



HAL
open science

From solvation to chemistry of aqueous surfaces : a combined DFT-MD simulations and SFG spectroscopy approach

Wanlin Chen

► **To cite this version:**

Wanlin Chen. From solvation to chemistry of aqueous surfaces : a combined DFT-MD simulations and SFG spectroscopy approach. Theoretical and/or physical chemistry. Université Paris-Saclay, 2023. English. NNT : 2023UPASF029 . tel-04188013

HAL Id: tel-04188013

<https://theses.hal.science/tel-04188013v1>

Submitted on 25 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

From solvation to chemistry of aqueous surfaces:
a combined DFT-MD simulations and SFG
spectroscopy approach.

*De la solvation à la chimie des surfaces aqueuses: une approche
combinée simulations DFT-MD et spectroscopie SFG*

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 571 : sciences chimiques : molécules, matériaux,
instrumentation et biosystèmes (2MIB)

Spécialité de doctorat: chimie

Graduate School : Chimie. Référent : Université d'Évry Val d'Essonne

Thèse préparée dans la unité de recherche **Laboratoire Analyse et Modélisation
pour la Biologie et l'Environnement (LAMBE) (Université Paris-Saclay, Univ
Evry, CNRS)**, sous la direction de **Marie-Pierre GAIGEOT**, Professeure.

Thèse présentée et soutenue à Evry,
le 21 Avril 2023, par

Wanlin Chen

Composition du jury

Laurent Joly Professeur, Université Lyon 1, France	Président
Sandra Luber Professeure, University of Zurich, Switzerland	Rapporteuse
Michael von Domaros Professeur, Philipps-Universität Marburg, Germany	Rapporteur
Christophe Humbert DR-CNRS, Université Paris-Saclay, France	Examineur
Weitao Liu Professeure, Fudan University, China	Examinatrice
Poul Petersen Professeur, Ruhr University Bochum, Germany	Examineur
Konstantin Smirnov CR-CNRS, Université de Lille, France	Examineur
Simone Pezzotti Post-Doctorate, Ruhr University Bochum, Germany	Invité

Contents

I	Introduction	1
II	Theoretical Methods	13
1	Molecular Dynamics	15
1.1	Molecular dynamics	15
1.1.1	Integration algorithms of the equations of motion	16
1.1.2	Temperature and pressure control	17
1.1.3	Initial conditions and choice of time step	19
1.1.4	Periodic boundary conditions	19
1.2	Born-Oppenheimer molecular dynamics (BO-MD) in the Density Functional Theory (DFT) framework	20
1.3	Density Functional Theory — From wave functions to electron density	21
1.3.1	Exchange-correlation functionals	24
1.3.2	Dispersion corrections	25
1.4	The CP2K software and the computational set-up chosen for the DFT-MD simulations in this thesis	27
1.5	Classical MD simulations	29
1.5.1	Intramolecular interactions	30
1.5.2	Intermolecular interactions	31
1.5.3	Force fields of different classes	34
1.5.4	The GROMACS package for CMD/FF-MD and the force fields chosen for the studied systems in this thesis	35
1.5.5	Details of FF-MD simulations of the PEG/water interfaces	36
1.6	Properties calculated based on the MD simulations	38
2	Nonlinear optical vibrational spectroscopy	41
2.1	Generalities on Sum Frequency Generation (SFG) spectroscopy	42
2.2	Computation of SFG vibrational spectra: DFT-MD vs FF-MD	44
2.3	Our approach to calculate SFG spectra from DFT-MD simulations	47
2.4	SFG spectroscopy of charged-interfaces: Deconvolution of $\chi_{BIL}^{(2)}(\omega)$ and $\chi_{DL}^{(2)}(\omega)$	51
2.4.1	Origins of $\chi_{BIL}^{(2)}(\omega)$ and $\chi_{DL}^{(2)}(\omega)$ contributions	51

2.4.2	Experimental approaches for deconvolving $\chi_{BIL}^{(2)}(\omega)$ and $\chi_{DL}^{(2)}(\omega)$ contributions	53
2.4.3	Theoretical approach for deconvolving $\chi_{BIL}^{(2)}(\omega)$ and $\chi_{DL}^{(2)}(\omega)$ contributions	55
2.5	Extraction of the surface potential from non linear optical techniques and vice versa	57
2.5.1	Finding the surface potential from extracted $\chi_{DL}^{(2)}(\omega)$ contribution to the experimental SFG spectrum	57
2.5.2	Reproducing $\chi_{DL}^{(2)}$ in Second Harmonic Generation (SHG) by calculating the surface potential through the modified Gouy-Chapman model	58
2.6	From O-H stretching of water to Si-O stretching at the silica surface	60
III Results		63
3	A metric for molecular hydrophobicity combining structural and spectroscopic descriptors	65
3.1	Development of a new descriptor for quantifying microscopic hydrophobicity and hydrophilicity: the H/V descriptor	66
3.1.1	Structural interpretation of DFT-MD calculated SFG spectra of water at the interface with the SAMs	70
3.1.2	Spectroscopic and structural descriptors of microscopic hydrophilicity/hydrophobicity: the H/V descriptor	73
3.1.3	How DFT-MD helps to understand experimental SFG spectra: extraction of the surface potential	75
3.1.4	Perspectives of our metric: SFG integrals & H/V descriptor	79
3.2	Interfacial organization of both SAMs and water	98
3.2.1	Statistically averaged density maps of interfacial atoms of both the surface (SAMs) and water	98
3.2.2	Time-resolved density maps of interfacial atoms at OT-S/water and PEG/water interfaces	101
3.2.3	Time-resolved PEG-water HBs density and H/V ratio: Oxygens of PEGs in pursuit of water	103
3.2.4	Number of hydrated PEG atoms counted chain by chain	105
3.3	Dynamics of PEG/water interfaces at larger time- and size-scales: from DFT-MD to Classical FF-MD	108
3.3.1	Average structural and spectroscopic properties	108
3.3.2	Dynamics of hydrophilicity at the PEG/water interface	110
3.3.3	Spatially-resolved x_H maps	112
3.3.4	Quantification of the time-scale of the heterogeneity in hydrophilicity/hydrophobicity at the PEG/water interface	113

4	Pop model: a simplified approach for theoretical SFG calculation and interpretation	119
5	Surface chemistry at the amorphous silica/water interface under different pH conditions	133
5.1	Historical but debated bimodal behavior of pKa values on silica/water interfaces	133
5.2	In situ experimental SFG spectra of silica surface at the interface with water from Y.R.Shen, W.Liu et al.: a missing band in the intermediate pH	135
5.3	Silica surface chemistry as a function of pH revealed by DFT-MD biased metadynamics simulations and theoretical Si-O stretch SFG calculations	136
5.3.1	DFT-MD metadynamics simulations: discovery of a new surface species	137
5.3.2	From pKa values to surface coverage vs. pH	140
5.3.3	Correction of the surface coverage with respect to bulk pH	142
5.3.4	Theoretical calculation of SFG spectra as a function of bulk pH	145
5.4	Consistency between the discovery of Si(5c) ⁻ species and SHG titration experiments	150
5.5	Requisites for the Si(5c) ⁻ species formation at the silica surface	152
5.5.1	Interaction between acidic silanols and water?	153
5.5.2	Specific local morphology in proximity of acidic silanols?	154
5.6	Existence of Si(5c) ⁻ on other amorphous silica and quartz model surfaces?	155
5.6.1	Other amorphous silica model surfaces	156
5.6.2	Crystalline Quartz model surface	156
5.7	Conclusions	157
5.8	Perspectives	159
6	General conclusions and perspectives	211

Part I

Introduction

In natural and technological processes, liquid water plays an enormous range of roles based on its ability to respond to different interfaces, in domains spanning from protein folding [1, 2] to materials [3, 4] to inhomogeneous catalysis [5, 6]. Surfaces can be classified as hydrophobic or hydrophilic depending on the subtle balance between water-water and water-surface interactions. Although the terms hydrophobicity and hydrophilicity are widely used, it remains a major challenge to precisely define hydrophobicity and its relation to water-mediated interactions, properties, and processes.

A conventional metric for quantifying macroscopic hydrophobicity and hydrophilicity is measuring contact angles for extended surfaces. However, the contact angle measurement fails to adequately capture local properties of extended surfaces at the sub-nanometer to few-nanometer length scales. An approach widely used in biophysics is the use of hydrophathy scales, which assign intrinsic hydrophobicity values to chemical functional groups [7]. However, it has become evident that such group-additive methods often fail to accurately predict the response of water to real surfaces. Studies have shown that the way a surface is patterned can affect its hydrophobicity and hydrophilicity in a complex manner [8, 9]. This is because water molecules have directional interactions and can form various hydrogen bonded networks with neighboring water molecules and with surface functional groups, resulting not only in local interactions but also in cooperative effects (long range effects). These properties are microscopic in essence and can only be captured by molecular/microscopic descriptors.

Therefore, a crucial question that remains unanswered is, what microscopic descriptors can be used to effectively and widely measure the molecular hydrophobicity and hydrophilicity of a surface?

The complexity of this issue arises from the various ways in which water can interact with a given surface, both microscopically and macroscopically. One aspect that has been extensively studied is the collective structure of water at the nanoscale. However, to understand water's structure, we must first develop a precise measure of its structural properties, such as density, water dipole orientation, strengths and numbers of hydrogen bonds and tetrahedrality, etc. Additionally, the dynamics of water (e.g. hydrogen bond collective dynamics) is also influenced by the surface.

The current microscopic knowledge of the specific collective structure of water at aqueous interfaces is still limited. The major cause originates from the need for techniques to selectively probe the thin interface between water and the surface with which it interacts. Infrared (IR) spectroscopy can not selectively probe the sub-nanometric interfacial region [10]. While scanning tunneling microscopy (STM) and atomic force microscopy (AFM) are capable of probing local structures at interfaces, their measurements are strongly affected by the dynamics [11]. X-ray absorption spectroscopy was for instance utilized to study water species at an electrode-water interface [12], but it is not capable of providing information on the organization of the H-bond

network at the interface.

Among the available experimental techniques, Sum Frequency Generation (SFG) [13] spectroscopy is a surface-specific technique. SFG spectroscopy, first reported by Shen *et al.* in 1987 [14, 15], is indeed a powerful tool for interface characterization. As a non-linear vibrational spectroscopy technique, SFG depends on the second-order susceptibility $\chi^{(2)}(\omega)$, which is zero in centrosymmetric media (e.g. bulk of liquids and solids). It is non zero only at interfaces where the 3D H-bonded network of liquid water is broken by the presence of a surface. Thanks to its surface specificity, SFG spectroscopy has been widely utilized to investigate aqueous interfaces to unravel their vibrational, structural, and dynamical properties on a molecular level [16–23].

SFG spectroscopy is capable of directly detecting the presence of dangling OH groups at interfaces, i.e. OH groups of water molecules that are not H-bonded to the surface, which are hence considered to be a key indicator for molecular hydrophobicity [24–26]. These dangling OH groups have a clear fingerprint in the 3550-3800 cm^{-1} frequency range, and their detection can reveal the presence of hydrophobic regions at macroscopically hydrophilic surfaces, such as heat-treated silica [23, 27] and 0001- α -alumina [28]. Recently, Kim *et al.* [29] have related the ratio of SFG bands of dangling OH groups versus H-bonded OH groups to surface hydrophobicity through surface adhesion energy. However, this fingerprint disappears at hydrophilic aqueous interfaces, where the free OH groups are not present anymore.

Furthermore, SFG spectroscopy alone cannot provide direct information on the structure and connectivity of the water H-bond network at the interface (in particular because this collective H-bond network is usually oriented parallel to the surface, thus SFG inactive). It cannot quantitatively track the transition from hydrophobic to hydrophilic surfaces and provides a direct quantitative measure of the hydrophobicity or hydrophilicity of the surface. This requires a deeper insight into the specific water network formed at the interface. This can only be obtained by molecular dynamics (MD) simulations.

A number of microscopic descriptors of hydrophobicity have been developed in recent years through molecular dynamics simulations, each one providing unique insights into the properties of interfacial water. One such descriptor is based on density fluctuations. As demonstrated by Godawat *et al.* [30], large local density fluctuations and higher probabilities of low water density are clear signatures of hydrophobicity. This led to the development of local compressibility as a powerful theoretical probe to map microscopic hydrophobicity/hydrophilicity [8]. Another approach to probing the interfacial structure of water is through statistical analysis of the orientational configurations of interfacial water molecules, as proposed by Shin *et al.* [31]. Additionally, recent studies of Pezzotti *et al.* have shown that the preferential orientation of hydrogen bonds formed by water in the topmost interfacial layer can be used to determine the degree of local hydrophilicity/hydrophobicity [26, 32, 33].

While all these descriptors developed through MD simulations can effectively characterize and quantify molecular hydrophobicity/hydrophilicity, they are not directly related to experimental observables. In particular, there is no direct relationship with SFG spectroscopy. This is one goal of this PhD to provide a molecular descriptor of hydrophobicity and hydrophilicity that is directly correlated with spectroscopic quantities.

In this manuscript, a metric combining a structural descriptor given by MD simulations and SFG spectroscopic quantities is developed for quantifying the molecular hydrophobicity/hydrophilicity of any aqueous interface. This metric enables to build the connection between structure and spectroscopy, allowing for a more comprehensive understanding of hydrophobicity and hydrophilicity at the molecular level.

The development of this metric (named H/V descriptor) is detailed in Chapter 3, section 3.1. We further extended the application of the structural H/V descriptor to time-resolved and spatially-resolved H/V, that successfully captured the dynamics and heterogeneity of surface hydrophobicity/hydrophilicity, as detailed in section 3.2 and 3.3.

Despite the fact that SFG is a powerful tool for interrogating interfaces, the interpretation of experimental SFG spectra in terms of microscopic surface structures is highly challenging. Theoretical analyses are mandatory in order to disentangle the observed features and fingerprints. Following the pioneer work of Morita *et al.* [34, 35], the second order susceptibility $\chi^{(2)}(\omega)$ of water can be theoretically calculated from molecular dynamics simulations by the Fourier transform of the correlation function of the total dipole moment and polarizability tensor of the system. Combining SFG spectroscopy with MD simulations allows for straightforward knowledge of the interfacial water organization and its direct signatures into spectroscopic fingerprints.

Both Density Functional Theory based MD (DFT-MD) and classical MD simulations have been employed in the literature to calculate theoretical SFG spectra [32, 33, 36–39], with different strengths and limitations. Since the molecular dipole moments and polarizability tensors are needed to be evaluated at each time step, the DFT-MD calculation of SFG spectra through localization of the electron density [40], though being precise, is highly computationally costly. On the other hand, the classical MD calculation of SFG spectra is much less computationally demanding, hence allows for a better sampling in both size- and time scales. However, in order to accurately calculate molecular dipole moments and polarizabilities, specific force fields (often polarizable and flexible) are required, which are not only difficult to design but also very much system dependent, thus with limited transferability.

To overcome the above limitations, another approach was introduced in the literature by using velocity-velocity correlation functions (VVCf) with Raman and atomic polar tensors (APTs) of water being parameterized from ab initio calculations once and for all (for water) [26, 27, 32, 33, 41–43]. There

is thus no need for the evaluation of molecular dipole moments and polarizabilities at each time step anymore, which highly reduces the computational cost of DFT-MD based calculation of SFG spectra. Furthermore, the VVCF method is based on the correlation functions of velocities only, fast to converge and immediate to generate as velocities are intrinsic ingredients of MD. This method has been validated by the good agreement between calculated and experimental spectra for numerous interfaces, e.g. air/water interface [32] and silica/water interfaces [27, 33, 43]. The theory is described in section 2.3 of this thesis.

However, the interpretation of SFG spectra remains challenging due to the overload of information contained in the signals. The following discussion is for the 3000–4000 cm^{-1} spectral range where OH groups have signatures.

The first problem of the existing methods for interpreting the SFG signals is the compensation of positive and negative intensity in peaks due to opposite orientations of the OH groups of water at the interface. Moreover, both top-surface OH groups (e.g. silanols on silica surfaces) and water OHs contribute to the SFG signal, which is hard to be dissected. Also, the SFG signals in experiments have spectral contributions from different water layers, each of these layers have distinct structural properties that are combined in the final bands of the SFG. Full interpretation of a SFG spectrum thus requires the deconvolution of the total spectrum and the assignment of the different spectroscopic fingerprints to the corresponding structural populations of OH groups of both water and top-surface.

Based on distinct structural properties that are now well defined [32, 33], an aqueous interface can be separated into three layers: the Binding Interfacial Layer (BIL), i.e. the top-surface and the water layer directly in contact with the surface; the Diffuse Layer (DL), i.e. the subsequent water layer with bulk-like water reoriented by the surface charges and associated electric field [44]; and the liquid bulk layer. Only BIL and DL are SFG active because of the broken centrosymmetry. In theoretical SFG, it is “easy” to calculate separately the SFG signals arising from the water in the BIL and from the water in the DL. We can thus deconvolve the total SFG spectrum with the following expression, all detailed in ref. [33] and in Chapter 2, section 2.4:

$$\chi^{(2)}(\omega) = \chi_{BIL}^{(2)}(\omega) + \chi_{DL}^{(2)}(\omega) \quad (1)$$

where $\chi_{BIL}^{(2)}(\omega)$ and $\chi_{DL}^{(2)}(\omega)$ are the spectral contributions from BIL-water and DL-water.

The physics behind $\chi_{BIL}^{(2)}(\omega)$ and $\chi_{DL}^{(2)}(\omega)$ are different. $\chi_{DL}^{(2)}(\omega)$ arises solely from the bulk-like water molecules reoriented by the electric field generated by the surface charges, leading to a universal two-band signal centered at 3200 cm^{-1} and 3400 cm^{-1} , whose sign (in Heterodyne-Detected SFG: HD-SFG) and amplitude depend on the surface potential value [33, 44]. This is detailed in Chapter 2, section 2.5.1. $\chi_{BIL}^{(2)}(\omega)$, on the contrary, is not universal. It is indeed composed of contributions from the water molecules that

directly interact with the surface and the subsequent layer (DL). Interactions between BIL-water and the surface depend on the types and the local distribution of surface functional groups, on the surface morphology and topology, etc. This will result in surface-specific fingerprints in a SFG spectrum that only simulations can properly dissect.

The second work presented in this manuscript is a simplified method that constructs a SFG spectrum of a given aqueous interface as the sum of the spectral contributions from populations not only located in different layers but also having different interactions with the top-surface of the solid. This method is named as “pop model”. This is detailed in Chapter 4. With this new method, the computational cost of SFG spectroscopy is further reduced, even from the VVCF method. Other physical and chemical advantages will be presented in Chapter 4.

The populations can be readily obtained from any numerical simulation (e.g. MD for us, but it could be Monte-Carlo as well), while the spectroscopic fingerprints of each population is pre-computed (by DFT-MD) into a database. We will show that a limited set of populations is sufficient to predict and interpret the spectra of a large variety of water-vapor, water-oxide, and water-organic interfaces, encompassing hydrophobic to hydrophilic interfaces, and neutral to charged environments. See Chapter 4.

By combining the spectroscopic contributions of all interfacial populations, we can obtain a complete theoretical SFG spectrum in the O-H stretch region for a selected aqueous interface (including signals arising from both water and surface) without calculating anymore any correlation function [34, 35, 42]. Moreover, this SFG calculation does not require anymore long time-scale convergence of the signal. Also, the SFG calculation does not require anymore a DFT-based (precise) dynamics. The “pop model” can be applied to a classical FF-MD trajectory, while the SFG spectrum is of quantum DFT quality. All this will be explained in Chapter 4. Besides providing computational efficiency, the “pop model” provides a direct interpretation of a SFG spectrum with a clear understanding of the spectroscopic contributions from each interfacial population.

Besides calculating SFG spectra of an aqueous interface in the O-H stretch region, the “pop model” can also be applied to calculate SFG in the phonon region to directly probe the top-surface structure.

The molecular understandings of surface chemistry of silica surfaces in contact with water have long been a hard nut to crack due to difficulties in probing buried oxide surfaces [45]. In 1992, Eienthal and coworkers [46] first employed the second harmonic generation (SHG) technique to monitor the silica/water interface under different pH conditions. They discovered two pKa values associated to the surface silanol deprotonation, giving rise to the historical bimodal titration behavior, which has been extensively discussed in the literature. The same bimodal behavior of silica surface deprotonation was further observed in many subsequent experimental works, e.g. refs. [47–49].

Moreover, a substantial amount of theoretical works (e.g. refs. [50–53]) have been dedicated to investigating the microscopic origins of the bimodal behavior of silica surface acidities. By performing DFT-MD simulations on the (0001) α -quartz surface, Sulpizi *et al.* [50] ascribed the bimodal acid-base behavior to the surface geometry, in particular to the different O-H orientations of silanols at the top-surface. Using the same method, Pfeiffer *et al.* [51] further confirmed the influence of surface geometry and the solvation structure of the conjugate base on acidities (pKas) of silanols on amorphous aqueous silica surfaces.

Despite extensive works done on the water side, only a limited amount of research has been conducted on the oxide side due to the difficulties of experimentally probing buried surfaces of insulating solids. Hence, the surface chemistry behind the bimodal behavior has been debated without the understanding of the surface chemistry triggered by the deprotonation of surface silanols under different pH conditions.

The third part of this manuscript, Chapter 5, consists of investigating the chemistry of the silica surface under varying pH conditions by combining DFT-MD simulations and both experimental and theoretical pH-dependent SFG spectra. These spectra are calculated through the new “pop model” discussed above. This chapter will show that such model is mandatory in order to access to the theoretical SFG spectra of the silica surface as a function of pH, without the need of possessing DFT-MD simulations under different pH conditions (pH-dependent simulations that are in any case out of reach of DFT-MD and also of FF-MD, as will be explained in Chapter 5).

From SFG experiments and DFT-MD, we demonstrated the existence of a new surface species at the aqueous silica surface during deprotonation processes. This joint theoretical/experimental framework is necessary to reach such final achievement.

Outline of this manuscript

In this manuscript we will first provide a general introduction to the basis of Molecular Dynamics (MD) simulations (Chapter 1) and Sum Frequency Generation (SFG) spectroscopy (Chapter 2). Regarding the MD simulations, both Density Functional Theory based MD simulations (DFT-MD) and classical MD simulations are presented, as these two theoretical methods were both used to investigate aqueous interfaces in this work. As regards SFG spectroscopy, we first provide a general introduction on the technique, followed by an overview on the theory and existing methodologies developed up-to-now to calculate theoretical SFG spectra from MD simulations (see section 2.3). We will then present our novel methodology to extract experimental surface potential by combining experimental and theoretical SFG spectra, based on the deconvolution of BIL/DL contributions to SFG signals, see chapter 2, section 2.4. The method for SFG calculation in the Si-O stretch region, which is used for the silica surface characterization in Chapter 5 is

presented in Chapter 2, section 2.6.

Chapter 3 focuses on the development of our metric for measuring molecular hydrophobicity of aqueous interfaces, that combines structural H/V and spectroscopic descriptors based on a platform composed of a set of self-assembled monolayer (SAM)/water interfaces of various hydrophobicity/hydrophilicity. The rest of this chapter is dedicated to investigating the dynamics and the heterogeneity of hydrophobicity/hydrophilicity of SAM/water interfaces by our developed time-resolved and spatially-resolved H/V descriptor.

In Chapter 4, a new method of theoretical SFG spectroscopy calculation, the “pop model” is presented in the form of a paper that is still in its working form.

In Chapter 5, we uncover the chemistry at the aqueous silica surface under different pH conditions through metadynamics DFT-MD simulations combined with a novel experimental SFG technique and innovative theoretical strategy for SFG calculations.

The most relevant results from this PhD work are summarized in section 6 and put in perspective with possible future directions in the research field dedicated to the characterization of aqueous interfaces.

The list of the publications related to this PhD work is provided on page 11 of this manuscript. These papers include published works as well as papers submitted or in preparation. They are also reported in the various chapters.

List of Publications

Papers

1. On the trail of molecular hydrophilicity and hydrophobicity at aqueous interfaces.
W. Chen, S.E. Sanders, B. Ozdamar, D. Louaas, F.S. Brigiano, S. Pezzotti, P.B. Petersen, M.-P. Gaigeot.
J. Phys. Chem. Lett., **14** 1301–1309, (2023).
reported pages 81–97
2. Unveiling structural evolution and reconstruction of oxide surface in water.
X. Li and F.S. Brigiano and S. Pezzotti and X. Liu and W. Chen and H. Chen and Y. Li and H. Li and Y.R. Shen and M.-P. Gaigeot and W.-T. Liu.
Under review by Nature, 2023.02.
reported pages 160–210
3. A simplified method for theoretical sum-frequency generation spectroscopy calculation and interpretation: the “pop model”.
W. Chen, S. Pezzotti, D. Louaas, F.S. Brigiano, MP Gaigeot.
Working form reported pages 122–132

Book chapter

1. Spectra of large molecules from direct dynamics, and some digression for condensed phase spectroscopies.
S. Bougueroua, V. Chantitch, W. Chen, S. Pezzotti, M.-P. Gaigeot.
Chap. 11, pp. 416–516 in “A computational approach to vibrational dynamics of molecules”, *J. M. Bowman (Ed.)*, World Scientific Press (2022).

Part II

Theoretical Methods

Chapter 1

Molecular Dynamics

I choose in this chapter to write only some of the essential elements of the theories of Density Functional Theory based Molecular Dynamics (DFT-MD) simulations and Classical Molecular Dynamics (CMD or FF-MD) simulations, applied in the present thesis work. More details can be found in other PhD manuscripts from the group (see for instance the thesis of Dr. Morgane Pfeiffer, NNT: 2016SACLE022, the one of Dr. Jerome Mahé, NNT: 2017SACLE043, the one of Dr. Simone Pezzotti, NNT: 2019SACLE008 and the one of Dr. Flavio Siro Brigiano, NNT: 2020UPASF025).

1.1 Molecular dynamics

The application of Molecular Dynamics (MD) simulations is a powerful tool in physics, chemistry, materials science, and multiple branches of engineering. They provide a direct route from the knowledge of microscopic details of a system to its macroscopic properties of experimental interest. In this section, the principles of the MD simulations will be briefly introduced.

The MD simulations theoretically predict the motion of all species of a molecular system for a given interval of time (from fs to ns and beyond). When the electronic representation is employed, in the most common cases, the Born-Oppenheimer approximation is adopted in the MD simulations, so that the motions of the nucleus and the electrons are treated separately. The nucleus is usually heavy enough to be described classically, maintaining in most of the cases a good description of the structural and dynamical properties of the investigated system. The motion of the nuclei can thus be simulated by solving the Newton's equations of motion step-by-step in discrete intervals of time:

$$m_i \frac{\partial^2 \vec{r}_i}{\partial t^2} = \vec{F}_i \quad \forall i, \quad (1.1)$$

where m_i is the mass and \vec{r}_i is the position vector (expressed in cartesian coordinates) of each i nucleus of the system. \vec{F}_i is the force acting on nucleus i , expressed as the negative of the energy gradient:

$$\vec{F}_i = -\frac{\partial V}{\partial \vec{r}_i} \quad \forall i,$$

where V is the potential energy function of all coordinates \vec{r}_i .

