces are shown. Both IR spectra were calculated including one-body and N-body (labeled as "1B + NB" in the main text), but not explicit short-range two-body, induced contributions to the total dipole moment. The quantum MB-MD simulations were carried out within the centroid molecular dynamics (CMD) formalism,[6–12] with simulations performed in the normal-mode representation using the adiabatic separation scheme of Ref. 9. Decoupling between the dynamics of the centroid and non-zero frequency normal modes was enforced with an adiabaticity parameter $\gamma = 0.1$,[13,14] and a timestep of 0.02 fs was found to be sufficient for energy conservation. The classical IR spectrum was obtained by averaging over 30 trajectories of 100 ps each, while the CMD IR spectrum was obtained by averaging over 20 trajectories of 27 ps each. Due to the neglect of nuclear quantum effects (e.g., vibrational zero point energy) in the classical MB-MD simulations, the IR spectrum in the OH stretching region is artificially blue shifted by $\sim$175 cm^{-1} relative to the quantum results.

2 Approximations to the time-correlation function

Figures 2 – 4 of the main text show the phase-sensitive vibrational sum-frequency generation (PS-vSFG) spectrum of the air/water interface calculated according to Eq. 1 from classical MB-MD simulations using the full time-correlation function (FCF) of the system dipole moment and polarizability. While most accurate, this approach requires extensive sampling ($\sim$20 ns) of the underlying molecular configurations to obtain converged results. Since the OH stretching region of the PS-vSFG spectrum primarily probes the intramolecular
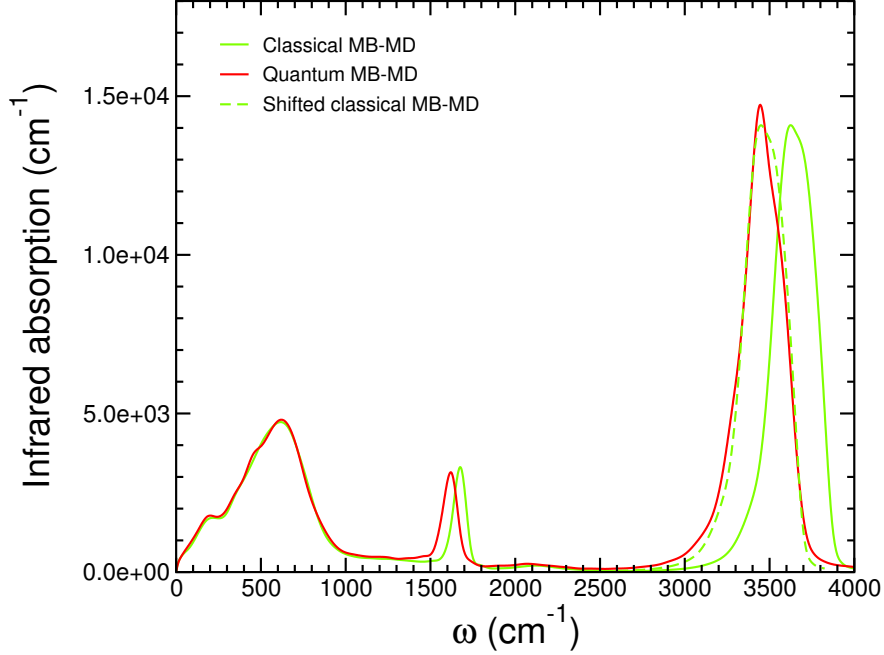
Figure S1: IR spectra of liquid water at ambient conditions (T = 298.15 K) calculated from classical MB-MD (green) and quantum MB-MD (red) simulations with the many-body MB-pol potential energy and MB-μ dipole moment surfaces, using the "1B + NB" approximation. The quantum MB-MD simulations were carried out within the CMD formalism. Also shown as a dashed green trace is the classical MB-MD OH vibrational band shifted by -175 cm^{-1} to approximately account for nuclear quantum effects.

vibrations, it was shown that the time-correlation function of Eq. 1 can be rewritten in a way that excludes correlations between molecules that are not nearby.[15] Neglecting these non-local correlations, which make a small contribution but require extensive sampling to converge, represents an effective strategy to minimizing the simulation time required to calculate converged PS-vSFG spectra.