There are different methods for determining the potential energy function V . Both the ab-initio way and the classical approach are introduced in this chapter. In ab-initio MD simulations, the electrons are explicitly treated and V is obtained by solving the time-independent electronic Schrödinger equation (eq-1.10). Such approach is accurate but also computationally expensive. Therefore, ab-initio simulations are usually limited in size- and time-scales: systems of hundreds of atoms simulated for tens of picoseconds (and nowadays a few 100' ps). When bigger systems and larger simulation timescales are needed, it is mandatory to use a more approximated estimation for the potential V to reduce the computational cost. This is classical MD simulations, where the electrons are not explicitly treated and V is approximated by parameterized functions, known as “force fields”.

In this thesis, both DFT-MD and Classical MD simulations have been carried out depending on needs of investigations.

1.1.1 Integration algorithms of the equations of motion

Independently of the kind of MD, either classical or ab-initio, the dynamical behavior of the simulated system (with classical nuclei) is obtained by numerical integration of the Newton's equations of motions. Propagation of all atomic positions and velocities requires the discretization of time in small time-intervals, defined as the time-step δt of an MD simulation. Many propagation algorithms have been developed, generally based on a Taylor expansion of the particle positions around the positions, resp. velocities, at a certain instant t (i.e. at each MD-step). Hereafter, we drop the i notation (of each nucleus) in \vec{r}_i and \vec{v}_i for simplicity. All the most commonly used propagation methods are derived from the seminal Verlet algorithm based on two Taylor expansions, one for $t + \delta t$ and the other one for $t - \delta t$:

$$\vec{r}(t + \delta t) = \vec{r}(t) + \vec{v}(t)\delta t + \frac{\vec{F}(t)}{2m}\delta t^2 + O(\delta t^3) + \dots, \quad (1.2)$$

$$\vec{r}(t - \delta t) = \vec{r}(t) - \vec{v}(t)\delta t + \frac{\vec{F}(t)}{2m}\delta t^2 - O(\delta t^3) + \dots. \quad (1.3)$$

Summing these two equations, we obtain:

$$\vec{r}(t + \delta t) = 2\vec{r}(t) - \vec{r}(t - \delta t) + \frac{\vec{F}(t)}{m}\delta t^2 + O(\delta t^4). \quad (1.4)$$

The estimate of the new position of nucleus i at time $t + \delta t$ contains an error of the order of δt^4 , where δt is the time step. Note that this algorithm does not use the velocity to compute the new position. However, we can derive the velocity at time t from the known positions at times $t - \delta t$ and $t + \delta t$:

$$\vec{v}(t) = \frac{\vec{r}(t + \delta t) - \vec{r}(t - \delta t)}{2\delta t}. \quad (1.5)$$

It is also possible to cast the Verlet algorithm in a form with the combination of positions and velocities computed at equal times, which is the so-called velocity Verlet algorithm:

$$\vec{r}(t + \delta t) = \vec{r}(t) + \vec{v}(t)\delta t + \frac{\vec{F}(t)}{2m}\delta t^2, \quad (1.6)$$

$$\vec{v}(t + \delta t) = \vec{v}(t) + \frac{\delta t}{2} \left(\frac{\vec{F}(t)}{2m}\delta t + \frac{\vec{F}(t + \delta t)}{2m}\delta t \right). \quad (1.7)$$

Another algorithm equivalent to the Verlet method is the so-called Leap-Frog algorithm, that evaluates the velocities at half-integer time steps and uses these velocities to compute new positions.

It should be noted that all common algorithms derived from the Verlet one are symplectic, thus able to conserve the total energy of the system in their basic form (if δt is sufficiently small). Therefore, a MD simulation performed with these symplectic algorithms without any constraints will provide the time evolution of the simulated system in the NVE (microcanonical) ensemble. However, other ensembles can be exploited (NVT, NPT) if needed. For example, even though the microcanonical ensemble NVE is the most natural for standard molecular dynamics simulations, it does not strictly represent experimental conditions where the temperature can be fixed instead of the total energy. In this respect, one can perform the MD simulations in the canonical NVT ensemble.

Specific algorithms commonly known as thermostat/barostat have been developed to fix temperature/pressure in the MD simulations, thus allowing to exploit NVT and NPT ensembles, which are briefly discussed in the following section.

1.1.2 Temperature and pressure control

In the canonical NVT ensemble, the macroscopic temperature has a specified average value, while the instantaneous observable representing the total energy of the system (i.e. the Hamiltonian H) can fluctuate. Similarly to the NVT ensemble, in the NPT ensemble the pressure has a specified average value, while the instantaneous volume V of the system can fluctuate. These fluctuations vanish in a macroscopic system, but are often non-negligible for the microscopic system sizes considered in the simulations.

Hence, performing a MD simulation in an ensemble other than microcanonical (NVE) requires keeping at least one intensive quantity constant (on average) during the simulation. A modification of the Newtonian MD scheme for generating a thermodynamical ensemble at constant temperature/pressure is called a thermostat/barostat algorithm. Several algorithms have been developed to control the temperature or the pressure of a system.

As regards the temperature control, the use of a thermostat requires the determination of an instantaneous temperature, that will be compared to the reference temperature T_0 of the heat bath to which the system is coupled. Following the equipartition theorem, the average internal kinetic energy of

a system is related to its macroscopic temperature T through the following equation:

$$E_{kin}(t) = \sum_{i=1}^N \frac{1}{2} m_i \bar{v}_i^2(t) = \frac{1}{2} (3N - 6) k_B T(t), \quad (1.8)$$

where k_B is the Boltzmann's constant and $(3N - 6)$ the number of internal degrees of freedom of the system, with N the total number of atoms. Since the instantaneous temperature is directly related to the atomic velocities, one can thus maintain the constant temperature in the MD simulations by imposing the controls on the rate of change of these velocities. Similarly to the temperature coupling, the system can also be coupled to a "pressure bath" to maintain the constant pressure in NPT ensembles.

The most popular thermostat was developed by Nosé [54] and Hoover [55] and further corrected by Martyna *et al.* [56]. Let us consider only one heat bath regarded as a new degree of freedom s with p_s momentum. The coupling of this bath with the studied molecular system is achieved via a fictitious "mass" Q , to be chosen with care. The new Hamiltonian includes an additional kinetic energy term for the heat bath, $p_s^2/(2Q)$, and a "temperature" term involving the number of dynamical degrees of freedom concerned by the thermostat (g), $gk_B T \ln s$. Moreover, momenta of the system coupled to the bath are virtual and have to be scaled by s : $\mathbf{p}_{real} = \mathbf{p}_{virtual}/s$. The full Hamiltonian for a NVT system in the Nosé-Hoover representation is thus presented below:

$$H = \sum_i \frac{\mathbf{p}_{real,i}^2}{2m_i s^2} + V(\mathbf{R}) + \frac{p_s^2}{2Q} + gk_B T \ln s. \quad (1.9)$$

We note that the thermostats should be applied cautiously due to their perturbations on the velocities of the system. In fact, if the average distribution of velocities obtained from NVT MD simulations is correct with respect to the target temperature, the time evolution of the velocities of all simulated species is not, as it is influenced by the artificial noise produced by the coupling with the thermostat. Therefore, if one wants to investigate properties that depend on the instantaneous velocities evolution as a function of time through MD simulations, as all vibrational spectroscopic properties, the microcanonical NVE ensemble has to be chosen. Since the purpose of this thesis work is related to Sum Frequency Generation (SFG) spectroscopy calculated as shown in chapter 2, section 2.3 through velocity-velocity correlation functions, most of the simulations have been performed in the NVE ensemble. Also, the NVE ensemble has to be seen cautiously for equipartition (see ref. [57], where this issue is discussed in details and where it is shown that equipartition is usually fulfilled for condensed matter.)

In the CP2K software for DFT-MD simulations used in this thesis, other strategies are used for temperature control: Canonical Sampling through Velocity Rescaling (CSVR) where a random factor is used to rescale velocities [58] or velocity rescaling (fast but not canonical). In our dynamics, rescaling of velocities is done at the very beginning of equilibration of NVE dynamics

(rescaling for 1 ps then relaxation until equilibrium is reached in the micro-canonical ensemble).

1.1.3 Initial conditions and choice of time step

At the beginning of a MD simulation, we need to input the initial coordinates and assign the velocities of all particles of the system. Despite the fact that a MD trajectory is independent of the initial configuration (after a certain amount of time for the system to reach equilibrium), the choice of the initial set of positions and velocities is still important, especially for ab-initio MD, where the equilibration time-scale is limited to achieve proper thermalization of the system. Initial velocities are generated by extracting the velocity vector components from a normal distribution, and the velocity magnitudes are generated from a Maxwell-Boltzmann distribution at the targeted simulation temperature. This distribution is generated from random numbers, according to the following equation:

$$p(\vec{v}_i) = \sqrt{\frac{m_i}{2\pi k_B T}} \exp\left(-\frac{m_i \vec{v}_i^2}{2k_B T}\right),$$

where $p(\vec{v}_i)$ is the probability of the i -th particle with the \vec{v}_i velocity, m_i the mass of this particle and T the targeted temperature of the system.

Another important parameter of a simulation is the time step δt , which is the crucial parameter in eq 1.4 in sec. 1.1.1 for solving the Newton's equations of motions. The maximum time step that could be chosen is determined by the rate of the fastest process in the system, and typically it is chosen an order of magnitude smaller than this fastest process. Molecular motions, such as rotations and vibrations, occur with frequencies in the range 10^{11} - 10^{14} s $^{-1}$, and this means that time steps of the order of fs (10^{-15} s) or less are required to model such motions with sufficient accuracy. In particular, fast vibrational motions as the OH-stretching investigated in this work, need a time step of the order of 0.1-0.5 fs to be correctly sampled. Furthermore, ab-initio MD requires δt of this order for a proper conservation of the total energy in the NVE ensemble. Since a correct description of the OH-stretching motion and vibration is crucial in this thesis, a time step of 0.4 fs is the standard choice for all DFT-MD simulations and some of the classical MD simulations for SFG spectroscopy calculations discussed in this work. To give an example, a total simulation time of 10 ps requires 25000 steps with this time step.

1.1.4 Periodic boundary conditions

To model continuous systems, like solids and liquids of interest in this work, periodic boundary conditions (PBC) are applied. The atoms of the system to be simulated are hence placed in a suitable box in the cartesian

space, generally cubic or parallelepipedic, which is replicated in all directions of space (the initial box is thus surrounded by 26 identical cubes/boxes, which are again surrounded by 98 boxes, etc). For example, if an atom leaves the central box through the right wall, its image will enter back the central box through the left wall from the neighboring box. The resulting model becomes quasi-periodic, with a periodicity equal to the dimension of the box. Note that the imposed artificial periodicity may cause errors when considering properties which are influenced by long-range correlations, such as for dipolar and charged systems. Corrections for the finite size of MD simulations are also necessary when evaluating long-range collective properties, as for instance surface tension. [59]

1.2 Born-Oppenheimer molecular dynamics (BO-MD) in the Density Functional Theory (DFT) framework

In Born-Oppenheimer (BO) DFT-MD simulations, the Born-Oppenheimer approximation is adopted to separate the motion of nuclei and electrons. In all this work, nuclei are treated classically. The time independent electronic Schrödinger equation is solved at each time step (at a given position of the nuclei) to determine the electronic energy of the system E_e (corresponding to the organization in space of the nuclei):

$$H_e \Psi_e(\{\vec{r}\}) = E_e \Psi_e(\{\vec{r}\}), \quad (1.10)$$

where $\Psi_e(\{\vec{r}\})$ is the electronic wavefunction of the system, a function of the $3N$ cartesian coordinates of the N electrons in the system ($\{\vec{r}\} = \{\vec{r}_i, \vec{r}_j, \dots, \vec{r}_N\}$), and H_e is the electronic Hamiltonian, which is defined as:

$$H_e(\{\vec{r}\}, \{\vec{R}\}) = T_e + V_{ee} + V_{Ne}, \quad (1.11)$$

where T_e is the operator that describes the kinetic energy of the electrons, while V_{ee} and V_{Ne} describe the potential energies of the electron-electron and nuclei-electron interactions, respectively, \vec{R} is the ensemble of the cartesian coordinates of the nuclei.

Once the time independent electronic Schrödinger equation is solved, and the electronic energy of the system E_e is known, the potential energy of the system E_{pot} is obtained by just adding to E_e the potential energy of the nuclei-nuclei interaction ($V_{NN} = \frac{1}{2} \sum_A^M \sum_{B \neq A}^M \frac{Z_A Z_B e^2}{|R_{AB}|}$, where R_{AB} is the distance between the nuclei A and B and $Z_{A/B}$ their atomic numbers).

However, an analytic solution to eq. 1.10 can only be obtained for the hydrogen atom and any hydrogenoid atom as well as the H_2^+ molecule (i.e. only one electron in the system) due to the existence of the V_{ee} term in multi-electronic systems. The presence of the V_{ee} term inhibits the decomposition of the electronic Hamiltonian into mono-electronic terms whose exact solution

of the time independent electronic Schrödinger equation is known. Moreover, for a multi-electronic system, the kinetic energy of one electron also depends on all other electrons, which makes it more complicated in correctly evaluating the T_e term. Therefore, in practice, approximations have to be applied to estimate the electronic wavefunction and energy.

Many approximated solutions have been proposed for eq. 1.10, starting from the seminal Hartree and Hartree-Fock approaches. All these methods differ in accuracy level of the estimation of E_e , which is inversely proportional to the computational cost, i.e. the amount of computational time needed to get the solution. The aim thus becomes to find the better compromise between accuracy and efficiency for a given system. Despite the differences between the various approaches, they all rely on rewriting the electronic wavefunction of the multi-electronic system in terms of mono-electronic wavefunctions, via one Slater determinant (needed to antisymmetrize the resulting total wavefunction since the electrons are fermions), so that:

$$\Psi_e(\{\vec{r}\}) = \frac{1}{\sqrt{N!}} \begin{bmatrix} \psi_1(\vec{r}_1) & \psi_2(\vec{r}_1) & \dots & \psi_j(\vec{r}_1) & \dots & \psi_N(\vec{r}_1) \\ \psi_1(\vec{r}_2) & \psi_2(\vec{r}_2) & \dots & \psi_j(\vec{r}_2) & \dots & \psi_N(\vec{r}_2) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \psi_1(\vec{r}_i) & \psi_2(\vec{r}_i) & \dots & \psi_j(\vec{r}_i) & \dots & \psi_N(\vec{r}_i) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \psi_1(\vec{r}_N) & \psi_2(\vec{r}_N) & \dots & \psi_j(\vec{r}_N) & \dots & \psi_N(\vec{r}_N) \end{bmatrix}, \quad (1.12)$$

where $\psi_i(\vec{r}_j)$ is the i -th spin orbital (mono-electronic wavefunction) populated by the j -th electron. All notations related to spins are removed for simplicity in eq. 1.12. The spin orbitals are then expressed on a basis set, based on either plane-waves or gaussian functions.

Among the different approximated solutions existing for the time-independent Schrödinger equation, the Density Functional Theory (DFT) approach is certainly the most used methodology adopted to run ab-initio MD-simulations, since it usually represents the best level of accuracy affordable for a reasonable computational cost, low enough to perform simulations typically of 100-1000 atoms for a few hundred ps time-scale.

1.3 Density Functional Theory — From wave functions to electron density

The Density Functional Theory (DFT) was introduced in Hohenberg and Kohn's work published in 1964. The original objective of the DFT was to rewrite the time-independent electronic Hamiltonian in terms of the electron density instead of the N -electrons wavefunction. The electron density is expressed as:

$$\rho(\vec{r}) = N \int \dots \int |\Psi_e(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)|^2 d\vec{r}_2 \dots d\vec{r}_N. \quad (1.13)$$

The complicated problem of $3N$ -dimensional space of the Slater determinant (i.e. eq. 1.10) is thus simplified to the 3-dimensional space (x,y,z) of $\rho(\vec{r})$.

The basis of the Density Functional Theory is established by the two Hohenberg and Kohn theorems discussed in the following:

1st. For any interacting many-body electronic system immersed in an external potential $V_{ext}(\mathbf{r})$ (this is the potential arising from the nuclei), $V_{ext}(\mathbf{r})$ is uniquely determined, upon addition of a constant, by the density $\rho(\vec{r})$ of this system in its ground state.

2nd. There is a universal energy $E^{HK}[\rho]$, functional of the density, whatever the external potential $V_{ext}(\mathbf{r})$. For each $V_{ext}(\mathbf{r})$, the ground-state energy is obtained via the variational principle applied to the energy functional: the obtained density corresponds to the exact ground-state density.

From the 1st theorem, we know that all the properties of the system can be determined if the density $\rho(\vec{r})$ is known. The ground-state electronic energy is hence a functional of the density, which is expressed as:

$$E[\rho] = T[\rho] + V_{ee}[\rho] + E_{ext}[\rho], \quad (1.14)$$

where $T[\rho]$ is the kinetic energy functional, $V_{ee}[\rho]$ is the electron-electron interaction energy functional, $T[\rho] + V_{ee}[\rho] = E^{HK}[\rho]$, and $E_{ext}[\rho]$ is the energy related to the external potential, so that $E_{ext}[\rho] = \int \rho(\vec{r})V_{ext}(\vec{r})d\vec{r}$.

The 2nd theorem validates the application of the variational principle on the energy functional in order to determine the ground-state density $\rho_0(\vec{r})$.

In principle, the two theorems from Hohenberg and Kohn provide all the means to calculate the electronic energy solely from the electron density, whereas the eq. 1.14 cannot be simply solved due to the many-body $T[\rho]$ and $V_{ee}[\rho]$ terms. While E_{ext} is known for each given spatial arrangement of the nuclei, the exact expressions of $T[\rho]$ and $V_{ee}[\rho]$ are unknown for interacting electrons, but are known for non-interacting electrons. This knowledge is the starting point for the Kohn-Sham approach, which is widely used nowadays to effectively solve the time-independent electronic Schrödinger equation within the DFT formalism, with specific approximations that are going to be detailed hereafter.

Within the Kohn-Sham approach, given a system with N (interacting) electrons and M nuclei in a given disposition, a model system is associated to it composed of the same M nuclei in the same spatial arrangement, but with the N electrons being non-interacting. The non-interacting model system moreover satisfies the constraint that the electron density of the non-interacting electronic model system is equal to the electron density of the real one (demonstrated by Kohn and Sham). The $T[\rho]$ term of the model system can be then exactly calculated by reintroducing the mono-electronic wave functions ψ_i , which are independent from each others (i.e. independent electrons). The electron density of the electronic non-interacting model system is thus expressed as $\rho(\vec{r}) = \sum_i^N |\psi_i(\vec{r})|^2$, where N is the number of electrons.

The total kinetic energy of the model non interacting electronic system $T_s[\rho]$ then becomes:

$$T_s[\rho] = -\frac{1}{2} \sum_i^N \langle \psi_i(\vec{r}) | \nabla^2 | \psi_i(\vec{r}) \rangle. \quad (1.15)$$

Knowing $\rho(\vec{r})$ of the electronic non-interacting model system, it is thus possible to calculate the classical electron-electron interaction energy, which is known as the Hartree energy $E_H[\rho]$:

$$E_H[\rho] = \frac{1}{2} \int \int \frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}d\vec{r}'. \quad (1.16)$$

However, based on the Kohn-Sham approach with non-interacting electrons, the $E_H[\rho]$ is only a part of the total electron-electron interaction energy $V_{ee}[\rho]$ of the real electronic system. The latter indeed also includes the contribution from electronic correlations (i.e. interactions) and quantum effects known as exchange interactions between electrons due to the fact that the electrons are interacting particles that must follow the Pauli exclusion principle. In the same way, also $T_s[\rho]$ is only a part of the total $T[\rho]$ of the real electronic interacting system, which also includes the contributions from the correlation between the motions of all electrons.

The total electronic energy of the real interacting system (eq. 1.14) can be now rewritten in terms of the known energy terms for the non-interacting model system, and the known and unknown terms can be separated as follows:

$$E[\rho] = T[\rho] + V_{ee}[\rho] + E_{ext}[\rho],$$

$$E[\rho] = T[\rho] + (T_s[\rho] - T_s[\rho]) + V_{ee}[\rho] + (E_H[\rho] - E_H[\rho]) + E_{ext}[\rho],$$

$$E[\rho] = \underbrace{T_s[\rho] + E_{ext}[\rho] + E_H[\rho]}_{known} + \underbrace{(T[\rho] - T_s[\rho]) + (V_{ee}[\rho] - E_H[\rho])}_{unknown}, \quad (1.17)$$

$$E[\rho] = T_s[\rho] + E_{ext}[\rho] + E_H[\rho] + E_{XC}[\rho],$$

where $T_s[\rho]$, $E_{ext}[\rho]$ and $E_H[\rho]$ are the known terms for the electronic model non-interacting system, while the $(T[\rho] - T_s[\rho])$ and $(V_{ee}[\rho] - E_H[\rho])$ are the unknown, non-classical contributions to $T[\rho]$ and $V_{ee}[\rho]$ respectively, which are grouped within the so-called exchange (X) and correlation (C) functional $E_{XC}[\rho]$.

In the Kohn-Sham approach the key point thus becomes the way the $E_{XC}[\rho]$ is expressed and calculated. The $E_{XC}[\rho]$ term can be generally expressed as:

$$E_{XC}[\rho] = \int d\vec{r} F(\rho(\vec{r}), \nabla\rho(\vec{r}), \nabla^2\rho(\vec{r}), \dots), \quad (1.18)$$

where F is a functional written as a Taylor expansion of the electronic density $\rho(\vec{r})$, its first gradient $\nabla\rho(\vec{r})$, its second gradient $\nabla^2\rho(\vec{r})$ and etc.

Many DFT functionals have been proposed by the scientific community in order to describe the $E_{XC}[\rho]$ term. Four main families of DFT-functionals have been developed up-to-now: the Local Density Approximation (LDA), the Generalized Gradient Approximation (GGA), the Hybrids and the meta-GGA. They differ by the order term of the Taylor expansion in eq. 1.18 in which the electron density $\rho(\vec{r})$ is expanded: LDA ($\rho(\vec{r})$), GGA ($\rho(\vec{r})$ and $\nabla\rho(\vec{r})$), meta-GGA ($\rho(\vec{r})$, $\nabla\rho(\vec{r})$ and $\nabla^2\rho(\vec{r})$). Hybrid functionals are based on a GGA expansion with a supplementary term of Hartree exchange that will be described in the next paragraphs.

1.3.1 Exchange-correlation functionals

Local Density Approximation (LDA)

The Local Density Approximation (LDA) was the first, and simplest approximation for the exchange and correlation functionals. These functionals are based on the assumption that the electron density is uniformly distributed over the space, and accordingly express $E_{XC}[\rho]$ as a simple functional of the density (i.e. independent of the density gradient). Since such approximation of a uniform electron gas is appropriate for metals but not for molecules, where the electron density has maxima between covalently bonded atoms and the density gradient is non negligible, LDA functionals provide satisfying energy estimations for metals but not for molecular systems. They are thus non appropriate in order to describe water at interfaces, goal of the investigations performed in this thesis work.

Generalized Gradient Approximation (GGA)

Generalized Gradient Approximation (GGA) functionals were developed as improvements from LDA functionals, by taking now into account not only the dependence of $E_{XC}[\rho]$ on the electron density ($\rho(\vec{r})$) but also on its gradient ($\nabla\rho(\vec{r})$). Within the GGA formalism, covalent bonds are well described, which makes this second generation of functionals a better choice for molecular systems compared to the LDA functionals. GGA functionals, in particular the BLYP functional, have been consequently chosen to perform all the DFT-MD simulations presented in this manuscript. This choice also relies on all the previous investigations performed in the group for aqueous interfaces [32, 33, 50, 60, 61]

Meta Generalized Gradient Approximation (Meta GGA)

Going beyond GGA, the Laplacian of the density $\nabla^2\rho$ has to be inserted in the expression of $E_{XC}[\rho]$. As such, the kinetic energy functional becomes

part of the parameters, increasing the degrees of freedom for fitting but at the price of a more difficult numerical convergence. One example is the Tao-Perdew-Staroverov-Scuseria (TPSS) meta-GGA functional [62], which is an improvement of the PBE GGA functional satisfying additional exact constraints avoiding empirical parameters. It is found to provide good descriptions of both molecules (bond lengths, harmonic frequencies) and solids (lattice constants, surface energies). The group of Truhlar designed several meta-GGA functionals like the Minnesota 06-Local or M06-L [63], in particular for metals and organometallics. However, the derivation strategy of this group is very different since the analytical form of the functionals is not imposed but is rather flexible and parametrization is achieved on a large number of benchmarks. Meta-GGAs are more complex than GGAs, they are also more expensive in terms of computational cost. We have not applied them to our works.

Hybrid Functionals

The motivation for the development of hybrid functionals originates from failures of the GGA functionals due to their incomplete treatment of the exchange interaction, thus requiring improvements in this direction. As a response, hybrid functionals are a class of approximations to the exchange-correlation functionals that incorporate a fraction of exact exchange from the Hartree-Fock theory, while keeping other elements of the exchange-correlation from other sources, for instance LDA or GGA. Despite the improved accuracy reached on average with this third generation of functionals, the higher computational cost than common GGA approximations makes hybrid functionals often too expensive for DFT-MD simulations. For instance, it is the case for the simulations of aqueous interfaces done in this work, which could not be performed with sufficiently large box-dimensions and simulations time-lengths using Hybrid functionals (roughly 40 times more computational expensive than GGA).

1.3.2 Dispersion corrections

Intermolecular interactions represent key quantities when real compounds of chemical and biological interest are concerned. One of the main goals of theoretical chemistry consists in calculating as accurately as possible such energies. Nevertheless, some contributions, namely the dispersion, appear to be much harder to describe than, for instance, electrostatic and exchange parts. Indeed, dispersion originates from quantum fluctuations of the charge distribution, generating instantaneous dipoles that interfere attractively according to the well-known $-\frac{C_6}{R^6}$ law for the energy, where R represents the intermolecular distance between nuclei and C_6 a specific coefficient of the system. Electronic correlation must be included either through special functionals that take this effect directly into account or van der Waals corrections

in DFT or through post-HF methods. Let us stress that the absolute requirements are numerical accuracy and speed of calculation. Indeed, we have to deal with energies of around 1 kcal/mol or 2 mHa, which is extremely small, and we do DFT-MD simulations where dispersion has to be evaluated at each time step.

Standard DFT, with LDA, GGA or hybrid functionals, is unable to take dispersion into account sufficiently properly because it arises from long-range electronic correlations, whereas DFT can only treat correlation at short distances (\sim a few Å, for overlapping charge densities). For instance, BLYP or B3LYP can provide purely repulsive interaction potentials for weakly bound complexes [64]. Besides, dispersion is needed to improve thermodynamics [65]. Several empirical corrections have been thus developed to overcome this problem, as for instance the Grimme D2 correction [66] used in our simulations, which is a semi-empirical vdW corrections to DFT functionals where the dispersion energy is included as an additional term to the Kohn-Sham energy. It is made of pairwise interactions and does not depend on the electron density but on atomic coordinates. This type of correction however does not correct the density but only the electronic energy.

In the D2 method of Grimme *et al.* [67], the correction term is expressed as:

$$E_{\text{disp}} = -\frac{1}{2} \sum_{i=1}^{N_{\text{at}}} \sum_{j=1}^{N_{\text{at}}} \sum_{\mathbf{L}} \frac{C_{6ij}}{r_{ij,\mathbf{L}}^6} f_{d,6}(r_{ij,\mathbf{L}}), \quad (1.19)$$

where the first two summations are over all N_{at} atoms in the unit cell and the third summation is over all translations of the unit cell $\mathbf{L} = (l_1, l_2, l_3)$. C_{6ij} is the dispersion coefficient for the atom pair ij , $r_{ij,\mathbf{L}}$ is the distance between atom i located in the reference cell $\mathbf{L} = 0$ and atom j in the cell L and the term $f(r_{ij})$ is a damping function whose role is to scale the interactions such as to minimize the contributions from interactions within typical bonding distances. Tobias and coworkers showed that BLYP+D2 strongly helped in reproducing the properties of the liquid water/air interface in molecular dynamics simulations with CP2K [68], like the Hbond population as a function of the interfacial depth.

Grimme *et al.* further proposed a refined D3 method [69], expressed as:

$$E_{\text{disp}} = -\frac{1}{2} \sum_{i=1}^{N_{\text{at}}} \sum_{j=1}^{N_{\text{at}}} \sum_{\mathbf{L}} \left(f_{d,6}(r_{ij,L}) \frac{C_{6ij}}{r_{ij,L}^6} + f_{d,8}(r_{ij,\mathbf{L}}) \frac{C_{8ij}}{r_{ij,\mathbf{L}}^8} \right). \quad (1.20)$$

where C_{8ij} is a higher order pairwise dispersion coefficient, dispersion with decay in r^{-8} .

In all the DFT-MD simulations presented in this work, we chose the D2 correction over the D3 correction due to the following reasons. In terms of the description of O-H vibrations, in the thesis of Dr. Morgane Pfeiffer, NNT: 2016SACLE022 (P 139), the VDOS analyses have been performed for 64 water molecules with BLYP-D2 and BLYP-D3, respectively. Both D2 and D3 dispersion corrections allow us to really improve the position of the OH

stretching band, i.e. around 30 cm^{-1} lower than the experiments (compared to the 250 cm^{-1} red shift with the BLYP functional without dispersion correction).

In terms of structural properties, the radial distribution functions of BLYP-D2 and BLYP-D3 water have been plotted and compared in the thesis of Dr. Morgane Pfeiffer, NNT: 2016SACLE022, which both correctly reproduce the liquid water. Both D2 and D3 corrections provide good descriptions for O-H vibrations and structural properties of water. However, D3 is more computationally consuming compared to D2 since it considers the third-order terms in the dispersion correction. The choice of D2 correction is a good compromise between the computational efficiency and the performance.

1.4 The CP2K software and the computational set-up chosen for the DFT-MD simulations in this thesis

All Born-Oppenheimer (BO) DFT-MD simulations presented in this thesis have been performed with the Quickstep/CP2K package [70, 71]. The methodology implemented in Quickstep uses a dual Plane Wave and Gaussian basis set over which the wavefunction of the simulated system is expanded. The advantage of such approach is that short-range interactions, like H-Bonds, which are well described by Gaussian basis sets, can be treated in the direct space using the Gaussian basis set, while long-range interactions, like electrostatic interactions, are more conveniently treated with the Plane-Wave basis set in the reverse space. Fast Fourier Transform is used to pass from one space to the other in the calculations.

For all MD-simulations performed in this thesis the BLYP [72, 73] GGA functional has been adopted in combination with mixed Gaussian/Plane Waves basis sets and GTH pseudopotentials [74] (treating all electrons would be computationally too expensive). Grimme D2 corrections [67, 75] are also taken into account for a better description of van der Waals interactions. This electronic set-up follows previous works in the group [26, 27, 32, 33, 43, 60, 61] on aqueous interfaces, adequately providing structural and vibrational properties of these interfaces. The DZVP-MOLOPT-SR gaussian basis set has been systematically used for the performed simulations (unless otherwise indicated), augmented with a 400 Ry plane wave basis, which generally provided the best compromise between accuracy and computational times (possible different choices for the plane-wave cut-off are properly indicated in the manuscript).

The box-size of $13\times 13\times 85\text{ \AA}^3$ has been chosen in our DFT-MD of SAMs/water interfaces in chapter 3 and $13\times 13\times 35\text{ \AA}^3$ has been chosen for amorphous silica/water interfaces presented in chapter 5. Periodic boundary conditions have been applied in all three spatial directions for all these simulations. A vacuum layer of 25 \AA along z-direction has been included in the SAM/water interfaces to separate the periodic replica. In the silica/wa-

ter interface, the vacuum layer is 10 Å. BO DFT-MD dynamics have been performed, the electronic wavefunction being calculated at each time step, the classical nuclei displacements being obtained through the Velocity-Verlet algorithm, with a time-step of 0.4 fs. The simulations protocol adopted for both the SAM/water interfaces and the silica/water interfaces consists of an equilibration dynamics of 10 ps duration (in the NVE ensemble, however allowing rescaling of velocities whenever necessary) followed by pure NVE production runs of 100 ps for SAM/water and 50 ps for silica/water, over which structural and spectroscopic properties have been calculated.

1.5 Classical MD simulations

Despite the accuracy of the DFT based molecular dynamics, dynamical properties at longer time scales and the spatial heterogeneity of an interface are not able to be captured due to the limited time scales and the box sizes of the DFT-MD simulations. Therefore, as mentioned in Section 1.1, we have also carried out a series of classical MD simulations (denoted CMD or FF-MD for Force Field MD in the text) with larger boxes and longer time-scales. In FF-MD dynamics, the electronic distribution is usually approximated in a rather coarse-grained way either by putting fixed partial charges on interaction sites or by using an approximate model for polarization effects [76]. The time scale of classical systems is determined by the time of intermolecular collision events, rotational motions or intramolecular vibrations, which are orders of magnitude slower than those of electron motions. Therefore, the integration time step of the MD equations is larger than in DFT-MD (usually 0.5 fs for flexible FFs to 1-5 fs) and trajectory lengths are of order of 100' ns and accessible lengths of order 10–1000' Å.

Neglecting the complex description of the electronic dynamics, the N atoms of the system are now described as point-like centers which interact through pairwise or multi-body interaction potentials. The system is modeled by a set of analytical functions (the so-called force field) that depend on the mutual position of the atoms in the configuration. The FF requires predefined parameters. The system energy is obtained through the FF model and the force acting on each particle is thus calculated by $\vec{F} = -\vec{\nabla}V$, which is further used in the numerical integration of the Newton's equations of motions, as explained in section 1.1.1 to obtain the dynamical behavior of the simulated system. The generalities of the principles of force fields and the choice of force fields depending on the specific systems and the needs of investigations are introduced hereafter.

A force field, in the molecular dynamics dictionary, is a combination of the functional forms used to express the intramolecular and intermolecular potential energy of a collection of N atoms, and the corresponding parameters which determine the energy of a given configuration. The functions and the parameters are obtained either from experimental data or from (accurate) quantum mechanics calculations. They are often refined by comparing the calculated condensed phase properties with experiments. The expression of a force field is generally divided into an intramolecular term and an intermolecular term:

$$V(\vec{R}_1, \dots, \vec{R}_N) = V_{\text{intramolecular}}(\vec{R}_1, \dots, \vec{R}_N) + V_{\text{intermolecular}}(\vec{R}_1, \dots, \vec{R}_N). \quad (1.21)$$

Both the intramolecular and intermolecular potentials can be described based on models of different accuracy and different computational costs, which will be briefly summarized in the following sections.

1.5.1 Intramolecular interactions

The intramolecular interactions consist of bond stretchings, bond angle bendings, and dihedral angular motions, all together representing the internal motions of the molecules, as illustrated in Fig. 1.1 and expressed by:

$$V_{intramolecular} = V_s + V_b + V_\phi, \quad (1.22)$$

where V_s , V_b , and V_ϕ represent the terms of bond stretching, bond angle, and torsional angle, respectively.

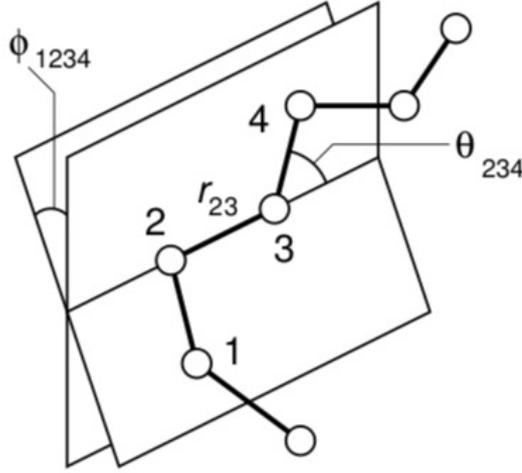


Figure 1.1: Geometry of a simple chain molecule, illustrating the definition of interatomic distance r_{23} , bend angle θ_{234} , and torsion angle ϕ_{1234} used in the intramolecular potential energy function of a Force Field.

The basic descriptions of the intramolecular terms are based on harmonic oscillators. By only considering the harmonic terms, we have:

$$\begin{aligned} V_{intramolecular}(\vec{R}_1, \dots, \vec{R}_N) = & \sum_{\text{bonds}(i,j)} \frac{1}{2} k_{r_{ij}} (\vec{r}_{ij} - \vec{r}_{0,ij})^2 \\ & + \sum_{\text{angles}(i,j,k)} \frac{1}{2} k_{\theta_{ijk}} (\theta_{ijk} - \theta_{0ijk})^2 \\ & + \sum_{\text{torsions}(i,j,k,l)} \sum_n k_{\phi,n} [\cos(n\phi_{ijkl} + \delta_n) + 1]. \end{aligned} \quad (1.23)$$

The first term in eq. 1.23 expresses the sum of the bond stretching motions, with $-\vec{r}_{0,ij}$ being the equilibrium bond length and r_{ij} being the bond length for every pair $i-j$ of covalently connected atoms. The second term is the sum of the bond angle bending terms where θ_{0ijk} is the bond angle at equilibrium, θ_{ijk} is the bond angle for each set of three covalently connected atoms $i-j-k$. The third term represents the sum of the torsional motions with ϕ_{ijkl} being the dihedral angle formed by four connected atoms. In principle, this latter is an expansion in trigonometric functions with different values of

n , the multiplicity (i.e. the number of minima in a rotation of 2π around the $j - k$ bond); many force fields fix $n = 3$ [77]. The ensemble of intramolecular parameters of the FF ($k_{r_{ij}}, \vec{r}_{0ij}, k_{\theta_{ijk}}, \theta_{0ijk}, k_{\phi,n}, \delta_n$) are obtained either from quantum mechanical simulations or empirically from experiments. The force constants in the bonded terms are usually parameterized based on the vibrational frequencies and conformational energies, and the geometries parameters are determined by computations or experiments (e.g. X-ray).

In addition to simple harmonic descriptions, anharmonic terms (e.g. Morse potential and cubic term for bond stretching) are added in some of the new generations of force fields for a more precise modeling, e.g. AMOEBA [78] for biomolecules. Moreover, since the bonds, angles and torsions are not independent in molecules, the implementation of the potential energy cross terms (e.g. bond–bond, bond–angle, angle–angle, bond–torsion, and angle–angle–torsion) is necessary for the accurate calculation of vibrational frequencies. This is rarely done in the FFs as this is a highly complex task to parameterize the associated parameters for the analytical expressions.

To give an example used in this work, the SPC-Marti force field [79] is a SPC-based flexible water model where a Morse potential is used for the O-H stretching and the cross terms of bond-bond, bond-angle, and angle-angle are incorporated. The parameters for intermolecular terms remain the same as in the rigid SPC model [80]. According to the previous work of our group in collaboration with the group of Prof. Predota [81], compared to the other SPC-based flexible water models (e.g. SPC-Smirnov [82], SPC-Ferguson [83]) where the cross terms are not incorporated, the water SFG spectra of various aqueous interfaces (air/water, fluorite/water, silica/water) calculated based on the SPC-Marti model are in best agreement with the experimental spectra and with the ab initio calculated ones. These results reiterate the importance of the implementation of cross terms in the intramolecular part of the force fields to more accurately describe the vibrational frequencies. The use of the SPC-Marti model in the classical MD simulations done in this work to calculate the SFG water spectra will be detailed in Chapter 3.

1.5.2 Intermolecular interactions

The intermolecular interaction is usually composed of van der Waals (vdW) interactions (repulsion and dispersion) and electrostatic interactions as:

$$V_{\text{intermolecular}}(\vec{R}_1, \dots, \vec{R}_N) = V_{\text{vdW}} + V_{\text{elec}}, \quad (1.24)$$

where the vdW interactions are normally represented by a Lennard-Jones potential or a Buckingham potential, and the electrostatic potential is expressed based on the Coulomb term by considering the electronic and nucleic densities as point charges centered on each atom. Note that the intermolecular interactions invariably exclude 1–2 and 1–3 pairs of atoms belonging to the same molecule. In this section, different models to describe both the vdW interactions and the electrostatic potential are introduced and compared.

Representation of vdW interactions: Lennard-Jones potential or Buckingham potential?

To describe the vdW interactions between two atoms in a force field, the most commonly used model is the Lennard-Jones (LJ) potential. The LJ potential is the building block of most of the force fields in the literature, composed of one steep repulsive term at short intermolecular distance and one smoother attractive term at longer intermolecular distance, representing the London dispersion forces (induced dipole-induced dipole interaction). The LJ expression is:

$$V(r_{i,j}) = 4\epsilon_{i,j}\left[\left(\frac{\sigma_{i,j}}{r_{i,j}}\right)^{12} - \left(\frac{\sigma_{i,j}}{r_{i,j}}\right)^6\right], \quad \forall pair(i,j) \quad (1.25)$$

where $\epsilon_{i,j}$ is the well depth and a measure of how strongly the two particles (i,j) attract each other, $\sigma_{i,j}$ is the distance at which the intermolecular potential between the two particles is zero, and $r_{i,j}$ is the distance between the two particles. We note that $\sigma_{i,j}$ gives a measurement of how close two nonbonding particles can get in the classical representation. It is referred to as the van der Waals radius, which is equal to 1/2 of the internuclear distance at the minimum of $V(r_{i,j})$. Because of the twelfth power of the repulsive term and the sixth power of the attractive term, the LJ model is also called the 12-6 LJ model.

Despite the computational expediency of the LJ potential, this model is not the most faithful representation of the potential energy interaction due to the too steep repulsive part, which often leads to an overestimation of the pressure in dense systems. According to quantum ab initio calculations, a more appropriate description of the repulsive potential between two atoms is an exponential form, which is the one adopted in the Buckingham potential, as expressed by:

$$V(r_{i,j}) = A_{i,j} \exp(-B_{i,j}r_{i,j}) - \frac{C_{i,j}}{r_{i,j}^6}, \quad (1.26)$$

where, $A_{i,j}$, $B_{i,j}$ and $C_{i,j}$ are constants. The two terms in $V(r_{i,j})$ refer to the repulsion and the attraction terms, respectively, between the two particles (i,j).

The Buckingham model describes the intermolecular interactions in a more accurate way by using the exponential type of the repulsive term, while at the same time it increases the computational costs due to the exponential function that is computationally more expensive to calculate than the $1/r^n$ term. Some of the force fields of the new generations (e.g. MM2 [84] and MM3 [85]) use the Buckingham model instead of the LJ potential for the vdW description. However, the LJ potential is usually preferred in biological force fields.

The force fields applied for the systems presented in this work all adopted the LJ potential for the vdW description.

Different descriptions of Electrostatic interactions through point-charge models

The electrostatic interaction based on the Coulomb interaction between two charges q_i and q_j is expressed as:

$$V_{i,j,Coulomb} = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}, \quad (1.27)$$

where r_{ij} is the distance between the two charges and ϵ_0 is the permittivity of vacuum. The atomic point charges that appear in the calculation of electrostatic interactions can be obtained through different methods ranging from ab initio, semi-empirical, to empirical approaches. One example of the ab initio approach is to carry out a large number of ab initio calculations and fit these to a potential function including electrostatic and van der Waals interactions as championed by the Clementi group [86]. To the other end, empirically, one can vary the charges to fit the properties of crystals or liquids. To give an example, the nonbonded parameters in the OPLS force field [87] are derived by fitting the enthalpy of vaporization and density of liquids determined by Monte Carlo calculations.

The ESP (electrostatic potential) method [88–92] has been widely used for the ones choosing to derive atomic point charges using ab initio calculations. The principle is to fit the charges on the atoms of a given molecular system to reproduce the electrostatic potential calculated at a large number of grid points around the molecule (outside the van der Waals surface) obtained from ab initio calculations. The advantages of the ESP derived charges are that they reproduce well the quantum mechanically determined multipole moments and optimally reproduce the intermolecular interactions with surrounding molecules, which are essential for simulating complex condensed systems. Like in other ab initio based approaches, the accuracy of the representation of the atomic charges derived by the ESP method depend on the choice of basis sets. The 6-31G* basis set has been proved to be an excellent choice for optimal reproduction of biomolecular properties in aqueous solution [93], which are suitable for our systems of investigations.

The main drawback of this method is that the ESP derived charges are less well suited for intramolecular properties and molecular conformational analyses while providing good descriptions for intermolecular properties. The reason for this drawback is due to the statistical nature of the fitting process as detailed in ref. [93]. In short, in a molecule, the more buried the atom is (i.e. far away from the van der Waals surface of the molecule), the more poorly determined the atom charge is. This makes the ESP derived charges highly conformational dependent, since one atom of a molecule could be buried in one conformer where the atom charge is thus poorly determined, while being exposed to the vdW surface of the molecule in another conformer, thus with well determined atom charge.

To overcome the conformational dependency of the ESP method, the works of Reynolds *et al.* [94] have calculated the ESP charges of each conformer, and fitted the charge for each site using the calculated ESP charges with a Boltzmann weighting according to the relative energy of each conformer. This approach successfully reduces the conformational dependency of the ESP derived charges while at the same time highly increases the computational costs of the parameterization of the charges in a Force Field.

An alternative and more pragmatic approach to minimize the artifacts generated during the fitting procedure of ESP method has been presented by Bayly *et al.* [93]. In ref. [93], an improved version of the ESP method, the restrained electrostatic potential (RESP) method is introduced, by restraining the ill-behaved charges to target values. The amber force field derives the atomic charges through this method. This amber force field using RESP derived atomic charges has been adopted for some of the investigated systems in this thesis, e.g. SAM polymer chains, which will be introduced in the following chapters.

1.5.3 Force fields of different classes

Force fields of different generations compute the intramolecular and intermolecular interactions with different expressions of the intra- and intermolecular interactions and with different parameters/parameterization strategies. They therefore have various accuracy. They are often divided into three classes, that are summarized below.

Class I force fields

The Class I force fields, e.g. the well-known CHARMM [95], AMBER [96], OPLS [87], and GROMOS [97] for biomolecules and polymers, are based on eqs. 1.23, 1.25 and 1.27 for the FF.

The functional form of the class I force fields is thus expressed as:

$$\begin{aligned}
 V = & \sum_{\text{bonds (i,j)}} \frac{1}{2} k_r (r_{ij} - r_0)^2 + \sum_{\text{angles (i,j)}} \frac{1}{2} k_\theta (\theta_{ijk} - \theta_0)^2 \\
 & + \sum_{\text{torsions (i,j,k,l)}} \sum_n k_{\phi,n} [\cos(n\phi_{ijkl} + \delta_n) + 1] \\
 & + \sum_{\text{non-bonded pairs (i,j)}} \left[\frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right],
 \end{aligned} \tag{1.28}$$

Class II force fields

In class II force fields, anharmonic terms (e.g. Morse potential) are usually used to describe the bond stretching potential energy and the cross

terms (e.g. bond–bond, bond–angle, angle–angle, bond–torsion, and angle–angle–torsion) are often taken into account in the intramolecular interaction functions. Some of the force fields of class II (e.g. MM2 [84] and MM3 [85]) replace the Lennard-Jones potential (LJ) by the Buckingham potential in eq. 1.26.

Class III force fields

The next generation, class III force fields, jump to the next accuracy level for the representations of the electrostatic interactions between molecules. Instead of using the point charges derived either *ab initio* or empirically (as detailed in section 1.5.2), a multipole expansion (e.g. dipole-dipole interaction, dipole-quadrupole interactions, quadrupole-quadrupole interactions) is included in this new generation force field for the description of electrostatic interactions. The distributed atomic multipoles can be extracted from the quantum wave function [98] or similarly to the point-charge methods by fitting to the electrostatic potential (ESP) [99, 100]. Moreover, the polarization term is included in the class III force fields, which can be represented based on three different models: induced dipoles model, the Drude model or the fluctuating charges model. To give an example, the AMOEBA force field [101] includes the distributed multipoles and the atom polarizabilities within the Thole model of induced dipole moments. This class also includes coarse-grained force fields, such as MARTINI [102] which is often used to model biological molecules. Moreover, force fields specifically designed for chemical reactions such as REAXFF [103] are also included in the class III force fields family.

1.5.4 The GROMACS package for CMD/FF-MD and the force fields chosen for the studied systems in this thesis

Based on a collaboration with Prof. Predota and Dr. Ondrej Kroutil of the University of South Bohemia in Czech Republic, all the classical MD simulations presented in this manuscript have been carried out using the GROMACS package version 18.7. The GROMACS package was originally developed in the Biophysical Chemistry department of University of Groningen in 1991, and is now maintained by worldwide contributors in research centers and in Universities [104, 105]. GROMACS can be run on both central processing units (CPUs) and graphics processing units (GPUs), and it is one of the fastest software packages of molecular dynamics simulations.

To run classical simulations using the GROMACS package, the following files are needed: the configuration file of the system in form of **.gro* or **.pdb*, the topology file (**.top*) where the molecules of the system are listed and force field files are included, the MD parameter file (**.mdp*) containing all

the parameters of the molecular dynamics. It should be noted that only one final input file (**.tpr*) combining all these information should be passed to GROMACS to run the simulation.

We use the following *grompp* command to compile the necessary files to the final input file:

```
gmx grompp -c *.gro -f *.mdp -p *.top -o *.tpr
```

After generating the final input file (**.tpr*), we pass it through the *mdrun* command to run FF-MD simulation, as:

```
gmx mdrun *.tpr
```

The trajectory files in form of **.trr* or **.xtc* are generated together with the energy file (**.edf*), which could be used to plot the energy or temperature evolution of the system during the simulation.

In this manuscript, we have carried out the classical simulations of a series of Self-Assembled Monolayer (SAM)-water interfaces with SAM polymer chains anchored on a silica substrate. The choices of force fields for the different parts of the system are based on the set up of ref. [106].

The CLAY force field (see ref. [107] for the details), which was shown to reproduce accurately the structure of water/quartz interfaces [108], has been adopted for the silica substrate of our system on which the SAMs are anchored. The CLAY-FF can be expressed by eq. 1.28.

SAM polymers were described using the same set-up as in ref. [106] with the AMBER force field [96], a class I force field expressed by eq. 1.28, coupled with the RESP method [93] (detailed in section 1.5.2) for the partial charges determination.

We used the SPC-Martini model [79], a flexible SPC based model demonstrated to provide good description for the molecular vibrations [81] thanks to the implementation of bond-bond, bond-angle, and angle-angle cross terms in the intramolecular part of the force field (detailed in 1.5.1) to describe the water slab of our system in contact with the SAMs. For all the systems, the Lorentz-Berthelot combining rules were adopted for interactions between unlike nonbonded atoms.

1.5.5 Details of FF-MD simulations of the PEG/water interfaces

For the PEG/water system, an equilibration run of 10 ns in the NVT ensemble has been carried out for all FF-MD trajectories, followed by an NVT production run (1 ns or 80 ns, see table 3.1 for details), both using a Nose-Hoover thermostat and with a time-step of 0.1 fs. The coordinates and velocities are saved at different frequencies depending on the length of the production runs

to save storage (see table 1.1 for details). Two box-sizes were chosen in the FF-MD simulations: $13 \times 13 \times 85 \text{ \AA}^3$ (same as the DFT-MD system, denoted 1 \times 1 in the following sections) and $40 \times 40 \times 300 \text{ \AA}^3$ (9 times the DFT-MD size, denoted 3 \times 3). Periodic boundary conditions have been applied in all three spatial directions of space.

Table 1.1: List of FF-MD simulations for SAM/water interfaces

No.	Size (\AA^3)	Length (ns)	Saved each (ps)	Ensemble
1	$13 \times 13 \times 85$	1	1	NVT
2	$13 \times 13 \times 85$	80	50	NVT
3	$40 \times 40 \times 300$	1	1	NVT
4	$40 \times 40 \times 300$	80	50	NVT

The FF-MD simulations with the same box size as that of the DFT-MDs allow us to capture dynamic properties of our interfaces at much longer time scales. Larger simulation boxes allow us to investigate the size effect on the structural properties of the PEG/water interface, as well as a better sampling of surface heterogeneity, as will be discussed in chapter 3, section 3.3.

1.6 Properties calculated based on the MD simulations

After collecting the DFT-MD and CMD/FF-MD trajectories, a series of structural and spectroscopic analyses have been carried out. This will be discussed in the different chapters of this manuscript.

In this manuscript, we first aim to characterize the interfacial water networks adapting to the presence of a surfaces of various hydrophobicity and hydrophilicity, and ultimately aim to develop a new microscopic descriptor to quantify the hydrophobicity and hydrophilicity of the interface.

To achieve this, building from the Fortran 90 codes originally developed by Dr. Simone Pezzotti during his PhD (NNT: 2019SACLE008), I have modified these codes and developed a series of new codes in both Fortran 90 and Python to investigate the properties of the interfaces of interests. After combining the modified codes of Dr. Pezzotti and my codes, we obtained a set of automatic codes to fully analyze the structure of interfaces for both the water and the surface. The set of codes has been applied to the trajectories of SAM/water interfaces described in chapter 3

The set of codes consist of four main parts:

- Generation of the Willard and Chandler (WC) instantaneous interface [109] at each time-step of the trajectory.
- Plot of the water density as a function of the vertical distances from the WC interface along the z direction, which is further used to define the different water layers (see ref. [32] and chapter 2, section 2.4 for the details about the layers separation).
- Calculation of horizontal and vertical H-bonds density ($/\text{nm}^2$) formed within the topmost Binding Interfacial Layer (BIL), that unveils the interfacial water network at different aqueous interfaces.
- Density maps of the atoms (of both the top surface and water) at the interface (within 3 Å to the WC interface), that helps us to understand the topology and morphology of the interface.