Following Ref. 15, two approximations to the full time-correlation function are examined in Figure S2: 1) an autocorrelation function (ACF) that includes only contributions from the time dependence of the dipole moment and polarizability on the same molecules, and 2) a truncated cross-correlation function (TCF) that includes the ACF contributions as well as correlations between the dipole moment and polarizability of neighboring molecules lying within a predefined cutoff distance. The use of the ACF and TCF approximations reduces

the simulation time required to obtain converged PS-vSFG spectra to approximately a few hundred picoseconds and a nanosecond, respectively. For the calculations presented in Figure S2, the TCF PS-vSFG spectrum was obtained including only contributions from pairs of molecules whose oxygen atoms were separated by less than 3.7 Å. This cutoff distance ensures that all molecules within the first solvation shell contribute to the TCF. It is important to note that both the ACF and TCF results rely on partitioning the total dipole moment and polarizability of the system into molecular contributions. Due to the similarity between the full MB-MD and "1B + NB" PS-vSFG spectra (shown as red and yellow traces in Figure 3 of the main text, respectively), only the "1B + NB" approximation in considered in the comparison presented in Figure S2 since it enables a natural definition of the dipole moment and polarizability of each water molecule.

The ACF approximation, including only the correlation of each molecule with itself, consistently underestimates the signal in the low-frequency portion of the spectrum where the vibrations are most strongly affected by intermolecular interactions. When the correlation of the dipole moment of one molecule with the polarizability of all other neighboring molecules is included within the TCF approximation, the free OH peak decreases in intensity while the hydrogen-bonded band between 3,100 cm^{-1} and 3,600 cm^{-1} becomes broader. In the FCF PS-vSFG spectrum obtained when all correlations are taken into account the hydrogen-bonded band appears to be slightly shifted towards lower frequencies as well as to lack some (negative) intensity on the blue side portion (i.e., more weakly hydrogen-bonded configurations). The comparison with the full MB-MD spectrum shown in Figure 3 of the main text suggests that the latter feature is likely associated with the "1B + NB" approximation being not capable of fully describing many-body electrostatic effects. Figure S2 shows that the TCF results represent a reasonable approximation to the FCF PS-vSFG spectrum requiring roughly 1/20th of the corresponding simulation time. Since, due to the associated computational cost, the calculation of the FCF from MB-MD simulations within the CMD formalism is currently not feasible, the TCF approximation effectively enables CMD calcula-
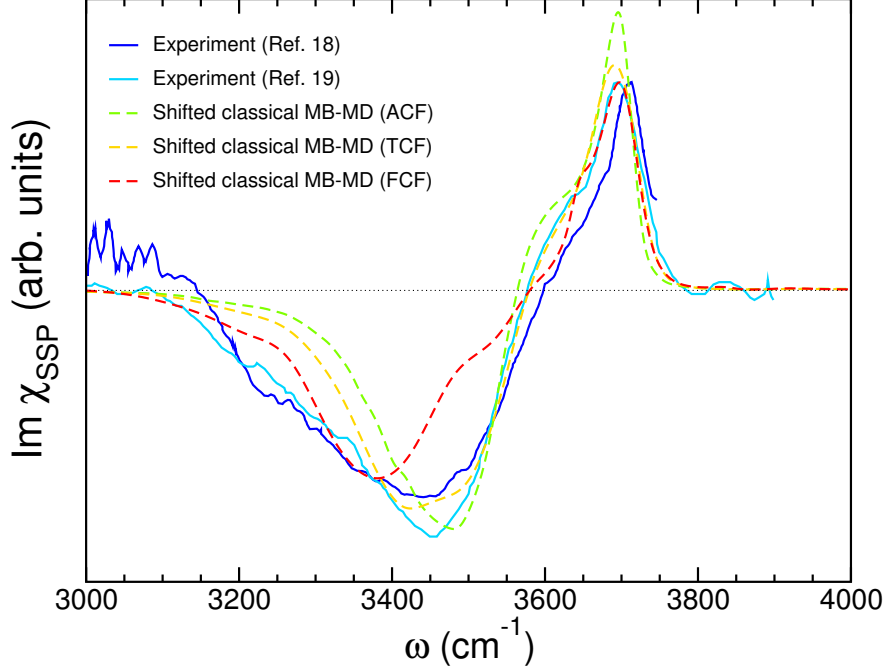
Figure S2: Comparison between the PS-vSFG spectra of the air/water interface calculated from classical MB-MD simulations with the many-body MB-pol water potential using different approximations to the time-correlation function (Eq. 1 in the main text). ACF: auto-correlation function approximation, TCF: cross-correlation function approximation, FCF: full correlation function. All calculations were carried out using the "1B + NB" approximation to the total dipole moment and polarizability (see text for details). The experimental spectra from Refs. 16 and 17 are shown in dark and light blue, respectively. As in Figures 2 – 4 of the main text, all classical spectra are shifted by -175 cm^{-1} to approximately account for nuclear quantum effects.