Based on the previous understandings of the interface hydrophobicity/hydrophilicity achieved in the group [26, 27, 32, 110] and the results extracted from our codes applied on the SAM/water interfaces of various hydrophilicity, we developed in this manuscript a new microscopic descriptor for hydrophilicity, the H/V descriptor. The development and the perspectives of this descriptor will be detailed in Chapter 3.

Despite the fact that numerous microscopic descriptors for hydrophobicity have been developed using the MD simulations in the past years [8, 30, 111], our descriptor is the first one directly connected to the spectroscopic quantities (detailed in Chapter 3), which naturally leads to our next topic, the surface specific non-linear optical technique: sum frequency generation (SFG) spectroscopy.

The SFG spectra of a set of aqueous interfaces were calculated based on the MD simulations, ranging from organic (SAM/water interfaces) to oxides surfaces (amorphous silica/water interfaces). In the next chapter, the principles of the SFG spectroscopy and different approaches of theoretical SFG calculation will be presented in chapter 2.

Chapter 2

Nonlinear optical vibrational spectroscopy

It is highly challenging to reveal interfacial water structures by conventional techniques, such as X-ray, neutron scattering and absorption [112]. Since the number of water molecules in the Binding Interfacial Layer (BIL) directly in contact with the solid surface is much smaller than the number of bulk water, the signals measured by these conventional techniques are hence hidden by that from the (numerous/dominant) bulk water. Therefore, surface specific techniques which directly probe interfaces, such as Sum Frequency Generation (SFG) spectroscopy, are necessary for surface characterization.

SFG spectroscopy, first reported by Shen *et al.* in 1987 [14, 15], is a powerful tool for interface characterization. As a non-linear vibrational spectroscopy technique, SFG depends on the second-order susceptibility $\chi^{(2)}(\omega)$, which is zero in centrosymmetric media (e.g. bulk of liquids and many solids). Thanks to its surface specificity, SFG has been widely utilized to investigate aqueous interfaces to unravel their structural, vibrational and dynamical properties on a molecular level [16–23]. Moreover, the coupling between molecular dynamics and theoretical SFG spectra calculation helped the community to gain more insights on dissecting SFG fingerprints and relate them to molecular organizations at the interface, i.e. in the BIL [33, 43, 50, 113, 114].

In this chapter, we will first briefly introduce the principles of both experimental and theoretical SFG spectroscopy, and we will present the new theoretical developments we made to fill the gap in the current techniques. These new developments led to the “pop model” whose applications are presented in chapter 4.

2.1 Generalities on Sum Frequency Generation (SFG) spectroscopy

In SFG spectroscopy, two lasers of frequencies ω_1 and ω_2 are overlapped in time and space at an interface between two media, and the signal of the sum frequency is detected at $\omega_{SFG} = \omega_1 + \omega_2$, as illustrated by fig. 2.1. One fixes ω_1 at a non resonant visible frequency ω_{VIS} and one scans ω_2 through a resonant infrared (IR) frequency ($\omega_2 = \omega_{IR}$), the sum frequency signal ($\omega_{SFG} = \omega_{VIS} + \omega_{IR}$) thus provides a vibrational spectrum of the interface as a function of ω_{IR} . In particular, by scanning the IR frequency in the range of resonance of O-H stretching motions (2800-4000 cm^{-1}), the vibrational response of the H-Bond network can be probed at aqueous interfaces. The hence recorded SFG intensity (I_{SFG}) is proportional to the square of the second order dipole susceptibility $|\chi^{(2)}(\omega)|^2$, where $\chi^{(2)}(\omega)$ is a third-order tensor.

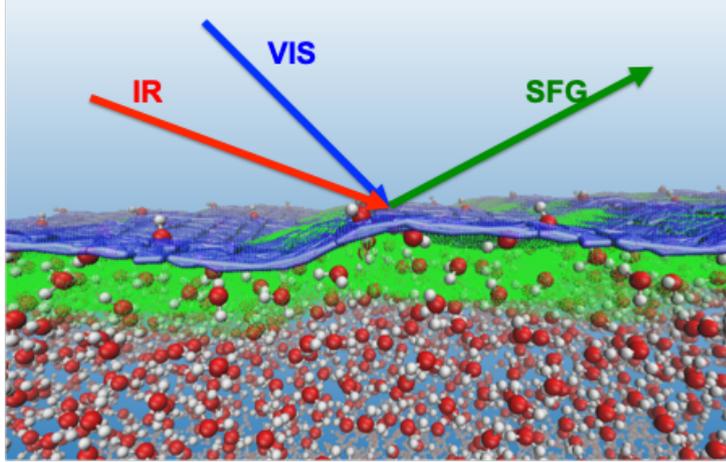


Figure 2.1: Scheme of the principle of the sum frequency generation (SFG) spectroscopy.

If the interface is assumed to be azimuthal symmetric, i.e. symmetric around a straight line, only 7 of the 27 tensor elements are nonzero, i.e. $\chi_{xxz}^{(2)}(\omega) = \chi_{yyz}^{(2)}(\omega)$, $\chi_{xzx}^{(2)}(\omega) = \chi_{yzy}^{(2)}(\omega)$, $\chi_{zxx}^{(2)}(\omega) = \chi_{zyy}^{(2)}(\omega)$ and $\chi_{zzz}^{(2)}(\omega)$, where xyz are the directions in the laboratory frame, z identifies the direction perpendicular to the surface. Each tensor element can be probed with different polarizations of SFG, visible, and infrared fields in SFG experiments. Each propagating light in the experiment can take either s or p polarization, where p denotes polarization of the electric field parallel to the plane of incidence and s denotes polarization perpendicular to it. The polarization combination of the SFG process is usually presented with three letters, such as ssp or ppp , denoting (in this order) the polarizations of the outgoing SFG, incoming visible, and incoming infrared lights respectively. A scheme of the very common ssp combination is illustrated in fig. 2.2.

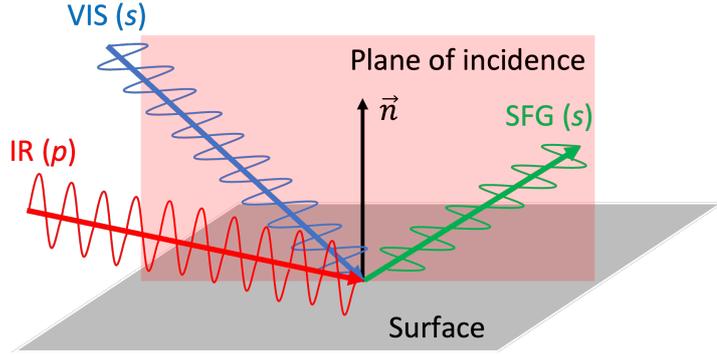


Figure 2.2: Scheme of the *ssp* combination for SFG, where the outgoing SFG (green) and the incoming visible take *s* polarization, i.e. perpendicular to the incidence plane, and the incoming IR takes the *p* polarization, i.e. parallel to the incidence plane.

Taking the *ssp* case as an example, the SFG intensity is proportional to the square of *xzx* (or *yyz*) tensor element of $\chi^{(2)}(\omega)$, so that:

$$I_{ssp} \propto |\chi_{xxz}^{(2)}(\omega)|^2. \quad (2.1)$$

It should be noted that the $\chi^{(2)}(\omega)$ tensor is zero for centrosymmetric media, e.g. the bulk of water and of many solids, which makes it a powerful tool to probe the interfacial water and solid surface structures, where the centrosymmetry is broken along the *z*-direction.

The non-linear susceptibility $\chi^{(2)}(\omega)$ can be decomposed into a vibrationally resonant part $\chi_R^{(2)}(\omega)$ and a non-resonant part $\chi_{NR}^{(2)}(\omega)$:

$$\chi^{(2)}(\omega) = \chi_R^{(2)}(\omega) + \chi_{NR}^{(2)}, \quad (2.2)$$

where the resonant part is mainly responsible for the dependence of $\chi^{(2)}(\omega)$ on the infrared frequency, while the non-resonant part can be assumed as a constant independent on ω_{IR} , hence only providing a shift in intensities of the SFG signal. The $\chi_R^{(2)}(\omega)$ is a complex quantity with a real part ($\Re(\chi^{(2)}(\omega))$) and an imaginary part ($\Im(\chi^{(2)}(\omega))$).

A $|\chi^{(2)}(\omega)|^2$ spectrum measured by an SFG experimental set-up provides direct information on the vibrational motions of interfacial water and on the vibrations of the surface of the solid interfaced with water. The initial groundbreaking $|\chi^{(2)}(\omega)|^2$ experimental homodyne spectrum was obtained by Shen and coworkers more than 20 years ago [115], while the first heterodyne detected HD-SFG spectrum was published by the same group almost 10 years ago [116]. In HD-SFG, it is $\Re(\chi^{(2)}(\omega))$ and $\Im(\chi^{(2)}(\omega))$ that are directly measured instead of $|\chi^{(2)}(\omega)|^2$. This gives the knowledge of a sign on $\chi^{(2)}(\omega)$ and therefore knowledge on the orientation of the O-H groups measured in the 2000–4000 cm^{-1} vibrational range.

Since then, the HD-SFG set-up has spread over numerous laboratories worldwide, including the Tahara's group in Japan, the Geiger's group in U.S. and the Bonn's group in Germany. The information on the orientation of

interfacial OH-groups contributing to the SFG signal is obtained in HD-SFG spectra [117, 118], by taking advantage of the interferences between SFG beams emitted by the sample and by a reference sample in order to separate the Real $\Re(\chi^{(2)}(\omega))$ and Imaginary $\Im(\chi^{(2)}(\omega))$ components of $\chi^{(2)}(\omega)$. The Imaginary $\Im(\chi^{(2)}(\omega))$ component is the relevant one for the interpretation of HD-SFG spectra. Note that both $\Re(\chi^{(2)}(\omega))$ and $\Im(\chi^{(2)}(\omega))$ parts can have positive and negative intensities. One further advantage of HD-SFG with respect to conventional SFG is that since $\chi_{NR}^{(2)}$ can be considered as a constant, it disappears in the interference process, so that only the resonant part of the second order susceptibility is present in HD-SFG spectra.

2.2 Computation of SFG vibrational spectra: DFT-MD vs FF-MD

Due to the combined information contained in a SFG spectrum, i.e. vibrational motions (ω values) and the orientation of the interfacial molecules (sign of the signal), it is often difficult to dissect the experimental SFG spectra without a theoretical model. The resonant component of the second order susceptibility $\chi_R^{(2)}(\omega)$ can be theoretically calculated from MD simulations by the Fourier transform of the correlation function of the total dipole moment (\overline{M}) and polarizability tensor (\overline{A}) of the system, following the time-dependent method introduced by Morita *et al.* [34, 35]:

$$\chi_{PQR}^{(2)}(\omega) = \frac{i\omega}{k_B T} \int_0^\infty dt \exp(i\omega t) \langle A_{PQ}(t) M_R(0) \rangle, \quad (2.3)$$

where (P,Q,R) are any x, y, z direction in the laboratory frame, and k_B and T are respectively the Boltzmann constant and temperature of the simulated system. $\langle \dots \rangle$ is a time-correlation function, ω is the frequency, $A_{PQ}(t)$ and $M_R(0)$ are the PQ component of the total polarizability tensor \overline{A} and the R component of the total dipole moment \overline{M} of the system respectively. Here and in the following text, the $\chi^{(2)}(\omega)$ refers to its resonant part for simplicity of notation. For the *ssp* polarization (i.e. PQR = xxz), for example, we have:

$$\chi_{ssp}^{(2)}(\omega) = \chi_{xxz}^{(2)}(\omega) = \frac{i\omega}{k_B T} \int_0^\infty dt \exp(i\omega t) \langle A_{xx}(t) M_z(0) \rangle \quad (2.4)$$

$$= \chi_{yyz}^{(2)}(\omega) = \frac{i\omega}{k_B T} \int_0^\infty dt \exp(i\omega t) \langle A_{yy}(t) M_z(0) \rangle \quad (2.5)$$

$\Im(\chi^{(2)}(\omega))$ and $\Re(\chi^{(2)}(\omega))$ are separately calculated. As inferred by the equations, the sign of the xxz element of the $\chi^{(2)}(\omega)$ tensor in eq. 2.4 solely depends on $M_z(0)$, which depends on the orientation of the SFG-active oscillators (e.g. O-H groups of water and of silanols SiOH at silica surfaces in the OH-stretching frequency range). While all the diagonal elements of the polarizability tensor (A_{xx} , A_{yy} , A_{zz}) are positive, including the $A_{xx}=A_{yy}$

element contributing to $\chi_{xxz}^{(2)}(\omega)$ in eq. 2.4, the M_z component of the dipole moment (along the z direction) can indeed be either positive or negative depending on the orientation of the interfacial water molecules (i.e. towards or outwards the surface). Adopting the convention used in the SFG community and reported in the following scheme (fig. 2.3), if interfacial water has a net



Figure 2.3: Convention in SFG: \vec{n} normal vector being directed from the aqueous phase to the other phase.

"up" orientation (i.e. water dipole pointing along the normal to the surface, normal vector being directed from the aqueous phase to the other phase), then M_z is positive and $\Im(\chi^{(2)}(\omega))$ is also positive, while if interfacial water has the opposite "down" orientation, M_z and hence $\Im(\chi^{(2)}(\omega))$ have negative sign.

The reverse is true for $\Re(\chi^{(2)}(\omega))$. The same conclusion is straightforwardly reached also for $\chi_{zzz}^{(2)}(\omega)$ (given by the $\langle A_{zz}(t)M_z(0) \rangle$ correlation function) probed by SFG in the *ppp* polarization, which is the most commonly used together with the *ssp* polarization (probing $\chi_{xxz}^{(2)}(\omega)$ and $\chi_{yyz}^{(2)}(\omega)$).

However, to reveal the structure-spectroscopy relationship it is desirable to evaluate the contribution of specific interfacial water populations to the total measured SFG spectra, i.e. a class of interfacial water molecules with distinct properties with respect to the other molecules, as for example the populations composed by water donating or accepting HBs with a solid surface. In order to achieve this, it is convenient to rewrite eq. 2.4 as the sum of the contributions from each water molecule, which leads to:

$$\chi_{xxz}^{(2)}(\omega) = \frac{i\omega}{k_B T} \sum_{i=1}^{N_W} \sum_{j=1}^{N_W} \int_0^\infty dt \exp(i\omega t) \langle \alpha_{xx}^j(t) \mu_z^i(0) \rangle, \quad (2.6)$$

where $\mu_z^i(0)$ and $\alpha_{xx}^j(t)$ now refer to the z and xx components of the *individual* dipole and polarizability of the i -th and j -th water molecules respectively, and N_W is the total number of water molecules in the simulation.

The above equation can be further rewritten by separating the self-correlation terms, i.e. the terms with $i = j$, from the cross-correlation terms, i.e. the terms with $i \neq j$:

$$\chi_{xxz}^{(2)}(\omega) = \frac{i\omega}{k_B T} \int_0^\infty dt \exp(i\omega t) \sum_{i=1}^{N_W} \left(\langle \alpha_{xx}^i(t) \mu_z^i(0) \rangle + \sum_{j \neq i}^{N_W} \langle \alpha_{xx}^j(t) \mu_z^i(0) \rangle \right), \quad (2.7)$$

where the cross-correlation terms (second part in the integral) are related to the inter-molecular vibrational couplings between the water molecules. These are zero for non-interacting molecules, while one expects these terms to be relevant and non zero for liquid water molecules, as they are all connected through the HB-network.

The great advantage of this formulation is that once the SFG signal is expressed in terms of molecular contributions, one can easily calculate the spectral contributions of a selected water population by just restricting the sum over i to the water molecules belonging to one targeted population only (instead of all the N_w water molecules in the system).

To compute such theoretical spectrum, it is however needed to calculate the molecular components $\mu_z^i(0)$ and $\alpha_{xx}^j(t)$ at each time t of the MD trajectory. They can be directly obtained from ab-initio simulations (like in DFT-MD), from localization and integration of the explicitly treated electron density [119], or from classical MD simulations with flexible and polarizable Force-Fields (FF) [37–39, 120].

Though being precise, the DFT-MD calculated SFG spectra through localization of the electron density is highly computationally costly, which limits the dimension of the simulated systems to hundreds of atoms and the simulation time-scales to tens of ps. This has been demonstrated in ref. [121] in a pioneering work.

We note that the cross-correlation terms in eq. 2.7 are known to converge within hundreds of ps [119], which exceeds the limit of DFT-MD simulation time-scales. Therefore, theoretical SFG spectra are often calculated from DFT-MD considering only the self-correlation term in eq. 2.7 (shown here for the ssp signal):

$$\chi_{xxz}^{(2)}(\omega) = \frac{i\omega}{k_B T} \int_0^\infty dt \exp(i\omega t) \sum_{i=1}^{N_w} \langle \alpha_{xx}^i(t) \mu_z^i(0) \rangle. \quad (2.8)$$

The classical FF-MD simulations with flexible and polarizable FFs, on the contrary, allow to significantly reduce the computational cost of the MD simulations, hence making it possible to investigate bigger systems with longer trajectories. However, the complex design of each system-specific FF and the limited transferability of the FFs from one system to another make the spectra calculation complicated. It is indeed well known that FFs are not precise enough for vibrational spectroscopy in general, and a specific FFs should be designed to that end [122, 123].

Another crucial limitation of FF-MD is that no chemistry reactions are simulated, so that the structural reorganization of the aqueous interfaces due to the surface chemistry, such as adsorption/desorption processes and proton transfers, can not be investigated.

Despite the relative limitations, both theoretical SFG spectra calculated

from DFT-MD and classical MD have been shown to be able to reproduce the experimental data with, respectively, very good agreement (DFT-MD) [27, 33, 40, 43] and reasonable agreement (FF-MD) [37, 38, 120, 124, 125]. Both methods will be used in this work.

2.3 Our approach to calculate SFG spectra from DFT-MD simulations

The original idea of the methodology presented in this section and used in my thesis works is to overcome the major limitations of theoretical SFG calculated from DFT-MD simulations, i.e. the high computational cost of extracting ab initio dipole moments and polarizability tensors for each molecule of the system at each time-step, and to go beyond the slow convergence of the dipole-polarizability correlation function. To achieve this, eq. 2.8 is rewritten in terms of the velocity-velocity correlation function (VVCF), which is trivial to calculate, and at low computational cost, as velocities are obtained as standard output from MD simulations, and fast to converge (few ps typically).

Let us start by introducing the time derivatives of $\mu_R^i(0)$ and $\alpha_{PQ}^i(t)$ (where (P,Q,R) are the directions in the laboratory frame) in eq. 2.8 (a time derivative introduces a $1/\omega$ factor in the Fourier Transform), so that (the reason for the time derivatives will become clear in a few lines):

$$\chi_{PQR}^{(2)}(\omega) = \frac{i}{k_B T \omega} \sum_{i=1}^{N_w} \int_0^\infty dt \exp(i\omega t) \langle \dot{\alpha}_{PQ}^i(t) \dot{\mu}_R^i(0) \rangle. \quad (2.9)$$

As shown in ref. [36], supposing that at the frequencies of interest only the OH stretching motions are contributing (which is more than reasonable in the 2000–4000 cm^{-1} spectral range), the dipole moment and polarizability tensors of each water molecule can be decomposed at each time t along the trajectory into individual O-H bond contributions ($\dot{\alpha}_{n,PQ}^i(t)$ and $\dot{\mu}_{n,R}^i(t)$), with the following equations:

$$\dot{\alpha}_{PQ}^i(t) = \sum_{n=1}^N \dot{\alpha}_{n,PQ}^i(t), \quad (2.10)$$

$$\dot{\mu}_R^i(t) = \sum_{n=1}^N \dot{\mu}_{n,R}^i(t), \quad (2.11)$$

where n identifies an O-H oscillator and N is the total number of O-H oscillators per molecule (i.e. $N = 2$ for water).

We use the rotational matrix D , of elements D_{Pl} , D_{Qm} and D_{Rn} , to project the molecular frame (x, y, z) , where z corresponds to the O-H bond axis, onto the laboratory frame (P, Q, R) , so that D_{Rl} is the cosine between the l -th

coordinate in the molecular frame and the R coordinate in the laboratory frame. An example of D_{Rz} is shown in Fig. 2.4.

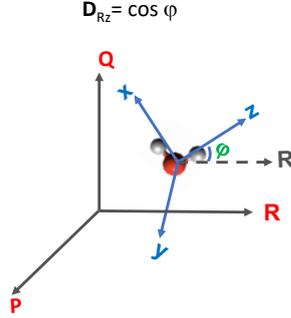


Figure 2.4: Scheme of an example of D matrix element: D_{Rz} .

Assuming that the O-H stretching is much faster than the modes involving a bond reorientation, for example the libration, we thus have $\frac{dr_z}{dt} = v_z(t) \gg \frac{dr_x}{dt} = v_x(t) \approx \frac{dr_y}{dt} = v_y(t)$, i.e. $v_z \gg v_x \approx v_y$. Based on this, the following expressions for molecular dipole moment and polarizability tensor components are obtained by omitting the terms with v_x and v_y :

$$\dot{\alpha}_{PQ}^i(t) \simeq \sum_l^{x,y,z} \sum_m^{x,y,z} D_{Pl}^i(t) D_{Qm}^i(t) \frac{d\alpha_{lm}^i(t)}{dr_z} v_z^i(t), \quad (2.12)$$

$$\mu_R^i(t) \simeq \sum_n^{x,y,z} D_{Rn}^i(t) \frac{d\mu_n^i(t)}{dr_z} v_z^i(t), \quad (2.13)$$

where of course we have written $\frac{d\mu_n}{dt} = \frac{d\mu_n}{dr_z} \frac{dr_z}{dt}$. Moreover, $v_z(t)$ is the projection of the OH bond velocity onto the z axis.

Substituting eqs. 2.12 and 2.13 into eq. 2.9, the final equation to calculate the theoretical SFG spectrum using this approach is thus:

$$\chi_{PQR}^{(2)}(\omega) = \frac{i}{k_B T \omega} \sum_{i=1}^{N_W} \int_0^\infty dt \exp(i\omega t) \sum_l^{x,y,z} \sum_m^{x,y,z} \sum_n^{x,y,z} D_{Pl}^i(t) D_{Qm}^i(t) D_{Rn}^i(0) \left\langle \frac{d\alpha_{lm}^i(t)}{dr_z} v_z^i(t) \frac{d\mu_n^i(0)}{dr_z} v_z^i(0) \right\rangle. \quad (2.14)$$

Furthermore, $\frac{d\alpha_{lm}^i(t)}{dr_z}$ and $\frac{d\mu_n^i(0)}{dr_z}$ can be shown to vary on a far longer time scale than the velocity terms $v_z^i(t)$. This has been shown in refs. [36, 126]. Therefore, eq. 2.14 becomes:

$$\chi_{PQR}^{(2)}(\omega) = \frac{i}{k_B T \omega} \sum_{i=1}^{N_W} \int_0^\infty dt \exp(i\omega t) \sum_l^{x,y,z} \sum_m^{x,y,z} \sum_n^{x,y,z} D_{Pl}^i(t) D_{Qm}^i(t) D_{Rn}^i(0) \frac{d\alpha_{lm}^i}{dr_z} \frac{d\mu_n^i}{dr_z} \langle v_z^i(t) v_z^i(0) \rangle. \quad (2.15)$$

The D matrix and the projection of the velocities on the O-H bond axis (v_z) can be directly obtained from the DFT-MD trajectory, while $\frac{d\alpha_{lm}}{dr_z}$ $\frac{d\mu_n}{dr_z}$ can be parameterized. Parameterization from refs. [36, 127], as shown in table 2.1, has been systematically applied to water to calculate all theoretical SFG spectra presented in this manuscript.

Note also that with the same approach, also the contribution from the O-H groups of (solid) surfaces in contact with water can be readily obtained, with the only additional requirement of a new parametrization for the given surface, as shown and detailed in refs. [27, 110].

Table 2.1: Parameterized derivatives of the dipole moment and polarizability of the O-H bond of water and SiOH from ref [36, 127] for water and ref [27, 110] for SiOH.

	$\frac{\partial\mu_x}{dr_z}$	$\frac{\partial\mu_y}{dr_z}$	$\frac{\partial\mu_z}{dr_z}$	$\frac{\partial\alpha_{xx}}{dr_z}$	$\frac{\partial\alpha_{yy}}{dr_z}$	$\frac{\partial\alpha_{zz}}{dr_z}$	$\frac{\partial\alpha_{xy}}{dr_z}$	$\frac{\partial\alpha_{xz}}{dr_z}$	$\frac{\partial\alpha_{yz}}{dr_z}$
H ₂ O	-0.15	0.0	2.1	0.40	0.53	1.56	0.0	0.02	0.0
(Si)OH	0.005	0.402	3.589	0.630	0.388	3.351	-0.105	0.118	0.296

It is finally important to stress here that the parameterization of $\frac{d\alpha_{lm}}{dr_z}$ and $\frac{d\mu_n}{dr_z}$ is made by evaluating the variation of the total polarizability (\bar{A}) and dipole moment (\bar{M}) of the system induced by the displacement along a given direction (in the molecular frame) of an O-H covalent bond of one water molecule, averaged over all waters in the simulation. Thus this naturally accounts for the water-water intermolecular couplings. This parameterization strategy not only takes into account the N-body polarization effects acting on all water dipoles and polarizabilities due to the water molecule undergoing the O-H motion (i.e. inter-molecular couplings), as the wavefunction of the modified system is systematically recalculated, but it also includes charge fluxes in between water molecules (which is highly important).

Within the dipole derivatives used for the parameterization, one has indeed (in the following, i labels the elongated/contracted water molecule, μ is the total dipole moment vector of the system, $\mu = \sum_i \mu_i$ with μ_i the dipole of molecule i): $\frac{d\mu}{dr_z^i} = \frac{d\mu_i}{dr_z^i} + \sum_{j \neq i} \frac{d\mu_j}{dr_z^i}$ where $\frac{d\mu_j}{dr_z^i}$ are charge fluxes in between molecules i and j [128]. The same is true for the polarizability derivatives. With charge fluxes hence included in the parameterized derivatives, inter-molecular couplings and cross-correlation terms are also included in a ‘mean field approach’, even though we formally calculate only self-correlation terms for the SFG signal in eq. 2.14.

The velocities also take into account intra- and inter-molecular couplings from the dynamics, hence again introducing cross-correlation terms in a ‘mean field approach’ into the SFG self-correlation function of eq. 2.14.

The present methodology to calculate theoretical SFG spectra from the weighted velocity-velocity correlation function (VVCF) has been shown to provide excellent agreement with respect to experimentally measured HD-SFG spectra on a vast range of water-solid and water-vapor interfaces

at different pH, electrolytes and surface hydrophilicity conditions [27, 32, 33, 129–131].

2.4 SFG spectroscopy of charged-interfaces: Deconvolution of $\chi_{BIL}^{(2)}(\omega)$ and $\chi_{DL}^{(2)}(\omega)$

2.4.1 Origins of $\chi_{BIL}^{(2)}(\omega)$ and $\chi_{DL}^{(2)}(\omega)$ contributions

The probing depth of the SFG spectroscopy is determined by the water layer(s) where bulk centrosymmetry is broken, therefore where $\chi^{(2)}(\omega)$ is not equal to zero. The Binding Interfacial Layer (BIL), defined by Tian *et al.* in ref [44], is the layer in direct contact with the surface, which is composed of an interface-specific H-bonding network (between the surface and the water and between the water molecules) dictated by the balance among various interplays, e.g. water-surface and water-water interactions, solvation of ions and adsorption of molecules. This interfacial region is typically the one where electrochemistry occurs at interfaces, and it directly controls many elementary processes, such as the uptake of species from the atmosphere, desolvation/resolvation processes (also called "stripping off hydration shells") of ions and (in)organic molecules and charge transfers in chemical reactions [132, 133]. As has been demonstrated by ref. [33], the BIL is usually composed of one water mono-layer, with a thickness of $\sim 2\text{-}4$ Å, independently of the boundary conditions of the interface (hydrophilicity/hydrophobicity, ionic strength, pH, surface charge...).

Only the BIL is probed by SFG spectroscopy in case of neutral interfaces, whereas a much thicker water slab is SFG active when the interfaces are charged [44, 134]. At charged interfaces, in addition to the effect of surface-water interactions, the interfacial charge (surface charge plus electrolytes in the BIL) plays an important role in reorienting the bulk water network, thus having non-zero $\chi^{(2)}(\omega)$, which enlarges the probing depth of the SFG spectroscopy.

In the BIL, the H-bond network reorganization due to the specific interactions with the other medium is dominant, so that this layer is defined with the same thickness at both neutral and charged interfaces [27, 32, 33, 130, 131]. On the contrary, as shown in e.g. refs [27, 33, 44, 129–131, 134], at longer distances from the surface, i.e. in the water layer(s) beyond the BIL, only the field-induced reorientation affects the structure of water in a region named Diffuse Layer (DL).

Both BIL and DL are SFG active. The total SFG signal of a charged interface is thus the sum of BIL and DL contributions (see ref. [44] for the first instance of this expression):

$$\chi^{(2)}(\omega) = \chi_{BIL}^{(2)}(\omega) + \chi_{DL}^{(2)}(\omega), \quad (2.16)$$

and the thickness probed by SFG is the sum of the BIL and DL thicknesses. Centrosymmetric bulk water is recovered beyond the DL (where the surface field does not reorient liquid water anymore).

A scheme for the BIL, DL and bulk layers is shown in Fig. 2.5.

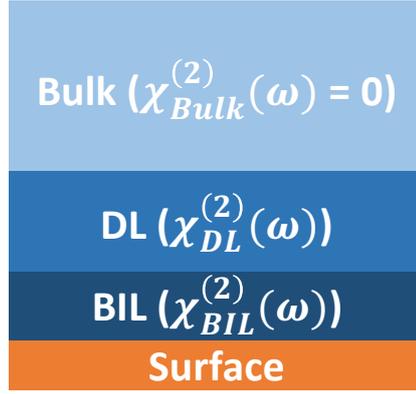


Figure 2.5: Scheme of BIL, DL and bulk water layers.

Although centrosymmetry is broken in both BIL and DL water layers, the microscopic origin of their SFG activity is different, and the distinct water arrangements thus lead to different BIL/DL SFG signatures. BIL-water, in particular, is characterized by a very different H-Bond network from the one found in the bulk (neutral and charged interfaces alike), because of H-Bonds formed with different strengths and orientations, either water-water or water-surface HBs. The $\chi_{BIL}^{(2)}(\omega)$ vibrational response from the BIL layer is thus unique for each interface, and it is the signature of how water adapts to the immediate perturbation from a solid.

The DL on the contrary is reoriented bulk liquid water, and the water-water HB-network in this layer is thus the same as the one in the liquid water, while the net orientation of the HB-network induced by the surface field breaks the centrosymmetry. Dominated by the effect of the surface field, the difference in the DL-SFG signal from various charged interfaces only originates from the value of the static field created by the surface charge (and by interfacial ions). The DL activity to SFG is dictated by the change of the water polarization and orientation in response to three electric fields: the IR and VIS fields applied in SFG experiments, plus the static field at charged interfaces.

As hypothesized by Tian *et al.* in ref. [44], assuming that the field-induced reorientation dominates the SFG signal of the DL, it is possible to express the $\chi^{(2)}(\omega)$ of the DL in terms of a third-order contribution $\chi_{Bulk}^{(3)}(\omega)$ of liquid water, through equation 2.17 [44, 134]:

$$\chi_{DL}^{(2)}(\omega) = \chi_{Bulk}^{(3)}(\omega) \int_a^b dr \cdot \mathbf{E}_{DC}(r) e^{i\Delta k_r r}, \quad (2.17)$$

where a and b are respectively the BIL-DL and DL-bulk z -boundaries (where z identifies the direction perpendicular to the surface), r is the (vertical) distance from the surface, $E_{DC}(r)$ is the field value arising from the surface charge plus electrolytes in the BIL at a given distance from the surface and

2.4 SFG spectroscopy of charged-interfaces: Deconvolution of $\chi_{BIL}^{(2)}(\omega)$ and $\chi_{DL}^{(2)}(\omega)$

$\int_a^b \mathbf{E}_{DC} dr$ is the electrostatic potential across the DL, denoted $\Delta\phi_{DL}$. Δk_r in this equation is a phase factor that takes into account the interferences between the emitted light at different depths from the surface. This factor becomes important only when the BIL+DL thickness exceeds 10 nm, as inferred by Roke's group in ref. [134].

By assuming that water in the DL behaves as a continuum medium and electrolytes are the major contributors to the surface field screening, the thickness over which the static field vanishes is usually estimated by the Debye-length (denoted by κ_D^{-1}), given by the formula:

$$\kappa_D^{-1} = \left(\frac{\epsilon k T}{e^2 \sum_i Z_i^2 n_i^B} \right)^{1/2}$$

where k is Boltzmann's constant, T is the absolute temperature, e is the proton charge, ϵ is the permittivity of the solvent, Z_i and n_i^B are the charge and concentration of ion species i in the bulk, respectively, and the sum extends over all ion species in solution. Following this assumption, interferences modulate the shape and magnitude of DL-SFG spectra, starting only at ionic strength lower than 10^{-3} M (corresponding to Debye-lengths larger than 10 nm) and going downwards (i.e. at lower ions concentrations), while they are negligible for more concentrated electrolytic solutions. [134, 135]

Neglecting the interferences term, the total SFG signal for a charged interface becomes:

$$\chi^{(2)}(\omega) = \chi_{BIL}^{(2)}(\omega) + \chi_{Bulk}^{(3)}(\omega) \cdot \Delta\phi_{DL}, \quad (2.18)$$

where $\Delta\phi_{DL} = \int_a^b dr \cdot \mathbf{E}_{DC}(r)$ is the potential difference across the DL.

2.4.2 Experimental approaches for deconvolving $\chi_{BIL}^{(2)}(\omega)$ and $\chi_{DL}^{(2)}(\omega)$ contributions

Since BIL and DL SFG signals both contribute to the total SFG spectrum of any charged interface, their deconvolution is mandatory for a correct interpretation of the spectral features. In a pioneering work on that subject, a deconvolution scheme has been proposed by Tian *et al.* in ref [44], from experimental data. Ref. [44] proved that the existence of surface charge has a predominant influence on the water molecules located further than a few Å from the surface, while the structure of the topmost layer (BIL) is barely changed at a low surface charge density σ (less than a few percent of a monolayer). Based on this, the difference between two measured SFG spectra of interfaces with two close surface charge values is equal to the difference between the two DL-SFG signals only, since the same $\chi_{BIL}^{(2)}(\omega)$ BIL-SFG spectra cancel each other: $\Delta\chi^{(2)}(\omega) = \Delta\chi_{DL}^{(2)}(\omega)$.

Once the (difference) $\Delta\chi_{DL}^{(2)}(\omega)$ spectrum is calculated, the $\chi_{bulk}^{(3)}(\omega)$ is thus obtained by eq. 2.17, so that: $\Delta\chi_{DL}^{(2)} = \chi_{bulk}^{(3)}(\omega) \cdot \Delta\Delta\phi_{DL}$, where $\Delta\Delta\phi_{DL}$ (which is estimated by the Gouy-Chapman (GC) theory) is the difference between the $\Delta\phi_{DL}$ (as defined in eq. 2.18) for the two surface charge values chosen for the two SFG measurements. The range of applicability of the GC theory hence sets the limits in the application of this deconvolution scheme. The $\chi_{bulk}^{(3)}(\omega)$, equal to $\frac{\Delta\chi_{DL}^{(2)}}{\Delta\Delta\phi_{DL}}$, is thus obtained.

In ref. [44], by hence extracting $\chi_{bulk}^{(3)}(\omega)$ for water at the interface with organic monolayers probed by SFG at various pH and electrolyte concentration conditions, Tian *et al.* have shown that the same $\chi_{bulk}^{(3)}(\omega)$ is obtained for all these different interfaces. Since the water organization in the DL is the one of bulk water at the only exception of the field-induced reorientation of the H-bond network, the $\chi^{(3)}(\omega)$ spectrum of any DL, obtained as $\chi_{DL}^{(2)}(\omega)$ divided by the potential difference inducing such water reorientation, has been thus proposed to be universal and equal to the $\chi^{(3)}(\omega)$ spectrum of bulk water $\chi^{(3)}(\omega) = \chi_{bulk}^{(3)}(\omega)$.

We insist here that, despite the fact that it can be rewritten in terms of the $\chi_{bulk}^{(3)}(\omega)$ third-order contribution, the DL-SFG activity is a pure second-order process, physically dominated by the water reorientation in the DL. [33, 44, 136–138] It is however convenient to rewrite $\chi_{DL}^{(2)}(\omega)$ as in eq. 2.17, since by doing this its dependence on the surface potential is explicitly expressed and the $\chi_{DL}^{(2)}(\omega)$ of any charged interface can be experimentally obtained using the above discussed procedure.

Alternatively, as is for instance shown in ref. [137], the BIL-SFG spectrum of a charged surface can be experimentally determined by performing a set of SFG-measurements at increasing ionic strength, by introducing a growing concentration of a selected electrolyte in the aqueous phase. The introduced ions can be expected to distribute in the liquid water in order to screen the surface charge (e.g. the cations closer to the surface if the latter is negatively charged), i.e. by forming an Electrical Double Layer (EDL). Increasing the ions concentration, the thickness over which the surface potential vanishes is progressively reduced (as is for instance predicted by the Debye-Length estimation commonly used in experiments [134]), up to the point that such thickness becomes lower than the thickness of the BIL layer (for ions concentrations greater than 2M, the Debye-Length is lower than 3 Å thus lower than the BIL thickness). At this stage, the field-induced reorientation would only affect water in the BIL, intrinsically accounted for in $\chi_{BIL}^{(2)}(\omega)$ signal which now would be equal to the total measured SFG-signal, since no $\chi_{DL}^{(2)}(\omega)$ intensity would be recorded anymore (as non SFG-active bulk water would be recovered right after the BIL layer).

This is indeed the strategy used in ref. [137], where for negatively charged silica-water interfaces the ions concentration is increased up to the point that the screening of the cations completely suppresses the DL-SFG signal, while

only the BIL layer still contributes to the measured SFG spectrum.

Despite being able to isolate the BIL-SFG signal via the ions-induced suppression of the DL-SFG signal, this technique has the disadvantage that the so-determined BIL-SFG spectrum is representative only of high electrolytes concentrations and different from the one at low electrolytes concentrations, this latter being out of the reach of this approach. However, this approach does not require passing through the knowledge of the $\chi_{bulk}^{(3)}(\omega)$ spectrum, contrary to the one proposed in ref. [44], since the BIL-SFG signal is not deconvolved from the DL-SFG one, but just obtained for the ionic strength conditions for which a DL layer does not exist anymore. This methodology thus has the advantage of not relying on any estimation of the surface charge and of the surface potential.

2.4.3 Theoretical approach for deconvolving $\chi_{BIL}^{(2)}(\omega)$ and $\chi_{DL}^{(2)}(\omega)$ contributions

A crucial advantage of theoretical SFG spectroscopy of charged interfaces calculated via MD-simulations is that the spectrum arising from a given water population can be isolated from the total signal, by just selecting the species over which the summation in eqs. 2.6 and 2.8 (for the self term only approach) runs. As a consequence, SFG contributions from different water layers, as the BIL and the DL discussed here, can be readily computed once the water molecules belonging to each of these layers have been identified. Therefore, the knowledge of which water molecules belong to each BIL/DL/bulk layer is required.

Pezzotti *et al.* demonstrated that a minimum set of three structural descriptors is sufficient to achieve this goal (see refs. [32, 33] for details). The descriptors include the water density profile along the z -direction perpendicular to the (water) surface, the average water coordination number and the H-Bonds orientation, all evaluated layer by layer at increasing distance from the water surface.

In ref. [33], the methodology to separate BIL-SFG and DL-SFG spectra is detailed and has been successfully applied to a wide range of aqueous interfaces, showing that such deconvolution scheme applies to any aqueous interface. By extracting the third order contribution from the DL-SFG signals of all the considered charged interfaces and by comparing it to the $\chi^{(3)}(\omega)$ spectrum calculated from bulk water under an external static field, our group has shown that the $\chi_{bulk}^{(3)}(\omega)$ spectrum of bulk water is at the heart of the SFG-activity of any DL. The same two-bands structure arising from $\text{Im}\chi_{bulk}^{(3)}(\omega)$ is hence observed in the DL-SFG $\text{Im}\chi_{DL}^{(2)}(\omega)$ spectra of all charged aqueous interfaces, just modulated in sign and magnitude by the potential difference $\Delta\phi_{DL}$ across the DL (as ruled by eq. 2.18).

The $\chi_{bulk}^{(3)}(\omega)$ spectra (imaginary and real components) obtained for all the aqueous interfaces investigated in ref. [33], including liquid water under an

external field, is in great agreement with the $\chi_{bulk}^{(3)}(\omega)$ spectrum extracted from experimental measurements by Tian *et al.* in ref. [44], as shown in ref. [33] and in Fig. 2.6, where the two-bands structure of the $\text{Im}\chi_{bulk}^{(3)}$ spectrum, with two positive bands centered at 3200 and 3400 cm^{-1} , is systematically recovered for all systems. This $\chi_{bulk}^{(3)}(\omega)$ spectrum will be of importance in some of my works reported in chapter 3 for SAMs/water interfaces with charged surfaces. It is discussed also in section 2.5.

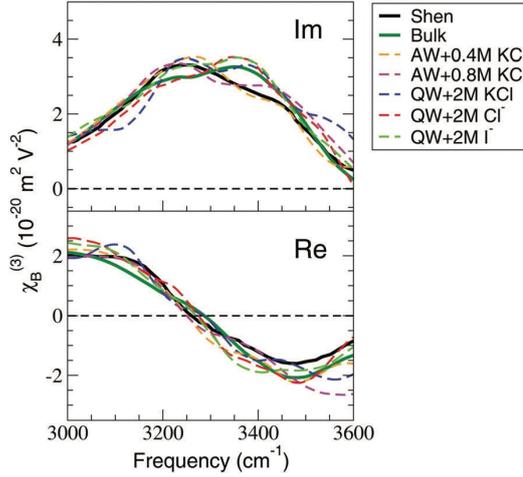


Figure 2.6: A comparison of the $\chi_{Bulk}^{(3)}(\omega)$ spectrum (imaginary and real components) calculated in ref. [33] for different aqueous interfaces, also for bulk liquid water under an applied external field and from Shen's experiments [44].

Moreover, all $\chi_{bulk}^{(3)}(\omega)$ spectra also have the same absolute intensities, in the order of $\sim 10^{-20} \text{ m}^2 \text{ V}^{-2}$, which is one order of magnitude bigger than the common intensity of BIL-SFG spectra ($\sim 10^{-21} \text{ m}^2 \text{ V}^{-1}$) [33, 44, 130, 136]. Since $\chi_{DL}^{(2)}(\omega) = \chi_{bulk}^{(3)}(\omega) \cdot \Delta\phi_{DL}$, this finding implies that $\chi_{DL}^{(2)}(\omega)$ and $\chi_{BIL}^{(2)}(\omega)$ have comparable intensities for $\Delta\phi_{DL} = 0.1\text{V}$, while BIL-SFG dominates for $\Delta\phi_{DL} \leq 0.1\text{V}$ (where $\chi_{BIL}^{(2)}(\omega) = 10 \cdot \chi_{DL}^{(2)}(\omega)$) and DL-SFG dominates for $\Delta\phi_{DL} \geq 0.1\text{V}$ (where $\chi_{DL}^{(2)}(\omega) = 10 \cdot \chi_{BIL}^{(2)}(\omega)$). The mixing of BIL/DL SFG signals in any measured $\text{Im}\chi^{(2)}(\omega)$ (or $\text{Re}\chi^{(2)}(\omega)$) spectrum can be thus simply estimated from the knowledge of the surface potential.

The very good agreement between $\chi_{bulk}^{(3)}(\omega)$ obtained from theoretical and experimental deconvolution schemes and the agreement with the $\chi_{bulk}^{(3)}(\omega)$ spectrum obtained from the simulation of bulk water under an external field, all validate the experimental deconvolution scheme from Tian *et al.* [44] and the assumption that $\chi^{(3)}(\omega) = \chi_{bulk}^{(3)}(\omega)$ for the DL layer at any aqueous interface.

As a result, reliable methodologies to deconvolve BIL/DL SFG signals now exist both from theory and experiments, and they can be combined to reveal the interfacial water arrangement at charged interfaces with the same accuracy as for neutral interfaces.

2.5 Extraction of the surface potential from non linear optical techniques and vice versa

Based on the eq. 2.17, one can easily calculate the potential difference across the DL region, i.e. the $\Delta\phi_{DL}$, once the $\chi_{DL}^{(2)}(\omega)$ spectrum is extracted, or reversely, reproduce the $\chi_{DL}^{(2)}(\omega)$ spectrum if the $\Delta\phi_{DL}$ can be determined. Both directions have been demonstrated by the two works presented in this manuscript, see section 3.1.3 and section 5.4 for details.

Here, we give a brief summary of the methods used to determine the surface potential $\Delta\phi_{DL}$ and on the other way around to reproduce the $\chi_{DL}^{(2)}(\omega)$ spectrum.

2.5.1 Finding the surface potential from extracted $\chi_{DL}^{(2)}(\omega)$ contribution to the experimental SFG spectrum

In the work presented in chapter 3, section 3.1, while studying the self-assembled monolayer (SAM)/water interfaces, we developed a method to extract the surface potential by combining the experimental SFG spectrum and our DFT-MD calculated SFG spectrum of water in contact with the SAMs. The details of the studied system are reported in section 3.1 and in the paper [*On the trail of molecular hydrophilicity and hydrophobicity at aqueous interfaces*, *J. Phys. Chem. Lett.* 2023, 14, 1301-1309] included in chapter 3.

The experimental spectra of the SAM-silica/water systems from our collaborator, the group of Prof. Poul Petersen [139] at the Ruhr-University Bochum in Germany, were measured at pH \sim 5.6, under which condition some of the silanol groups at the silica substrate on which the SAMs are anchored are dehydroxylated. The interfaces are thus negatively charged, leading to a $\chi_{DL,exp}^{(2)}(\omega)$ contribution to the experimental SFG spectra (denoted $\chi_{total,exp}^{(2)}(\omega)$). However, our DFT-MD simulated SAM/water interfaces are under the point of zero charge (PZC) condition for the silica (\sim pH 2-4), where the silica substrate is fully hydroxylated. In our simulations, the interfaces are thus neutral, hence only $\chi_{BIL}^{(2)}(\omega)$ contributes to our calculated spectra. Our DFT-MD calculated spectra of the SAM/water interfaces are thus denoted $\chi_{BIL,DFT-MD}^{(2)}(\omega)$.

To reconcile experiment and simulation, something has to be done.

We assume $\chi_{BIL}^{(2)}(\omega)$ to be invariant at the pH from 2 to 5.6 based on the previous researches [44, 140] showing that the $\chi_{BIL}^{(2)}(\omega)$ of a charged interface with a low surface charge density σ (less than a few percent of a monolayer) is barely modified compared to the neutral interface. Therefore, the $\chi_{DL}^{(2)}(\omega)$

contribution to the experimental SFG spectrum ($\chi_{DL,exp}^{(2)}(\omega)$) is found by subtracting the $\chi_{BIL,DFT-MD}^{(2)}(\omega)$ from the experimental SFG spectrum based on eq. 2.19:

$$\chi_{DL,exp}^{(2)}(\omega) = \chi_{total,exp}^{(2)}(\omega) - \chi_{BIL,DFT-MD}^{(2)}(\omega) \quad (2.19)$$

The potential difference across the DL region at the experimental SAM/water interfaces at pH \sim 5.6 [139], $\Delta\phi_{DL}$, can hence be obtained by eq. 2.20:

$$\Delta\phi_{DL} = \frac{\chi_{DL,exp}^{(2)}(\omega)}{\chi_{Bulk}^{(3)}(\omega)} \quad (2.20)$$

where the $\chi_{bulk}^{(3)}(\omega)$ spectrum is the universal spectrum reported in ref. [33, 44] and shown in the previous section and fig. 2.6. The subtleties of the extraction of $\Delta\phi_{DL}$ are detailed in chapter 3, section 3.1.3 and in our paper [*On the trail of molecular hydrophilicity and hydrophobicity at aqueous interfaces*, *J. Phys. Chem. Lett.* 2023, 14, 1301-1309] included in chapter 3.

2.5.2 Reproducing $\chi_{DL}^{(2)}$ in Second Harmonic Generation (SHG) by calculating the surface potential through the modified Gouy-Chapman model

As discussed at the beginning of this section, once the surface potential can be deduced, one can calculate the $\chi_{DL}^{(2)}$ contribution. This holds true not only for the SFG but also for Second Harmonic Generation (SHG).

SHG is a particular case of SFG spectroscopy. In SHG, two photons at the same frequency are absorbed and generate an output beam at the doubled frequency as illustrated in Fig. 2.7.

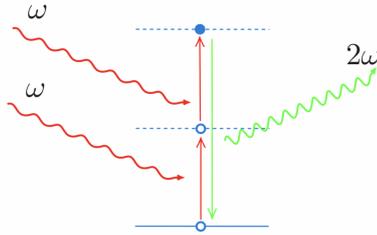


Figure 2.7: Scheme of the principle of the second harmonic generation (SHG).

Unlike in the SFG spectroscopy where a scan of the laser in the infrared (IR) frequency is done, in SHG the frequencies of the incident beams are fixed (same frequency for the two photons). Hence the SHG signal is not a function of the frequency as in the SFG.

According to ref. [141], the effective surface nonlinear polarization at 2ω is expressed by $\chi_{eff}^{(2)}$:

$$\chi_{eff}^{(2)} = \chi_{BIL}^{(2)} + \chi_{DL}^{(2)}, \quad (2.21)$$

$$\text{with } \chi_{DL}^{(2)} = \chi_{Bulk}^{(3)} \Delta\phi_{DL}, \quad (2.22)$$

$$\text{and } \Delta\phi_{DL} = \int_a^b E_0(z') e^{i\Delta k_z z'} dz', \quad (2.23)$$

where Δk_z is the phase mismatch of the reflected SHG, $\chi_{BIL}^{(2)}$ is the second-order nonlinear susceptibility of the BIL water, $\Delta\phi_{DL}$ is the potential difference across the DL region, and $\chi_{Bulk}^{(3)}$ is the third-order nonlinear response of water molecules to optical and static fields, which was also shown to be constant for SHG [141]. For SHG, $\chi_{Bulk}^{(3)} = (9.56 \pm 1.87) \times 10^{-22} \text{ m}^2/\text{V}^2$, as determined in ref. [141]. The minimum and maximum (a and b) of the integral in eq. 2.23 correspond respectively to the border between BIL and DL and the one between DL and bulk water. The integral between these two borders along the z direction is thus in the DL region, where the $E_0(z')$ is the electric field across this region.

In the work presented in chapter 5, we proposed a new interpretation of the SHG for the silica surface as a function of increasing pH (and increasing deprotonation of the surface), that answered a historically debated question: what is behind the bimodal behavior of the silica surface acidities initially found by Eisenthal *et al.* [46]. To not spoil too much here this interesting discovery, we refer the readers to chapter 5 and to our paper [“*Unveiling structural evolution and reconstruction of oxide surface in water*” in the review by Nature] included by the end of chapter 5 for the full demonstrations.

The relevant part of this work for this section is that we deduced the surface potential through the modified Gouy-Chapman model [141, 142] based on the knowledge and input parameters from DFT-MD simulations. We then managed to calculate the $\chi_{DL}^{(2)}$ contribution to the SHG signals of the silica/water interface under different pH conditions by eq. 2.22.

With the calculated $\chi_{DL}^{(2)}$ contribution, we successfully calculated $\chi_{eff}^{(2)}$ for the SHG signal, and successfully reproduced the evolution of SHG as a function of pH measured by Eisenthal in 1992 [46] at the silica/water interface, including the newly discovered species at the silica surface in certain pH ranges. This part is detailed in section 5.4.

2.6 From O-H stretching of water to Si-O stretching at the silica surface

Until now, all the theoretical methods of calculating SFG spectra based on MD simulations presented in this chapter (and in the literature) have been focused on the O-H stretching frequency region from 2000 cm^{-1} to 4000 cm^{-1} in order to probe the interfacial water organization.

However, in some of the works presented in this manuscript, we also managed to calculate SFG spectra of more than just the water part.

While collaborating with Prof. Wei-Tao Liu's Group at the University of Fudan in Shanghai, China and Prof. Y. Ron Shen at UC Berkeley on the investigations of the silica/water interface (detailed in chapter 5), we developed a methodology for calculating SFG spectra in the Si-O stretch region (i.e. 800–1200 cm^{-1} vibrational range) based on the approach introduced in section 2.3 for water O-H stretch. i.e. based on the VVCF (Velocity-Velocity Correlation Function) method.

The DFT-MD SFG spectrum of the silica/water interface in the Si-O stretching region is obtained by the sum of the individual $\chi^{(2)}(\omega)$ signals arising from the different Si-OX species existing at the silica surface (Si-OH, Si-O⁻, and Si-OSi), through:

$$\chi_{\text{total}}^{(2)}(\omega, pH) = \sum_j^{\text{SiOXgroups}} C_j(pH) \bar{\chi}_j^{(2)}(\omega), \quad (2.24)$$

written here at a given pH value, and where $C_j(pH)$ is the coverage of species j at a given pH (determined based on parameters from DFT-MD simulations, see chapter 5 for details) and $\bar{\chi}_j^{(2)}(\omega)$ is the Si-O SFG signal of the corresponding Si-OX species.