tions of quantum PS-vSFG spectra. In the present study, the quantum PS-vSFG spectrum of the air/water interface shown in Figure 1 of the main text was calculated from CMD simulations within the TCF approximation described above. Although a cutoff distance of 3.7 Å was used in the CMD calculations, the results were found to be effectively insensitive to the choice of the cutoff, which is consistent with the local nature of the PS-vSFG signal for frequencies above 3,000 cm^{-1}.

3 Analysis of hydrogen bonding at the air/water interface

The analysis of the hydrogen-bonding structure at the air/water interface was carried out using the definition of the instantaneous interface introduced in Ref. 18. Specifically, the differences in the classical MB-MD PS-vSFG spectra calculated with and without including explicit short-range three-body contributions (Figure 4 in the main text) can be rationalized in terms of different of hydrogen-bond topologies as defined in Ref. 19. In panel a) of Figure S3, the water density relative to the bulk density is shown as a function of the distance from the instantaneous interface (located at $z = 0$ Å). While the density of liquid water at ambient conditions predicted by classical MB-MD simulations with the MB-pol water potential without explicit short-range three-body contributions is significantly higher (1.08 g/cm^3) than the value obtained when all many-body effects are taken into account (1.004 g/cm^3), the density profiles with respect to the instantaneous interface are remarkably similar in the two cases. Further examination of the local hydrogen-bonding motifs as a function of the distance from the interface, however, reveals that classical MB-MD simulations without explicit three-body contributions predict a larger fraction of tetrahedrally coordinated water molecules (labeled as "4d" molecules)[19] than classical MB-MD simulations with the full MB-pol potential, which instead favor configurations with molecules both donating and accepting a single hydrogen bond ("2s" molecules)[19] near the interface as reported in Table S1.

Table S1: Fractions of water molecules within 3.2 Å of the instantaneous interface in "2s" and "4d" hydrogen-bonding configurations predicted by classical MB-MD simulations with and without including explicit three-body contributions (no 3B).