The species-specific DFT-MD SFG spectrum is obtained by (in the ssp polarization):

$$\bar{\chi}_{j,xxz}^{(2)}(\omega) = \frac{i}{k_B T \omega} \int_0^\infty dt e^{-i\omega t} \langle \dot{\alpha}_{j,xx}(t) \dot{\mu}_{j,z}(0) \rangle, \quad (2.25)$$

where $\mu_{j,z}(0)$ and $\alpha_{j,xx}(t)$ refer to the z and xx components of the dipole and polarizability of a given Si-O species in the laboratory frame and $\dot{\mu}_{j,z}(0)$ and $\dot{\alpha}_{j,xx}(t)$ are their time derivatives. Following what was done in section 2.3 for the water O-H vibration, using basic geometry considerations, we can express the dipole moment of the Si-O bond in the laboratory frame from the calculated one in the molecular frame by:

$$\dot{\mu}_{j,z}(0) \approx \sum_i^{x',y',z'} D_{j,z,i}(0) \dot{\mu}_{j,i}(0) \approx D_{j,z,\text{SiO}}(0) \dot{\mu}_{j,\text{SiO}}(0), \quad (2.26)$$

where x' , y' , z' are any set of three directions in the molecular frame, and D is a cosine matrix projecting the molecular frame onto the laboratory frame (see scheme in fig. 2.4 in section 2.3). The second equality is written assuming that the dipole component along the Si-O bond direction (i.e. z') is much larger than that in the perpendicular directions (i.e. x' , y').

By introducing a change of variable in the derivative $\dot{\mu}_{j,i}(0)$, we hence obtain:

$$\dot{\mu}_{j,z}(0) \approx D_{j,z,SiO} \frac{d\mu_{j,SiO}}{dr_{j,SiO}} v_{j,SiO}(0), \quad (2.27)$$

where $r_{j,SiO}$ is the Si-O bond coordinate that is one of the three directions in the molecular frame and $v_{j,SiO}(0) = \frac{dr_{j,SiO}}{dt}$. We assume that the orientation of the Si-O bond does not change substantially over time for a given Si-OH or Si-O⁻ group, therefore the dependence of $D_{j,z,SiO}$ on time can be neglected. This is true for solids, even at the top-surface of interest here.

With the same methodology for the polarizability, we have:

$$\dot{\alpha}_{j,xx}(t) = D_{j,x,SiO}^2 \frac{d\alpha_{j,SiOSiO}}{dr_{j,SiO}} v_{j,SiO}(t), \quad (2.28)$$

Substituting Eq. 2.27 and Eq. 2.28 into Eq. 2.25, we hence obtain:

$$\begin{aligned} \bar{\chi}_{j,xxz}^{(2)}(\omega) &= D_{j,x,SiO}^2 D_{j,z,SiO} \frac{d\alpha_{j,SiOSiO}}{dr_{j,SiO}} \frac{d\mu_{j,SiO}}{dr_{j,SiO}} \\ &\quad \frac{i}{k_B T \omega} \int_0^\infty dt e^{-i\omega t} \langle v_{j,SiO}(t) v_{j,SiO}(0) \rangle, \end{aligned} \quad (2.29)$$

where $D_{j,x,SiO}$, $D_{j,z,SiO}$, and $v_{j,SiO}(t)$ are obtained from the simulations, while $\frac{d\alpha_{j,SiOSiO}}{dr_{j,SiO}}$ and $\frac{d\mu_{j,SiO}}{dr_{j,SiO}}$ are parameterized based on reference ab initio simulations, see refs [36, 43] for water.

For the works done in chapter 5, we set them to unity since we are more interested in how the changes in species populations (as it happens when changing the pH) influence the SFG intensity than in the change in the actual SFG activity of the species. For a more accurate description of the SFG activity, $\frac{d\alpha_{j,SiOSiO}}{dr_{j,SiO}}$ and $\frac{d\mu_{j,SiO}}{dr_{j,SiO}}$ must be rigorously evaluated for all x' , y' , z' components as done for O-H of water in refs [36]. This has not been done in this work.

Once the species-specific $\bar{\chi}_{j,xxz}^{(2)}(\omega)$ signals are calculated (shown in chapter 5), we are then able to determine the total SFG spectrum as a function of pH by:

$$\begin{aligned} \chi_{\text{total}}^{(2)}(\omega, pH) &= C_{SiOH}(pH) \bar{\chi}_{SiOH}^{(2)}(\omega) + C_{SiO^-}(pH) \bar{\chi}_{SiO^-}^{(2)}(\omega) \\ &\quad + C_{SiOSi}(pH) \bar{\chi}_{SiOSi}^{(2)}(\omega), \end{aligned} \quad (2.30)$$

where Si-O-Si bonds are centrosymmetric thus SFG-inactive, hence:

$$\chi_{\text{total}}^{(2)}(\omega, pH) = C_{SiOH}(pH) \bar{\chi}_{SiOH}^{(2)}(\omega) + C_{SiO^-}(pH) \bar{\chi}_{SiO^-}^{(2)}(\omega). \quad (2.31)$$

Therefore, as long as the coverage in each surface species (SiOH and SiO⁻) is known, the SFG spectrum of the silica surface in contact with water can be calculated using the eq. 2.31. Applying this approach, we managed to obtain the theoretical SFG spectra of the silica/water interface in the Si-O stretch region as a function of pH, which is in great agreement with the experimental spectra measured by Liu *et al.* . This is shown in chapter 5. This method could be further used to investigate other silica/water interfaces with or without the presence of electrolytes.

Part III

Results

Chapter 3

A metric for molecular hydrophobicity combining structural and spectroscopic descriptors

The interfacial water organization at a solid surface plays an important role in many phenomena, such as in geochemistry, diffusion of pollutants, electrochemical catalysis [5, 143], electric double layer formation, biomolecule recognition [1, 2, 144], to name a few. The liquid bulk water breaks its centrosymmetry when it comes to the surface, and rearranges depending on the balance between surface-water and water-water interactions. This balance ultimately dictates the hydrophobicity and hydrophilicity of an interface. In other words, it dictates the properties of wetting the surface.

Macroscopically, the hydrophobicity/hydrophilicity of an interface is mostly evaluated through contact angle measurements. However, it fails to capture the local hydrophobicity/hydrophilicity at a heterogeneous interface at the molecular level. It is this local hydrophobicity that matters for a substantial amount of physical and chemical processes that happen at the interface [27, 110]. The search for microscopic descriptors is thus necessary. Numerous microscopic descriptors for hydrophobicity/hydrophilicity have been determined either through sum-frequency generation (SFG) spectroscopy [24–26] or by molecular dynamics (MD) [8, 30, 111].

In spectroscopy, the free OH peak in SFG [24, 25], i.e. a positive peak at above 3600 cm^{-1} , has been considered as a spectroscopic fingerprint for molecular hydrophobicity. However, this fingerprint fails to capture the complexities of hydrophobicity in terms of e.g. the effect of local morphology, topology, connected H-bond network at the interface. Also, it is unable to quantify the degree of hydrophilicity for an interface where no free OH peaks are present, i.e. for weak hydrophobic interfaces and hydrophilic interfaces. From the MD side, Godawat *et al.* [30] demonstrated that large local density fluctuations and higher probabilities of low water density provide clear signatures of hydrophobicity. Based on that, local compressibility has been

developed as a powerful theoretical probe to map microscopic hydrophilicity [8].

All these descriptors are capable of quantifying molecular hydrophobicity/hydrophilicity, but there is no link between these theoretically determined descriptors and spectroscopic fingerprints. This is where our developments are targeted.

In this chapter, we summarize the previous investigations of our group on hydrophilicity and hydrophobicity of aqueous interfaces ranging from the hydrophobic air/water interface to the hydrophilic silica/water interface. Based on these studies, we then introduce a metric to quantify molecular (local) hydrophilicity, which for the first time connects structural descriptors and spectroscopic fingerprints in the Sum-Frequency Generation (SFG). This metric was developed by coupling DFT-MD simulations and both experimental and theoretical SFG calculation for a set of self-assembled monolayer (SAM)/water interfaces. The application of this metric in keeping track of the dynamics and spatial resolution of interfacial molecular hydrophilicity is further presented.

3.1 Development of a new descriptor for quantifying microscopic hydrophobicity and hydrophilicity: the H/V descriptor

To unravel the specific interfacial water structure in contact with different surfaces of various hydrophobicity/hydrophilicity, we first focus on the very hydrophobic interfaces and very hydrophilic interfaces. The air/water interface, commonly considered as the hydrophobic prototype, has been broadly studied because of its relevance in many fields, such as in atmospheric chemistry [145–147]. The coupling of SFG spectroscopy with MD simulations has vastly broadened our understanding of the H-bond network organization of water in contact with the vacuum. The consensus reached by both experimental [148] and theoretical works [32, 149, 150] in the literature is the discovery of the dangling O-H groups in the topmost water layer pointing towards the vacuum and of the O-H groups H-bonded to the subsequent water layer. The SFG signature for the dangling O-H groups, which is a positive band located at 3700 cm^{-1} , has from there been adopted by the spectroscopic community as a microscopic hydrophobicity descriptor for an aqueous interface.

While dissecting the SFG spectrum of the air/water interface, several interpretations of the structural organization of water molecules responsible for the SFG signals have been proposed. Paesani *et al.* [149] reported that the interfacial water molecules are made of tetrahedral water molecules plus doubly H-bonded water molecules. Skinner *et al.* [150, 151] interpreted the SFG spectrum in terms of water pairs being hydrogen-bond donor and acceptor. However, neither the actual thickness of the interfacial water layer nor the explicit 3D view of the water network as a whole reorganized at the interface

was mentioned in these interpretations.

Previous studies of Dr. Simone Pezzotti (a former Ph.D. student in the group of Prof. M.P. Gaigeot, NNT: 2019SACLE008) and coworkers have identified the collective 2-dimensional H-bond network (denoted 2DN) [26, 32], formed by water-water H-bonds parallel to the Willard and Chandler instantaneous interface [109] at the air/water interface, see Fig. 3.1. Fur-

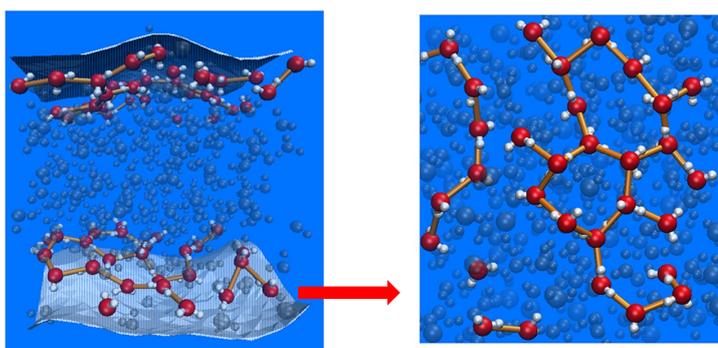


Figure 3.1: Collective 2-dimensional H-bond network (2DN) formed at the air/water interface [32]. The H-bonds are schematized by the orange connectors in the picture. The H-bonds are oriented in a plane, parallel to the surface.

thermore, while the 2DN is SFG inactive due to the orientation of the O-H groups being perpendicular to the SFG normal vector (see scheme in fig. 2.3 in chapter 2), its fingerprint in the THz-IR spectroscopy was observed. It is the band located at 164 cm^{-1} arising from the intermolecular stretching of H-bonds forming the 2DN, demonstrated by coupling theoretical and experimental THz-IR spectroscopy [26, 152].

The presence of the 2DN is thus considered another hydrophobic proof in addition to the SFG fingerprint of dangling O-H groups. This has indeed been generalized in ref. [32] by Pezzotti *et al.*

As regards a hydrophilic surface wetted by water, one can easily imagine that at the molecular level the surface sites (their nature, type, pKa, quantity over the surface, etc) are crucial parameters for the H-bonds that can be formed between surface and water. Therefore, one anticipates different H-bonds patterns, and therefore hydrophilicity should take different forms depending on the surface. As an example, the silica/water interface, known as a typical hydrophilic interface in the macroscopic scale, has been widely investigated because of the diversity of sites at the surface, different pKa values of surface silanols, and morphology, etc.

As revealed by Shen's group from SFG spectroscopy in the phonon region [153], three possible sites can be exposed at silica surfaces: SiOH (silanols) groups, SiO⁻ (deprotonated) groups and Si-O-Si siloxane bridges. While the silica/water interfaces are macroscopically hydrophilic, local microscopic hydrophobic patches can be present at the silica surface in the form of regions made by siloxane sites.

Ref. [27] discovered the signature of dangling O-H groups of water at

hydrophilic silica/water interfaces. DFT-MD simulations demonstrated that the water molecules in the topmost layer above the silica surface hydrophobic patches form preferential water-water HBs parallel to the surface. This can ultimately lead to the formation of the 2-dimensional H-bond network (2DN) as the one discovered at the air/water interface, if a sufficient amount of siloxane patches are present over the surface. This is the case for a low density of silanols (3.5 SiOH/nm^2) at the surface [27, 43, 110], as shown in Fig. 3.2 on the left. While increasing the number of hydrophilic sites by

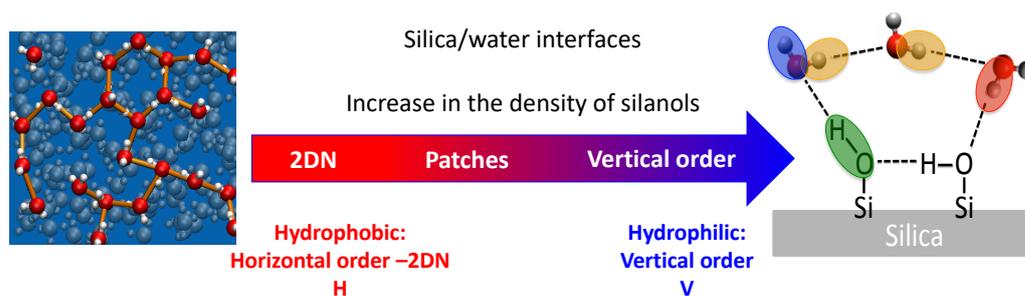


Figure 3.2: From hydrophobic to hydrophilic silica/water interfaces: horizontal order (2DN) to vertical order (surface-water H-bonds) of the organization of water.

tuning, e.g. the surface morphology, the density of silanols, and the degree of crystallinity, more and more water-water H-bonds parallel to the surface are replaced by H-bonds formed between water and the hydrophilic sites at the surface [43, 50], leading to a “vertical” order in the water organization, as illustrated in Fig. 3.2 on the right.

In a nutshell, as shown in Fig. 3.2, above a hydrophobic interface (e.g. air/water, silica/water of low silanol density), horizontal H-bonds formed between water molecules in the topmost layer in direct contact with the surface, i.e. the Binding Interfacial Layer (denoted BIL, see section 2.4 for details) are dominant, leading to a horizontal ordered water organization (2DN) [26, 32]. However, a more vertical ordered water organization exists above a hydrophilic interface (e.g. silica/water interface of high silanol density) induced by surface-water H-bonds [154]. The preferential water network orientation in the BIL, i.e. horizontal or vertical order, can thus be used to quantify the hydrophobicity/hydrophilicity of an aqueous interface.

Inspired by these knowledge, we chose to use a platform composed of a series of Self-Assembled Monolayers (SAM)/water interfaces in MD simulations, where the hydrophilicity can be easily tuned, to validate our theory of horizontal and vertical ordered water organization and develop a new descriptor to quantify hydrophobicity/hydrophilicity. Self-Assembled Monolayers have attracted augmenting interest within the last decade as a “flexible tool for tuning molecular surface properties”. SAM coatings have been widely applied in numerous domains, i.e. from protein adsorption prevention in anti-fouling material design [155–157] to the production of gate dielectric surface

3.1 Development of a new descriptor for quantifying microscopic hydrophobicity and hydrophilicity: the H/V descriptor

treatments in organic thin-film transistors (OTFT) [158]. The SAMs/water interfaces have attracted considerable attention as a model platform to study hydrophobicity/hydrophilicity and have been extensively characterized in a number of theoretical [159–161] and SFG experimental [113, 139, 140, 162] works. Here, we investigate a set of SAM/water interfaces of well-controlled and tunable hydrophilicity.

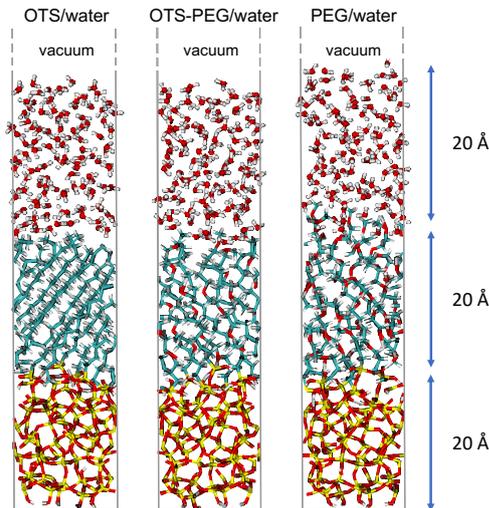


Figure 3.3: Snapshots of the three SAM/water interfaces from the DFT-MD simulations. In each snapshot, from bottom to top: amorphous silica, SAMs anchored chemically on silica, water, and vacuum (Si: yellow, C: cyan, O: red, H: white). Box dimensions: $13.386 \times 13.286 \times 85 \text{ \AA}^3$.

We have carried out DFT-MD simulations for three different silane-based SAM/water interfaces (with the SAMs anchored on a silica substrate): a monolayer made of octadecyltrichlorosilane (**OTS**, $\text{CH}_3(-\text{CH}_2)_{17}-\text{SiCl}_3$) polymer chains (intrinsically hydrophobic), a monolayer made of polyethylene glycol (**PEG**, $\text{CH}_3(-\text{O}-\text{CH}_2-\text{CH}_2)_7-\text{CH}_2-\text{SiCl}_3$) polymer chains (intrinsically hydrophilic), and mixed **OTS-PEG** monolayer (with a choice of 1:1 mixing ratio, alternately distributed in the simulation box). The three systems correspond to the three experimental samples (OTS/water, mixed OTS-PEG/water and PEG/water) prepared by Sanders *et al.* [139] for SFG spectroscopy.

For the mixed OTS-PEG monolayer, it should be noted that we only have made one (ideal) scenario for the mixing of the two types of polymers in the DFT-MD simulations, which is a homogeneous mixing, where 1 PEG alternates with 1 OTS all over the surface. This is unfortunately not enough to compare to the experiment, where not only the rate of mixing is unknown but also the extent of homogeneous/heterogeneous disposition of the monolayer is unknown.

For DFT-MD, the simulation boxes of $13.386 \times 13.286 \times 85.0 \text{ \AA}^3$ are each composed of eight SAM chains anchored to an amorphous silica surface with a hydroxylation degree of 4.5 SiOH/nm^2 , solvated by 120 water molecules (see

Fig. 3.3). The amorphous silica model is taken from Ugliengo *et al.* [163], it can accommodate up to 8 SiOH groups at the surface at PZC conditions, see refs [27, 43]. The 8 SiOH groups are replaced by 8 polymeric chains. Three different systems are built with 8 OTS chains (pure OTS/water), alternately mixing of 4 OTS and 4 PEG chains (mixed OTS-PEG/water), and 8 PEG chains (pure PEG/water), respectively. Each box is periodically repeated in the x and y directions and separated by a vacuum layer of 25 Å from the replica in the vertical z-direction (also replicated in space).

For each system, 4 individual trajectories are accumulated with initial configurations extracted from different geometry optimization steps (that have relaxed all elements in space). Each of the 4 DFT-MD simulations are composed of 10 ps of equilibration for each trajectory before the production run of 100 ps, both in the NVE ensemble (with velocity rescaling and a target temperature of 300 K in the equilibration part only). In total, 400 ps of trajectories have been accumulated for each SAM/water interface (OTS/water, PEG/water, and mixed OTS-PEG/water).

Both the interfacial water organization and the structure of SAMs at the interface are dissected from the MD simulations. The water theoretical SFG spectra of the three interfaces are calculated from the simulations (detailed in section 2.3) and compared with the experimental measurements [139]. The method of theoretical SFG spectra calculation is detailed in section 2.3. This work was performed under PRACE (Partnership for Advanced Computing in Europe) under No. 2018194688. 33 million CPU hours have been consumed for these big systems in DFT-MD simulations.

3.1.1 Structural interpretation of DFT-MD calculated SFG spectra of water at the interface with the SAMs

All our simulations are carried out under the point of zero charge (PZC) condition for silica (pH \sim 2), which means that all simulated interfaces are neutral without surface charges. Therefore, the calculated SFG signals of water only contain the contribution from the BIL-water (\sim 20 water molecules), see section 2.4 for definition of BIL-water. However, the experimental SFG spectra are measured at pH \sim 5.6 [139], where both $\chi_{BIL}^{(2)}(\omega)$ and $\chi_{DL}^{(2)}(\omega)$ contribute to the SFG signal (see section 2.4 for BIL and DL contributions to SFG). The way to reconcile theoretical and experimental SFG spectra in this context is detailed later in section 3.1.3.

We first focus on comparing the OTS/water interface and the PEG/water interface in terms of the interfacial water structure and their spectroscopic signatures. Fig. 3.4B presents the DFT-MD calculated $\text{Im}\chi_{BIL}^{(2)}(\omega)$ (red lines) of the OTS/water interface (top) and of the PEG/water interface (bottom) over \sim 400 ps of simulation of each (4 trajectories for each interface, each trajectory \sim 100 ps).

In each figure, the deconvolution of the (red) signal into water OH-up and

3.1 Development of a new descriptor for quantifying microscopic hydrophobicity and hydrophilicity: the H/V descriptor

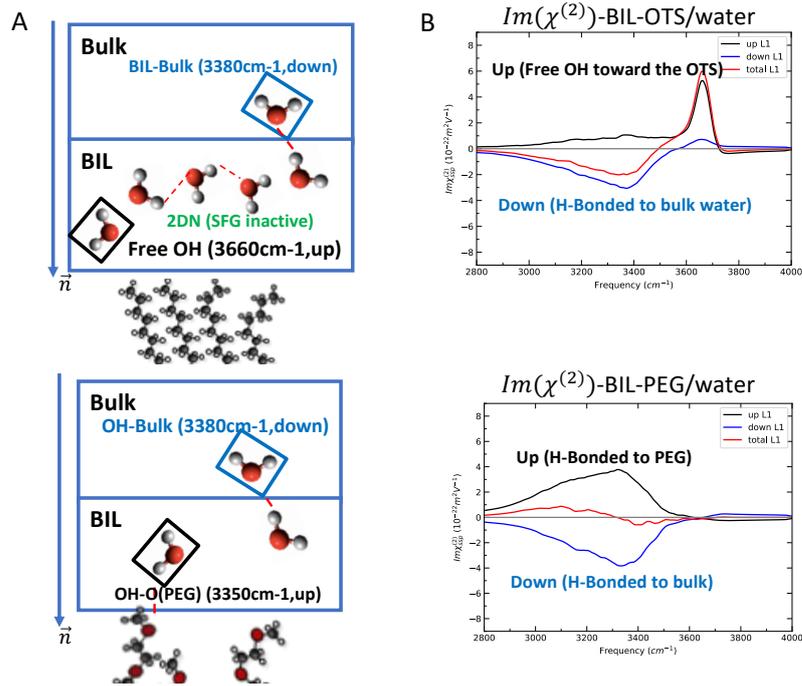


Figure 3.4: (A) Schemes of interfacial water motifs with their corresponding SFG fingerprints. (B) DFT-MD calculated spectra of $Im\chi_{BIL}^{(2)}(\omega)$ under PZC (pH 2~4) conditions of OTS/water interface (top) and PEG/water interface (bottom). Total $Im\chi_{BIL}^{(2)}(\omega)$ are depicted in red and the deconvolution of the (red) signal into water OH-up and OH-down (respectively directed along the normal \vec{n} , outward the normal \vec{n}) is presented.

OH-down (respectively directed along the normal \vec{n} , outward the normal \vec{n}) is presented. In Fig. 3.4B top (red), we recognize the positive peak at 3660 cm^{-1} corresponding to the dangling O-H group of water pointing towards the OTS monolayer in the $Im\chi_{BIL}^{(2)}(\omega)$ spectrum of OTS/water, proving the hydrophobicity of this interface. There is also a negative band centered at 3380 cm^{-1} , corresponding to water OH groups pointing and H-bonded to the subsequent water layer. This spectrum resembles the one of the air/water interface with the same molecular interpretation. The frequency of the free OH peak is slightly lower than that of the air/water interface, possibly due to the van der Waals interactions between the dangling OH groups and the OTS monolayer as suggested by Tian *et al.* [140].

The SFG spectrum of the PEG/water interface in Fig. 3.4B-bottom (red) surprisingly presents an almost zero signal of $Im\chi_{BIL}^{(2)}(\omega)$ in the whole OH stretch region. There is no signal in the free OH region as expected for a hydrophilic PEG surface. The zero signal in the H-bonded region was unforeseen.

Thanks to the strength of DFT-MD calculated SFG spectra, we can easily deconvolve the total spectra into contributions from water molecules with specific properties, as detailed in section 2.3. Knowing that the normal in SFG is the vector perpendicular and pointing to the solid surface (\vec{n} in fig. 2.3),

we here deconvolve the total $\text{Im}\chi_{BIL}^{(2)}(\omega)$ spectra based on the orientation of the water dipole moments into contributions from water-up (dipole pointing towards the monolayer, black lines in Fig. 3.4B) and from water-down (dipole pointing towards the subsequent water layer, blue lines in Fig. 3.4B).

The deconvolved water-up and water-down spectra of the OTS/water interface respectively show a positive peak at 3660 cm^{-1} , a positive band centered at $\sim 3380\text{ cm}^{-1}$ of low intensity (black line in Fig. 3.4B), and a negative band at 3380 cm^{-1} (blue line in Fig. 3.4B). This corresponds to dangling OH groups pointing towards the OTS and H-bonded OH groups to bulk water, respectively. We further find that the water molecules in the BIL are organized with the 2-dimensional H-bond network (2DN) [26, 32] formed by parallel H-bonds among BIL-water molecules, but this can not be captured by the SFG spectrum due to the H-bond orientation, while it is captured by a structural descriptor for hydrophobicity, H/V descriptor, that will be presented in section 3.1.2.

In the PEG/water spectrum, two bands (up in black and down in blue) with opposite signs and same intensities centered at the same frequency are observed after the up-down deconvolution. The water-up (black line) band corresponds to the O-H groups pointing towards the PEGs (along the normal \vec{n}) and H-bonded to the PEGs (as will be shown later in section 3.1.2), while the water-down (blue line) represents the O-H groups pointing towards the subsequent bulk water layer (opposite to the normal \vec{n}) and H-bonded to it, see Fig. 3.4A-bottom.

We thus understand that the zero signal in the total $\text{Im}\chi_{BIL}^{(2)}(\omega)$ of PEG/water is due to the compensation of these two bands. It is amazing to see that the strength of H-bonds made between water-water and PEG-water is the same as revealed by the same frequency of the up and down bands. This has not been reported in the literature up to now.