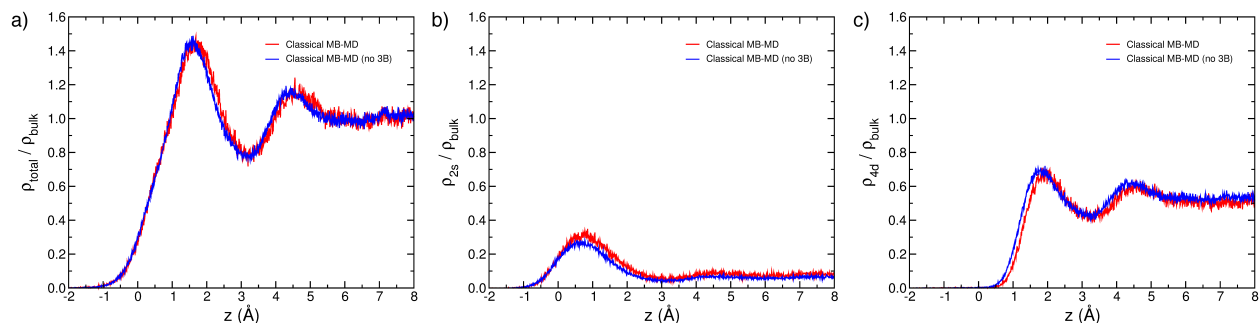|	MB-MD	MB-MD (no 3B)	Δ
2s	19.6%	16.7%	-2.9%
4d	34.8%	37.8%	+3.0%

Figure S3: Water density relative to the bulk density as a function of the distance from the instantaneous interface located at $z = 0$ Å. The results from classical MB-MD simulations with the MB-pol water potentials including all many-body effects (red) and without including explicit short-range three-body contributions (blue) are shown. Panel a) shows the total density of the water molecules, regardless of their hydrogen bonding environment. Panels b) and c) show the density of molecules in "2"s and "4d" hydrogen-bonding environments, respectively.

References

(1) Babin, V.; Leforestier, C.; Paesani, F. *J. Chem. Theory Comput.* **2013**, *9*, 5395–5403.

(2) Babin, V.; Medders, G. R.; Paesani, F. *J. Chem. Theory Comput.* **2014**, *10*, 1599–1607.

(3) Medders, G. R.; Babin, V.; Paesani, F. *J. Chem. Theory Comput.* **2014**, *10*, 2906–2910.

(4) Medders, G. R.; Paesani, F. *J. Chem. Theory Comput.* **2013**, *9*, 4844–4852.

(5) Medders, G. R.; Paesani, F. *J. Chem. Theory Comput.* **2015**, *11*, 1145–1154.

(6) Cao, J.; Voth, G. A. *J. Chem. Phys.* **1994**, *100*, 5093–5105.

(7) Cao, J.; Voth, G. A. *J. Chem. Phys.* **1994**, *100*, 5106–5117.

(8) Cao, J.; Voth, G. A. *J. Chem. Phys.* **1994**, *101*, 6157–6167.

(9) Cao, J.; Voth, G. A. *J. Chem. Phys.* **1994**, *101*, 6168–6183.

(10) Cao, J.; Voth, G. A. *J. Chem. Phys.* **1994**, *101*, 6184–6192.

(11) Jang, S.; Voth, G. A. *J. Chem. Phys.* **1999**, *111*, 2357–2370.

(12) Jang, S.; Voth, G. A. *J. Chem. Phys.* **1999**, *111*, 2371–2384.

(13) Hone, T. D.; Rossky, P. J.; Voth, G. A. *J. Chem. Phys.* **2006**, *124*, 154103.

(14) Rossi, M.; Liu, H.; Paesani, F.; Bowman, J.; Ceriotti, M. *J. Chem. Phys.* **2014**, *141*, 181101.

(15) Nagata, Y.; Mukamel, M. *J. Am. Chem. Soc.* **2010**, *132*, 6434–6442.

(16) Nihonyanagi, S.; Ishiyama, T.; Lee, T.; Yamaguchi, S.; Bonn, M.; Morita, A.; Tahara, T. *J. Am. Chem. Soc.* **2011**, *133*, 16875–16880.

(17) Nihonyanagi, S.; Kusaka, R.; Inoue, K.-i.; Adhikari, A.; Yamaguchi, S.; Tahara, T. *J. Chem. Phys.* **2015**, *143*, 124707.

(18) Willard, A. P.; Chandler, D. *J. Phys. Chem. B* **2010**, *114*, 1954–1958.

(19) Auer, B.; Kumar, R.; Schmidt, J. R.; Skinner, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 14215–14220.