Based on the understanding of the 2 SFG spectra of the purely hydrophobic and purely hydrophilic interfaces, we now interpret the total and the up-down deconvolved $\text{Im}\chi_{BIL}^{(2)}(\omega)$ spectra of all three systems, including the 1:1 mixed OTS-PEG/water (see Fig. 3.5).

Since the interactions between BIL-water molecules and the surface dictates the hydrophobicity/hydrophilicity of the interface, we fix our attention on the **deconvolved water-up signal**, which corresponds to the O-H groups of water pointing towards the monolayers and interacting with the monolayers.

With the increase of hydrophilic sites at the surface, i.e. from OTS to mixed OTS/PEG to PEG, we observe a decrease in the intensity of the peak in the free OH region (above 3600 cm^{-1}) and an increase of intensity of the one in the H-bonded region ($3000\text{--}3550\text{ cm}^{-1}$). We reiterate that the O-H groups parallel to the surface forming the 2DN above a hydrophobic surface are SFG inactive since they are perpendicular to the SFG normal vector. Therefore, the more pronounced the SFG signal is, the more vertical ordered the water network is at the interface. As the surface becomes more

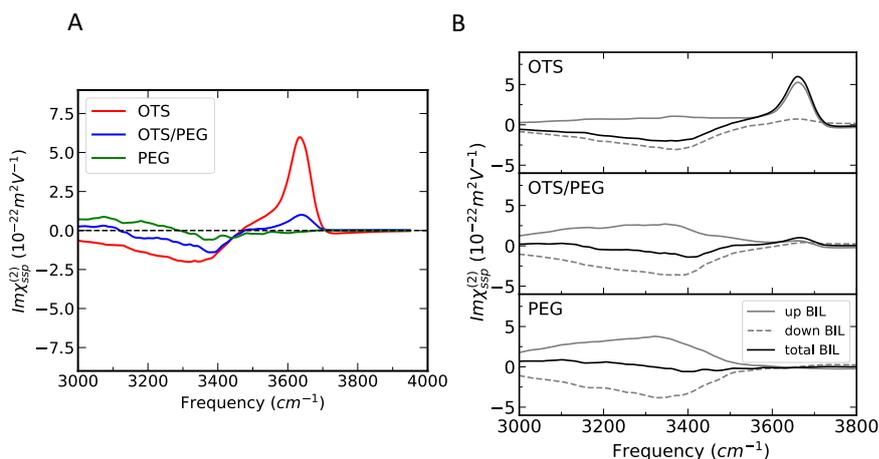


Figure 3.5: DFT-MD calculated SFG spectra of the three SAM aqueous interfaces (OTS/water, OTS-PEG/water, PEG/water) in the SSP polarization: (A) Total $\text{Im}\chi_{BIL}^{(2)}(\omega)$ and (B) Deconvolved up and down $\text{Im}\chi_{BIL}^{(2)}(\omega)$.

and more hydrophilic, the initially dangling O-H groups in the 2DN of the BIL form vertical H-bonds with oxygen atoms of the PEGs, leading to the disappearance of the free OH peak in the SFG spectrum and a growth of the deconvolved water-up SFG signal in the H-bonded region below 3600 cm^{-1} . The more hydrophilic the interface is, the more vertical the H-bonds formed between water and the surface, thus the more intense the water-up spectrum is in the H-bonded region.

3.1.2 Spectroscopic and structural descriptors of microscopic hydrophilicity/hydrophobicity: the H/V descriptor

The quantification of the balance between the horizontal and vertical H-bonds at the interface can provide the information on the hydrophobicity and hydrophilicity of an interface at the molecular scale. We hereby introduce our descriptors for measuring molecular hydrophobicity/hydrophilicity by quantifying the horizontal and vertical H-bonds both spectroscopically and structurally.

Spectroscopic descriptor

Based on the understandings of DFT-MD calculated BIL-SFG spectra of the three SAMs/water systems in section 3.1.1, we here introduce our spectroscopic descriptors for molecular hydrophobicity/hydrophilicity, by integrating the water-up SFG spectra both in the free-OH region ($3550\text{--}3800\text{ cm}^{-1}$) and in the H-bonded ($3000\text{--}3550\text{ cm}^{-1}$) region. Since the horizontal H-bonded OH groups are SFG inactive, the integral of the H-bonded signal in the SFG spectrum solely contains the vertical components of the H-bonded OH groups. However, as shown above, the signals arising from OH groups

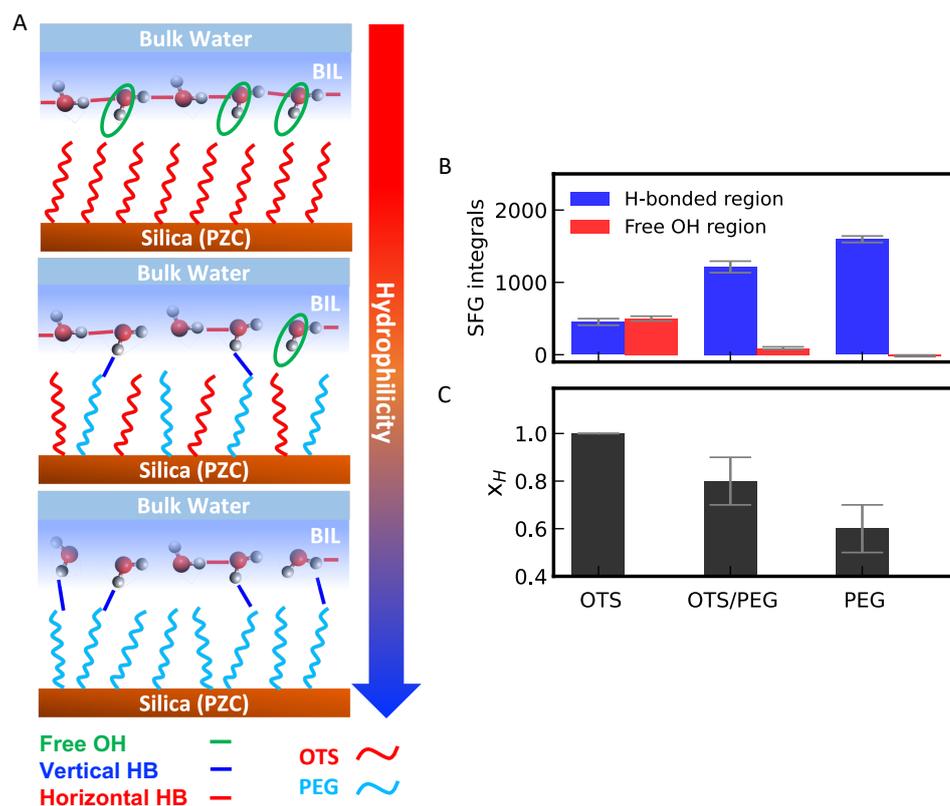


Figure 3.6: Hydrophobicity/hydrophilicity descriptors investigated in this work. (A) Scheme of horizontal H-bonds (red) and vertical H-bonds (blue) in the 3 SAMs/water systems investigated here. Two microscopic descriptors determined from SFG signals and DFT-MD simulations: (B) SFG intensity integrals of the deconvolved water-up spectra in the spectral regions of the free-OH ($3550\text{--}3800\text{ cm}^{-1}$, red bars) and H-bonded ($3000\text{--}3550\text{ cm}^{-1}$, blue bars) OH stretching motions, and (C) structural H/V descriptor based on the preferential orientation within the water network in the binding interfacial layer (BIL).

H-bonded to the O atoms of PEGs are washed out in the total BIL spectrum of PEG/water, because of the compensation between water-up and water-down contributions due to the same strength of PEG-water and water-water H-bonds. We hence choose to calculate the integral of the deconvolved water-up signal only, in order to account for the water-surface H-bonds only.

The larger the integral of the free OH peak is, the more hydrophobic the interface is. And the larger the integral of the water-up H-bonded peak is, the more vertical H-bonds are formed between BIL-water and the SAMs, hence the more hydrophilic the surface is. The results are shown in Fig. 3.6B, red bars for free OH peak integrals and blue bars for H-bonded peak integrals. As expected, the integrals of the H-bond signals (blue) are anticorrelated with the ones of the free OH signal (red), i.e. the larger the integrals of the H-bond signal, the smaller that of the free OH signal. One can nicely see that OTS is the most hydrophobic, however with a non negligible contribution from the H-bonded region. The mixed OTS/PEG is slightly less hydrophilic than the PEG system.

Structural H/V descriptor

Alternatively, we can directly quantify the horizontal (inactive in SFG) and vertical H-bonds in a more direct way from the DFT-MD simulations. We hence complement the spectroscopic descriptor with a structural one, namely the H/V descriptor. To that end, we calculate the number of horizontal HBs, n_H , defined as water-water H-bonds in the BIL, parallel $\pm 30^\circ$ to the instantaneous water surface [109] in contact with SAMs, red lines in Fig. 3.6A. We also calculate the number of vertical HBs, n_V , defined as water-surface H-bonds [32], blue lines in Fig. 3.6A).

We then define the molar fraction x_H of horizontal H-bonds as our H/V descriptor to quantify the hydrophobicity/hydrophilicity of the surface. The H/V descriptor is thus:

$$H/V = x_H = \frac{n_H}{n_H + n_V} \quad (3.1)$$

The lower the x_H is, the less horizontal H-bonds are and the more vertical H-bonds are formed at the interface, hence the more hydrophilic the interface is. The H/V descriptor values of the three systems are presented in Fig. 3.6C. The x_H values range from 1 for the pure OTS/water interface to 0.63 for the pure PEG/water interface. A value of 0.77 is obtained for the mixed OTS-PEG/water interface in between the two extreme values. The values are nicely correlated with the integral values of the free OH SFG signal (Fig. 3.6B, red bars) and are nicely anti-correlated with the integral values of the (vertical) H-bonded OH SFG signal (Fig. 3.6B, blue bars).

This work provides for the first time a framework to connect a structural descriptor for molecular hydrophobicity that quantitatively informs on the changes in the H-bond network, with spectroscopic quantities measured by SFG.

More details for the development of our metric for quantifying molecular hydrophobicity/hydrophilicity are presented in our paper [*On the trail of molecular hydrophilicity and hydrophobicity at aqueous interfaces, J. Phys. Chem. Lett. 2023, 14, 1301-1309*], reported on pages 81–97 of this thesis.

3.1.3 How DFT-MD helps to understand experimental SFG spectra: extraction of the surface potential

Our presented metric combining spectroscopic fingerprints and structural analyses can furthermore be used for a detailed analysis of the experimental SFG. Here, for SAM/water interfaces, we use the experimental SFG spectra measured by our collaborator, the group of Prof. Poul Petersen [139] in Bochum, Germany.

While our simulations are under PZC conditions for the silica substrate, i.e. no surface charge, the experimental SFG signals are measured at $\text{pH} \sim 5.6$, where the silica surface is negatively charged due to the deprotonation of some of the silanol groups at the silica substrate. Hence the experimental SFG spectra of water include both the contributions from the Binding

Interfacial Layer (BIL) and from the Diffuse Layer (DL), see section 2.4. Due to the small box dimension limit imposed by DFT-MD simulations, we are not capable of simulating the whole diffuse layer since it requires a much thicker water slab in the simulation box. Hence we are not able to directly calculate the $\text{Im}\chi_{DL}^{(2)}(\omega)$ contribution to the total SFG spectrum at the experimental pH (~ 5.6) directly from our simulations.

However, we can calculate this DL contribution in the following way.

As already explained in Chapter 2, section 2.5, and refs. [44, 140], the $\chi_{BIL}^{(2)}(\omega)$ of a charged interface with a low surface charge density σ (less than a few percent of the water monolayer) is barely modified compared to the neutral (PZC) interface, because the structure of water in the BIL is barely affected by the small change in the surface charge. We thus can assume $\text{Im}\chi_{BIL}^{(2)}(\omega)$ to be invariant from pH 2–4 to 5.6, at which the experimental spectra are measured. Therefore, the BIL-water contribution obtained in the DFT-MD spectra (pH ~ 2) of the SAMs/water systems investigated here is the same as that in the experimental spectra at pH ~ 5.6 . The difference between experimental and theoretical spectra is then solely due to the DL contribution at pH ~ 5.6 .

As introduced in Chapter 2, section 2.4, the $\chi_{DL}^{(2)}(\omega)$ equals to the $\chi_{Bulk}^{(3)}(\omega)$, the third order susceptibility of bulk water (proved to be a universal constant spectrum independent on systems [33, 44]), multiplied by the potential difference $\Delta\phi_{DL}$ across the DL region. In the reverse way, if $\chi_{DL}^{(2)}(\omega)$ is known experimentally, one can extract $\Delta\phi_{DL}$ by:

$$\Delta\phi_{DL} = \frac{\chi_{DL}^{(2)}(\omega)}{\chi_{Bulk}^{(3)}(\omega)} \quad (3.2)$$

However, life is never so simple in reality. The experimental SFG spectra measured by Sanders *et al.* [139] for the SAM/water systems are reported in arbitrary units instead of in absolute values, contrary to our SFG calculations. Because of this difference in units of the SFG intensity, we have to do something in order to exploit the experiments to our advantage and hence be able to extract $\Delta\phi_{DL}$ from experiments. A fitting procedure is thus required. To that end, we use the SFG spectrum (both experimental and theoretical) of the water at the interface with OTS/water as a reference for the fitting procedure that we explain below.

In the experimental spectrum of the OTS/water interface, a broad band with two peaks centered at 3200 and 3400 cm^{-1} and a free OH peak at 3660 cm^{-1} are recorded, see Fig. 3.7B-top, dashed black. The 2-peaks band at 3200 and 3400 cm^{-1} is the $\chi_{DL}^{(2)}(\omega)$ contribution to the SFG spectrum of water at the interface with OTS.

Since the shape of $\chi_{DL}^{(2)}(\omega)$ is determined by the shape of $\chi_{Bulk}^{(3)}(\omega)$ (two band spectrum centered at 3200 and 3400 cm^{-1} , as demonstrated both experimentally [44] and theoretically [33] and shown in Fig. 3.8), the addition of

3.1 Development of a new descriptor for quantifying microscopic hydrophobicity and hydrophilicity: the H/V descriptor

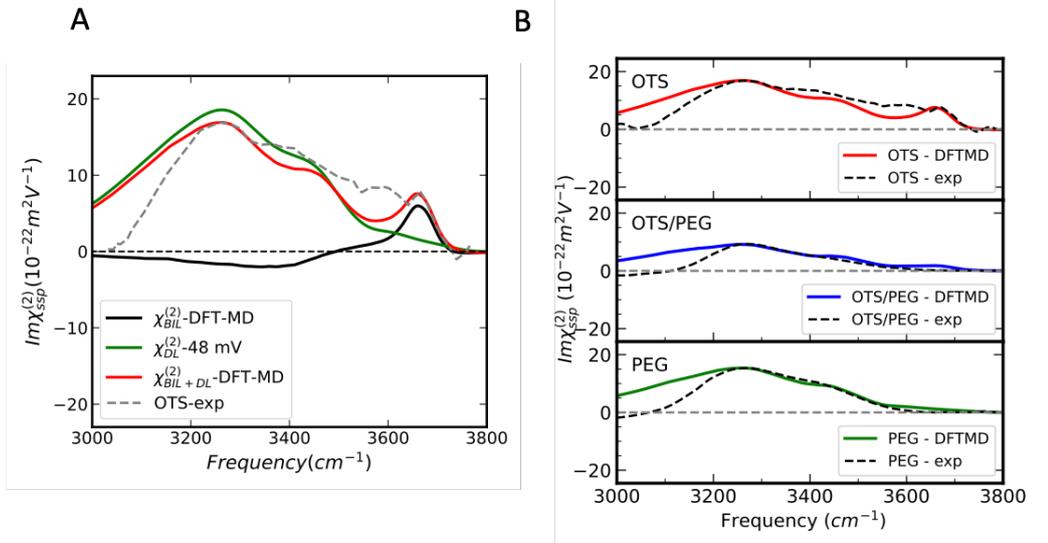


Figure 3.7: (A) SFG spectra in the ssp polarization of the OTS/water interface with DFT-MD calculated $\text{Im}\chi_{\text{BIL}}^{(2)}(\omega)$ contribution (black), fitted $\text{Im}\chi_{\text{DL}}^{(2)}(\omega)$ contribution (green), see text for details, and the sum of the two ($\text{Im}\chi_{\text{tot}}^{(2)}(\omega)$, in red). The experimental SFG spectrum of the OTS/water interface from Sander *et al.* [139] is plotted in dashed grey. (B) SFG spectra of the three SAM aqueous interfaces (OTS/water, OTS-PEG/water, PEG/water) in the SSP polarization: DFT-MD calculated spectra of $\text{Im}\chi_{\text{BIL}}^{(2)}(\omega)$ under PZC (pH 2~4) conditions plus fitted $\text{Im}\chi_{\text{DL}}^{(2)}(\omega)$ contribution are depicted in colors (see text for details) and experimental spectra measured by Sanders *et al.* [139] at pH ~ 5.6 are plotted in dashed black.

the $\chi_{\text{DL}}^{(2)}(\omega)$ to the $\chi_{\text{BIL}}^{(2)}(\omega)$ solely modifies the intensity and shape of the band in the H-bonded region ($3000\text{--}3550 \text{ cm}^{-1}$) and does not change the intensity of the free OH peak at 3660 cm^{-1} .

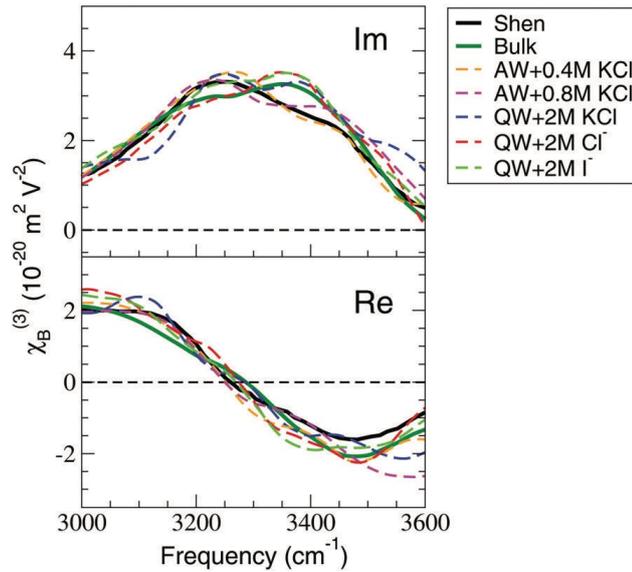


Figure 3.8: A comparison of several $\chi_{\text{Bulk}}^{(3)}(\omega)$ spectra (imaginary and real components) calculated in ref. [33] for different interfaces, for bulk liquid water and from Shen's experiments [44]. See section 2.4 for more explanations.

Based on this, the experimental relative intensity of the broad band in the H-bonded region and the free OH peak can be used as an objective function (OF) to obtain $\Delta\phi_{DL}$ from a fitting procedure, as expressed by:

$$\begin{aligned}
 OF &= \frac{\chi_{exp}^{(2)}(\omega_{freeOH})}{\chi_{exp}^{(2)}(\omega_{Hbonded})} \\
 &= \frac{\chi_{BIL,theory}^{(2)}(\omega_{freeOH})}{\chi_{BIL,theory}^{(2)}(\omega_{Hbonded}) + \Delta\phi_{DL} \cdot \chi_{bulk}^{(3)}(\omega_{Hbonded})} \quad (3.3)
 \end{aligned}$$

where $\chi_{exp}^{(2)}(\omega)$ is the experimental SFG spectrum, $\chi_{BIL,theory}^{(2)}(\omega)$ and $\chi_{bulk}^{(3)}(\omega)$ are theoretical spectra calculated from the DFT-MD simulations, and $\omega_{freeOH} = 3660 \text{ cm}^{-1}$, $\omega_{Hbonded} = 3250 \text{ cm}^{-1}$. In the fitting procedure, we adjust the $\Delta\phi_{DL}$ value until we obtain the same ratio between the intensity of peak at $\omega_{freeOH} = 3660 \text{ cm}^{-1}$ and at $\omega_{Hbonded} = 3250 \text{ cm}^{-1}$ in the theoretical OTS/water SFG spectrum as in the experimental SFG spectrum of the OTS/water interface (our OF in eq. 3.3). The $\Delta\phi_{DL}$ at experimental OTS/water interface is hence extracted.

For the OTS/water interface, Fig. 3.7A presents the $\chi_{BIL}^{(2)}(\omega)$ calculated from DFT-MD simulations (in black), the fitted $\chi_{DL}^{(2)}(\omega)$ based on the presented methodology above (in green) and the sum of the two, i.e. $\chi_{total}^{(2)}(\omega)$ (in red). The experimental spectrum of OTS/water interface is plotted in grey dashed line for comparison. Note that the experimental spectrum (in arbitrary unit) of the OTS/water is scaled such that the maximum intensity is the same as that obtained for the theoretical $\chi^{(2)}(\omega)$ spectrum after the fitting procedure. The experimental spectra (in arbitrary unit) of the PEG/water interface and the mixed OTS-PEG/water interface are normalized by the same scaling factor.

We are hence able to determine the $\chi_{DL}^{(2)}(\omega)$ contributions of the other two interfaces (mixed OTS/PEG and pure PEG), and calculate the $\Delta\phi_{DL}$ by eq. 3.2.

The extracted $\Delta\phi_{DL}$ values for the three interfaces are reported in Fig. 2C in the paper [*On the trail of molecular hydrophilicity and hydrophobicity at aqueous interfaces, J. Phys. Chem. Lett. 2023, 14, 1301-1309*]. The screening effect of the SAM chains on the $\Delta\phi_{DL}$ induced by the deprotonated sites at the silica substrate is discovered and discussed in the paper, page 1304.

The $\chi_{total}^{(2)}(\omega)$ of the three systems composed of DFT-MD calculated $\chi_{BIL}^{(2)}(\omega)$ and fitted $\chi_{DL}^{(2)}(\omega)$ are compared to the experimental spectra and presented in Fig. 3.7B. The theoretical spectra of the three interfaces obtained by this way are in good agreement with the experimental SFG spectra at above 3200 cm^{-1} . We however see that our calculations are always overestimating the experiment below 3200 cm^{-1} . Why is that?

Apart from water, SiOH groups on silica also contribute to SFG spectra in the O-H stretch range. The theoretical SFG signature of SiO-H of an

3.1 Development of a new descriptor for quantifying microscopic hydrophobicity and hydrophilicity: the H/V descriptor

amorphous silica/water interface presents a negative peak at low frequency (3000–3200 cm^{-1}), corresponding to surface SiO-H groups H-bonded to water, as reported in ref. [27, 43]. We present in Fig. 3.9 the calculated SFG signal of SiO-H stretch at the amorphous silica/water interface, from refs.[27, 43]. However, in our systems, silica is not directly in contact with water since it is

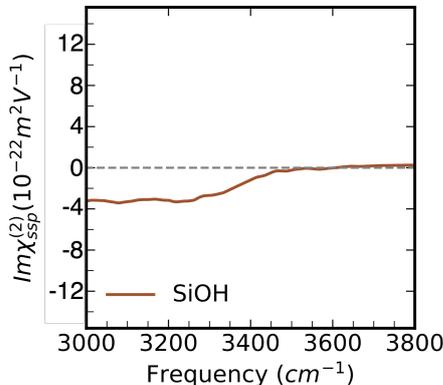


Figure 3.9: SFG signal arising from SiOH groups on our model silica surface put in contact with liquid water, as previously calculated from DFT-MD in refs. [27, 43].

covered by SAMs, and no water is observed to penetrate the monolayer and reach silica in the simulations. The spectral contributions of SiOH groups are therefore not taken into account in the theoretical SFG spectra of SAM/water interfaces.

Although the water penetration is not observed in the simulations, where the monolayers are densely packed, it is still possible to have water penetrating the monolayer and H-bonds to SiOH in the experiments, due to the possible less dense packing and more disorder of the monolayers generated during the deposition procedures. As shown in Fig. 3.7B, our DFT-MD calculated SFG spectra of SAM/water interfaces are in good agreement with the experimental ones [139] at above 3200 cm^{-1} . There is a systematic discrepancy between the theoretical and experimental SFG spectra in the low-frequency region (3000–3200 cm^{-1}), where theoretical SFG spectra present positive tails that are not observed in the experimental ones. If we add up the SiOH positive contribution to the SFG spectrum of the interface to the $\chi_{total}^{(2)}(\omega)$ of the water for each SAM/water interface, then the tail in each simulated DFT-MD spectrum in Fig. 3.7B will be reduced. The discrepancy at low frequency would then be removed. We believe that therefore the negative peak in the low-frequency range arising from SiOH groups of the silica surface is the possible contribution to the experimental SFG spectrum, compensating for the positive tail in the water OH SFG spectrum.

3.1.4 Perspectives of our metric: SFG integrals & H/V descriptor

Our H/V descriptor, that was shown in section 3.1.2 to be quantitatively correlated with spectroscopic descriptors for hydrophilicity quantification, can

further serve at more situations. The H/V descriptor presented in this chapter and in our paper [*On the trail of molecular hydrophilicity and hydrophobicity at aqueous interfaces, J. Phys. Chem. Lett.* 2023, 14, 1301-1309] is the statistically averaged value over the whole trajectory of each system and over the whole box. If we cut the trajectories into pieces, we obtain the time evolution of the interface hydrophilicity. The dynamics of the surface hydrophilicity revealed by the interfacial water network reorganization can hence be quantified by calculating the relaxation time of the H/V descriptor. And if we divide the box into grids, we can get the space-resolved H/V maps where the heterogeneity of the interface hydrophilicity and hydrophobicity can be unraveled.

However, a longer trajectory and a larger box beyond the DFT-MD limits are required for these developments. Hence, we go from DFT-MD simulations to classical MD simulations, which will be presented in the following section 3.3. Before section 3.3, we dissect the water organization in the BIL as well as the SAMs structures at the top-layer in the BIL in section 3.2.

We now report our paper[On the trail of molecular hydrophilicity and hydrophobicity at aqueous interfaces, J. Phys. Chem. Lett. 2023, 14, 1301-1309] where all results presented in this section are fully discussed.

On the Trail of Molecular Hydrophilicity and Hydrophobicity at Aqueous Interfaces

Wanlin Chen,* Stephanie E. Sanders, Burak Özdamar, Dorian Louaas, Flavio Siro Brigiano, Simone Pezzotti, Poul B. Petersen, and Marie-Pierre Gaigeot*



Cite This: *J. Phys. Chem. Lett.* 2023, 14, 1301–1309



Read Online

ACCESS |



Metrics & More

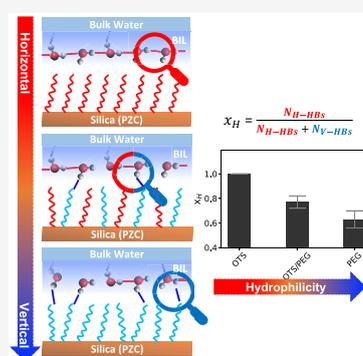


Article Recommendations



Supporting Information

ABSTRACT: Uncovering microscopic hydrophilicity and hydrophobicity at heterogeneous aqueous interfaces is essential as it dictates physico/chemical properties such as wetting, the electrical double layer, and reactivity. Several molecular and spectroscopic descriptors were proposed, but a major limitation is the lack of connections between them. Here, we combine density functional theory-based MD simulations (DFT-MD) and SFG spectroscopy to explore how interfacial water responds in contact with self-assembled monolayers (SAM) of tunable hydrophilicity. We introduce a microscopic metric to track the transition from hydrophobic to hydrophilic interfaces. This metric combines the H/V descriptor, a structural descriptor based on the preferential orientation within the water network in the topmost binding interfacial layer (BIL) and spectroscopic fingerprints of H-bonded and dangling OH groups of water carried by BIL-resolved SFG spectra. This metric builds a bridge between molecular descriptors of hydrophilicity/hydrophobicity and spectroscopically measured quantities and provides a recipe to quantitatively or qualitatively interpret experimental SFG signals.



Understanding liquid water organization at solid surfaces is of central importance in many phenomena, encompassing the ability of interfaces to catalyze chemical reactions,^{1,2} electric double layer formation and adsorption processes,^{3–5} the dielectric response of the interface,^{6,7} ultrafast energy transfer and vibrational dynamics,^{8,9} and surface acidity and conductivity.^{10–12} Following the termination of the bulk hydrogen-bonding (HB) network, the water structure is reorganized at an aqueous interface depending on the balance between water–surface and water–water interactions. Such balance dictates the hydrophilicity/hydrophobicity of an interface. It is subtly connected to the nature, morphology, and topology of the hydrated surfaces.¹³ These subtle yet fundamental connections cannot be captured by the conventional macroscopic metric—contact angle—due to the nanoscale structural heterogeneity and dynamics of extended interfaces.^{3,14,15} Therefore, there is the need for searching molecular signatures to comprehend microscopic hydrophobicity/hydrophilicity at the nanoscopic scale. Despite enormous progress, this remains a challenge for both theory and experiment.^{16–18}

In particular, from the experimental side, the surface-sensitive nonlinear optical technique vibrational sum-frequency generation (SFG) spectroscopy has been widely utilized to characterize aqueous interfaces at the molecular level,^{19–26} since the second-order susceptibility $\chi^{(2)}(\omega)$ is zero for centrosymmetric media (such as the bulk of liquids and solids). SFG spectroscopy can directly probe dangling OH groups at the interface, which are regarded as a key indicator of

molecular hydrophobicity,^{27–29} as they provide a clear fingerprint in the 3550–3800 cm^{-1} frequency range. Local hydrophobicity at macroscopically hydrophilic surfaces, such as heat-treated silica^{15,26} and 0001- α -alumina,^{29,30} was revealed by the detection of the SFG-active dangling OH groups. The ratio between dangling OH and H-bonded OH SFG bands has been recently correlated by Kim et al.³¹ to surface hydrophobicity through surface adhesion energy. However, the SFG signature of dangling OH groups provides little information on the actual structure and connectivity of the water HB network and does not allow us to quantitatively follow the transition in the interfacial water arrangement from hydrophobic to hydrophilic surfaces. This demands deeper insight into the specific water network formed at the interface, which can be obtained from molecular dynamics (MD) simulations.

Numerous microscopic descriptors of hydrophobicity have been developed in the last decades through MD simulations based on density fluctuations,^{32–34} bond interactions,³⁵ or orientation probes³⁶ of interfacial water. Godawat et al.³² demonstrated that large local density fluctuations and higher probabilities of low water density provide clear signatures of

Received: October 31, 2022

Accepted: January 27, 2023

hydrophobicity. Based on that, local compressibility has been developed as a powerful theoretical probe to map microscopic hydrophilicity.³⁴ Shin et al.³⁶ proposed another approach of probing the interfacial structure through a statistical analysis of the orientational configurations of interfacial water molecules. In a recent work, we have shown that the preferential orientation of HBs formed by water in the topmost interfacial layer can be used to determine the degree of local hydrophilicity/hydrophobicity.²⁹ All these descriptors manage to characterize and quantify the molecular hydrophobicity/hydrophilicity, but the connections between these theoretically determined descriptors and spectroscopic fingerprints are still missing.

Bridging between theoretical and spectroscopic fingerprints of molecular hydrophobicity/hydrophilicity is the aim of the present work. To this end, we investigate a set of aqueous interfaces where water is in contact with self-assembled monolayers (SAM) of well-controlled and tunable hydrophilicity. These interfaces have attracted considerable attention as a model platform for studying hydrophobicity/hydrophilicity and have been extensively characterized in a number of theoretical^{13,37,38} and SFG experimental^{39–42} works. We have carried out DFT-MD simulations for three different silane-based SAM/water interfaces (with SAMs anchored on a silica substrate as illustrated in Figure 1): a pure octadecyl-

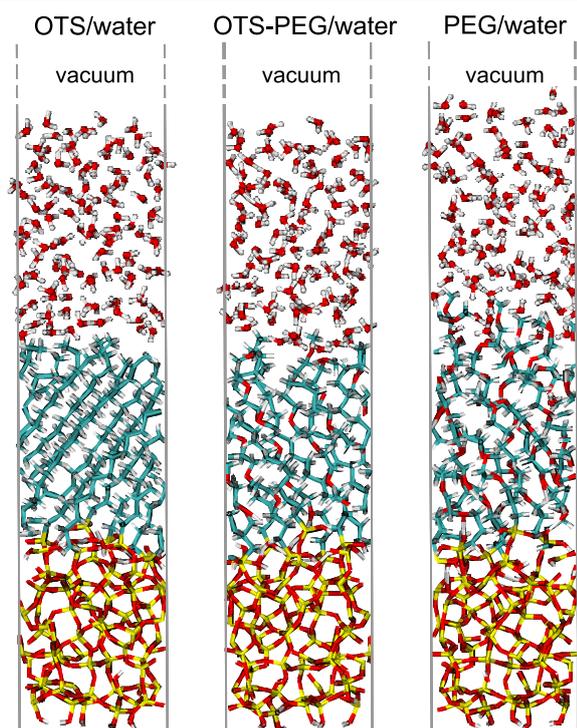


Figure 1. Snapshots of the three SAM/water interfaces from the DFT-MD simulations. From bottom to top: amorphous silica, SAMs, water, and vacuum (Si, yellow; C, cyan; O, red; and H, white). Box dimensions: $13.386 \times 13.286 \times 85 \text{ \AA}^3$.

trichlorosilane (OTS, $\text{CH}_3(-\text{CH}_2)_{17}-\text{SiCl}_3$) monolayer (intrinsically hydrophobic), a pure poly(ethylene glycol) (PEG, $\text{CH}_3(-\text{O}-\text{CH}_2-\text{CH}_2)_7-\text{CH}_2-\text{SiCl}_3$) monolayer (intrinsically hydrophilic), and a mixed OTS–PEG monolayer (1:1 mixing ratio), corresponding to the experimental samples prepared by Sanders et al.⁴² The water theoretical SFG spectra of the three

interfaces are calculated from the simulations and compared with the experimental measurements.⁴² The deconvolution of the water SFG spectra aids in elucidating the macroscopic and molecular response of the interfacial water with increasing hydrophilicity of the SAM surface. On the macroscopic scale, we evaluate the ability of the interface to screen the surface field arising from the charged silica + SAM surface, as obtained by introducing a mixed theoretical and experimental approach to extract the potential drop in the liquid from SFG spectra. On the molecular scale, we combine quantitative descriptors of hydrophilicity/hydrophobicity based on the preferential orientation within the water HB network in the topmost interfacial layer, hereafter named the H/V descriptor, and on the integration of the SFG intensity in the spectral regions containing the stretching motions of either H-bonded ($3000\text{--}3550 \text{ cm}^{-1}$) or dangling ($3550\text{--}3800 \text{ cm}^{-1}$) water OH groups. The synergic evaluation of these molecular and macroscopic descriptors allows us to establish a connection between the structural and spectroscopic fingerprints of molecular hydrophobicity and to map the transition of the water HB network from the most hydrophobic to the most hydrophilic interface. *Surface Potential Determination from SFG Spectra.* Figure 2A displays the DFT-MD calculated $\text{Im}\chi_{\text{ssp}}^{(2)}(\omega)$ spectra of water

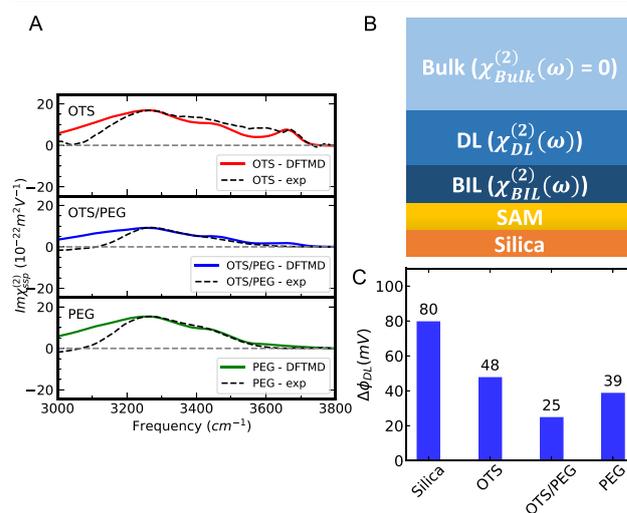


Figure 2. (A) SFG spectra of the three SAM aqueous interfaces (OTS/water, OTS–PEG/water, and PEG/water) in the ssp polarization with DFT-MD calculated spectra of $\text{Im}\chi_{\text{BIL}}^{(2)}(\omega)$ under PZC (pH 2–4) conditions plus the fitted $\text{Im}\chi_{\text{DL}}^{(2)}(\omega)$ contribution in colors (in absolute intensities) and experimental spectra measured by Sanders et al.⁴² at pH ~ 5.6 in dashed black (in a.u.). (B) Scheme of the deconvolution of water SFG spectra into signals arising from water in the binding interfacial layer (BIL) and in the diffuse layer (DL). (C) Potential differences across the liquid region (mV) for the three different systems at pH ~ 5.6 together with the reference value for the silica/water interface at pH ~ 6 .⁵

within the OH stretching region from 3000 to 3800 cm^{-1} for the three SAM/water interfaces illustrated in Figure 1. The theoretical spectra omit the (negative) contributions of SiOH groups (details in section S2 of the SI), which are also active in the same frequency range of the water O–H stretch but are negligible for frequency $>3200 \text{ cm}^{-1}$ according to previous studies.^{15,42,43} In the frequency range $>3200 \text{ cm}^{-1}$ (where only water contributes), the theoretical spectra (Figure 2A) are in good agreement with the corresponding experimental spectra

measured by Sanders et al.⁴² (including the correction for Fresnel factors; see section S1 of the SI). For all systems, the SFG spectra show a band centered at 3200 and a shoulder at 3400 cm⁻¹, with the intensities decreasing in the order water/OTS > water/PEG > water/mixed OTS and PEG. An extra peak at around 3660 cm⁻¹ corresponding to the signal of the free OH groups is seen at the OTS/water interface.

The free OH peak in the experimental OTS spectrum is combined with a shoulder at a slightly lower frequency (3600 cm⁻¹), leading to a seemingly double free OH band, whereas only one single free OH peak is observed in our theoretical SFG spectrum. The double free OH band observed in the experimental OTS spectrum shows that there is diversity in the local hydrophobicity of the surface, which could arise, for instance, from the local morphology.⁴⁴ This is not captured by our DFT-MD simulation. The diversity in the local hydrophobicity revealed by the experimental spectrum from Sander et al.⁴² might depend on how the monolayer is prepared experimentally. In some other SFG spectra of the OTS/water interface, e.g., the ones reported in refs 20 and 39, a sharp single free OH peak without any shoulder was observed, as in our DFT-MD spectrum.

The consistency between the experimental and the DFT-MD calculated SFG spectra validates our simulations, but the interpretation of the spectra in terms of the water structures interacting with SAMs is still not straightforward due to the following reason. With the silica substrate (below the SAM monolayers, see Figure 1) negatively charged in the experimental environment (pH ~5.6), the centrosymmetry of bulk-like water located further than the first few angstroms from the interface is broken by the surface electric field. That leads to a non-negligible surface-charge-induced contribution to the total SFG signal, which veils the spectrum arising from the specific interfacial structure (the binding interfacial layer, denoted as BIL).^{45,46} Therefore, the understanding of SFG spectra requires the separation of the spectroscopic contributions into a $\chi_{\text{BIL}}^{(2)}(\omega)$ contribution arising from the BIL, the layer directly in contact with the SAMs, and into the contribution from the DL (diffuse layer), $\chi_{\text{DL}}^{(2)}(\omega)$, the subsequent layer with bulk-like water reoriented by the surface charges (Figure 2B):^{45,46}

$$\chi^{(2)}(\omega) = \chi_{\text{BIL}}^{(2)}(\omega) + \chi_{\text{DL}}^{(2)}(\omega) \quad (1)$$

The methodology of the deconvolution of $\chi_{\text{BIL}}^{(2)}(\omega)$ and $\chi_{\text{DL}}^{(2)}(\omega)$ contributions based on the identification of BIL and DL according to three structural descriptors (density, HB network, and orientation) is presented in ref 46.

The $\text{Im}\chi_{\text{BIL}}^{(2)}(\omega)$ water spectra are directly calculated from the DFT-MD simulations of the three SAM/water interfaces under the PZC (point of zero charge) conditions (pH 2–4) where only water in the BIL is noncentrosymmetric and thus SFG-active in the OH stretching region (3000–3800 cm⁻¹). The calculation of the $\text{Im}\chi_{\text{DL}}^{(2)}(\omega)$ contribution at the experimental pH ~5.6 condition requires much thicker water slabs than the length scale affordable in the DFT-MD simulations. We introduce below a novel approach to circumvent such a limitation. This approach builds on the combination of theory and experiment and requires the following hypothesis for the $\text{Im}\chi_{\text{BIL}}^{(2)}(\omega)$ contribution. At the experimental pH ~5.6, the silica surface is only ~1% dehydroxylated.⁴⁷ Since the $\chi_{\text{BIL}}^{(2)}(\omega)$ of a charged interface with a low surface charge density σ (less than a few percent of a monolayer) is barely modified

compared to the neutral interfaces, as proved by Tian et al.,^{39,45} we assume $\text{Im}\chi_{\text{BIL}}^{(2)}(\omega)$ to be invariant from pH 2–4 to 5.6. The experimentally measured $\text{Im}\chi^{(2)}(\omega)$ is thus equivalent to the sum of the theoretical $\text{Im}\chi_{\text{BIL}}^{(2)}(\omega)$ of the fully hydroxylated surface and the charge-induced $\text{Im}\chi_{\text{DL}}^{(2)}(\omega)$ contribution.

The field-induced reorientation of water dominates the SFG signal of the DL, which thus contains the interfacial electrostatic information. As demonstrated in refs 43, 45, 46, and 48, $\chi_{\text{DL}}^{(2)}(\omega)$ is equal to the third-order susceptibility of liquid water, $\chi_{\text{bulk}}^{(3)}(\omega)$, multiplied by the potential difference across the DL region, $\Delta\phi_{\text{DL}}$, as expressed by

$$\chi_{\text{DL}}^{(2)}(\omega) = \chi_{\text{bulk}}^{(3)}(\omega)\Delta\phi_{\text{DL}} \quad (2)$$

where $\Delta\phi_{\text{DL}}$ is the potential drop across the diffuse layer,^{43,45,46,48} defined as the difference between the surface potential measured at the plane separating BIL and DL regions and that measured in the bulk. It is equal to the surface potential minus the potential drop across the BIL region (i.e., between the surface plane and the plane separating BIL and DL). Interference contributions are omitted from the above equation since they are negligible under the experimental condition (ionic strength >10⁻² M). $\chi_{\text{bulk}}^{(3)}(\omega)$ is a constant spectrum for any aqueous interface with two bands centered at 3200 and 3400 cm⁻¹, as demonstrated both experimentally⁴⁵ and theoretically.^{46,48} $\Delta\phi_{\text{DL}}$ solely determines the sign and magnitude of $\chi_{\text{DL}}^{(2)}(\omega)$. Although $\Delta\phi_{\text{DL}}$ is actually the surface potential across the DL region, it is usually regarded as the surface potential in the literature. The difference between the surface potential and $\Delta\phi_{\text{DL}}$ is determined by the screening within the BIL region. $\Delta\phi_{\text{DL}}$ can be extracted from the measured SFG intensities by using eq 2 once $\chi_{\text{bulk}}^{(3)}(\omega)$ and $\chi_{\text{DL}}^{(2)}(\omega)$ are known.

To achieve this, we combine the total experimental SFG intensity and the theoretical $\chi_{\text{BIL}}^{(2)}(\omega)$ spectrum. We start by making a simple consideration: the intensity of the SFG signal in the H-bonded region (3000–3550 cm⁻¹) is the result of the combination of $\chi_{\text{BIL}}^{(2)}(\omega)$ and $\chi_{\text{DL}}^{(2)}(\omega)$ contributions, while the signal at 3660 cm⁻¹ arising from the free OH groups solely contains the $\chi_{\text{BIL}}^{(2)}(\omega)$ part. With the calculated $\chi_{\text{BIL}}^{(2)}(\omega)$ intensity being fixed, the relative intensity of the free OH signal and the H-bonded band is hence a function of $\chi_{\text{DL}}^{(2)}(\omega)$ intensity, which in turns depends on $\Delta\phi_{\text{DL}}$ according to eq 2. Based on this, the experimental relative intensity of the two bands can be used as an objective function (OF) to determine $\Delta\phi_{\text{DL}}$ from a fitting procedure, as expressed by

$$\begin{aligned} \text{OF} &= \frac{\chi_{\text{exp}}^{(2)}(\omega_{\text{free OH}})}{\chi_{\text{exp}}^{(2)}(\omega_{\text{H bonded}})} \\ &= \frac{\chi_{\text{BIL}}^{(2)}(\omega_{\text{free OH}})}{\chi_{\text{BIL}}^{(2)}(\omega_{\text{H bonded}}) + \Delta\phi_{\text{DL}}\chi_{\text{bulk}}^{(3)}(\omega_{\text{H bonded}})} \quad (3) \end{aligned}$$

where $\chi_{\text{exp}}^{(2)}(\omega)$ is the experimental SFG spectrum, $\chi_{\text{BIL}}^{(2)}(\omega)$ and $\chi_{\text{bulk}}^{(3)}(\omega)$ are theoretical spectra calculated from the DFT-MD simulations, with $\omega_{\text{free OH}} = 3660$ cm⁻¹, and $\omega_{\text{H bonded}} = 3250$ cm⁻¹. This methodology makes it possible to extract the $\Delta\phi_{\text{DL}}$ of the interface directly from the experimentally measured SFG spectrum, without knowing its absolute intensity. The experimental spectrum (in a.u.) of the OTS/water interface is taken as a reference, which is scaled such that the maximum intensity is the same as that obtained for the theoretical

$\text{Im}\chi_{\text{BIL}}^{(2)}(\omega)$ spectrum after the fitting procedure. All other experimental spectra are normalized by the same scaling factor. The $\Delta\phi_{\text{DL}}$ values providing the best fit for all three interfaces are plotted in Figure 2C.

Compared to the reference $\Delta\phi_{\text{DL}}$ value (80 mV) at the silica/water interface at pH ~ 6 from ref 5, the obtained $\Delta\phi_{\text{DL}}$ values for the three SAM/water interfaces are well reduced. This indicates that the presence of monolayers between water and the silica substrate has a screening effect on the surface charge, with $\Delta\phi_{\text{DL}}$ decreasing in the order OTS (48 mV) > PEG (39 mV) > mixed OTS–PEG (25 mV). Intrinsically, PEG polymer chains screen better than OTS chains because of their larger dipole moments due to the presence of oxygen atoms. However, this still cannot explain the strongest screening effect of the mixed OTS–PEG monolayer. With the assumption that the different degrees of screening effects might also be related to the organization of the SAM chains, we refer to the results of Sanders et al.,⁴² where a qualitative investigation of the spatial order of the three monolayers is introduced by evaluating the relative intensity of methylene to methyl SFG bands. Their results show that the OTS monolayer is uniform and well ordered, whereas the PEG chains are significantly more disordered than the OTS due to the variable chain length of the commercial silane (six to nine PEG units), which allows the longer chains to lean over shorter chains. The mixed OTS–PEG monolayer is the most disordered monolayer, as expected, due to the implicit PEG disorder and the defects and disorders generated by the mixture of OTS and PEG. The decrease in the SAM chain order (OTS > PEG > OTS/PEG) is qualitatively correlated with the increase in the screening effect: the more disordered the monolayer, the more strongly the surface charge is reduced.

Another possible explanation can be thought of by considering that the screening comes from not only the monolayer but also the organization of BIL-water. BIL-water is horizontally ordered (2DN) above a hydrophobic surface^{29,49} and vertically ordered at a hydrophilic one.^{15,43,50} At the mixed interfaces with the coexistence of microscopic hydrophobic and hydrophilic patches, it is disordered with neither a stable 2DN nor a stable vertically ordered structure. The BIL-water organization at the three SAM–water interfaces is characterized in the next section.

Interpretation of $\chi_{\text{BIL}}^{(2)}(\omega)$. The specific interfacial structures of water adapting to various surfaces are revealed by the SFG signal of the BIL. Figure 3B shows the $\text{Im}\chi_{\text{BIL}}^{(2)}(\omega)$ spectra of the three different interfaces calculated from the DFT-MD simulations over ~ 400 ps for each system. The BIL spectrum of the OTS/water interface resembles that of the hydrophobic prototype—the air/water interface⁴⁹—with a positive peak at 3660 cm^{-1} , corresponding to the dangling OH groups pointing toward the monolayer as shown by green circles in Figure 3A on top and a broad negative band centered at around 3350 cm^{-1} . It is noted that the frequency of the free OH peak is lower than that of the air/water interface, which we ascribe to the van der Waals interactions between the dangling OH and the OTS monolayer as suggested by Tian et al.³⁹ A similar shift is also observed in the SFG spectra of water in contact with graphene and boron nitride (BN).²⁹ The mixed OTS–PEG/water SFG spectrum also presents a free OH peak and a broad negative band in the H-bonded region, but with strongly reduced intensities. The PEG spectrum unexpectedly shows an almost zero signal over the whole O–H stretching region. The

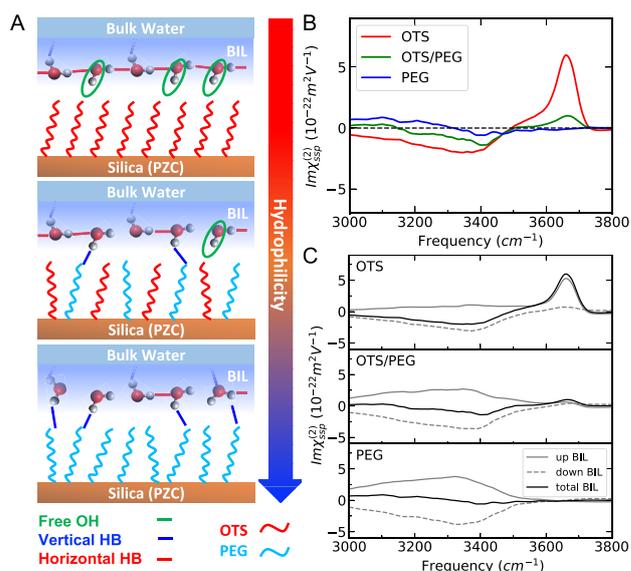


Figure 3. Characterizations of BIL water structures in contact with the three different monolayers: (A) Scheme of different types of OH groups of BIL water. (B) DFT-MD calculated $\text{Im}\chi_{\text{BIL}}^{(2)}(\omega)$ signals. (C) Deconvolved water-up and water-down BIL spectra.

disappearance of the free OH peak is reasonable, but the minuscule signal in the H-bonded region is astonishing.

As mentioned above, one strength of the DFT-MD SFG compared to the experimental one is that we can further deconvolve the BIL-water spectrum by calculating the spectral contributions arising from each BIL-water population identified from the simulations based on structural properties. Knowing that the normal in SFG is the vector perpendicular and pointing to the solid surface, we have deconvolved the BIL spectra into signals arising from water molecules with their dipole moments pointing toward the SAMs (along the normal, thus denoted as up) and the ones pointing toward the subsequent bulk water (opposite to the normal, thus down). The two deconvolved spectra are denoted as water-up and water-down, respectively, and displayed in Figure 3C.

Strikingly, the water-up and water-down spectra of the PEG/water interface are both broad bands centered at $\sim 3300\text{ cm}^{-1}$, with the same intensity and opposite sign, and the compensation of the two leads to an approximate zero signal in total. The same frequency of the water-up and water-down spectra indicates that the strength of H-bonds formed between the oxygen atoms of PEG and water in the BIL is the same as that of water–water H-bonds between BIL and bulk. For mixed OTS–PEG/water, we observe a similar compensation of water-up and water-down broad bands in the H-bonded region, which explains the very low intensity of the negative band seen in the total BIL spectrum compared to that in the OTS/water spectrum.

Spectroscopic and Structural Descriptors of Microscopic Hydrophilicity/Hydrophobicity. The free OH peak revealed by SFG spectroscopy is a well-recognized microscopic hydrophobic fingerprint.^{27–29} It is thus quantified here by calculating the $\text{Im}\chi_{\text{BIL}}^{(2)}(\omega)$ integral in the $3550\text{--}3800\text{ cm}^{-1}$ region for each SAM interface (Figure 4A, red bars). The integrals decrease with the increase in the number of PEG chains in the monolayer composition. Interestingly, the integral of the mixed OTS–PEG/water free OH signal is ~ 5 times smaller than that of the pure OTS/water interface, despite OTS and PEG being

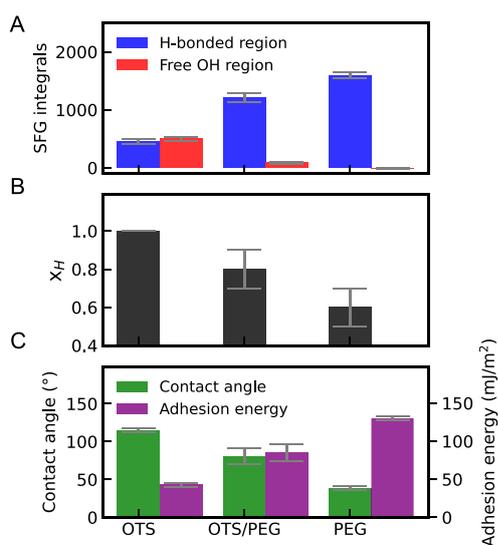


Figure 4. Hydrophobicity/hydrophilicity descriptors investigated in this work. Two microscopic descriptors determined from SFG signals and DFT-MD simulations: (A) SFG intensity integrals of the deconvolved water-up spectra in the spectral regions of the free OH ($3550\text{--}3800\text{ cm}^{-1}$) and H-bonded ($3000\text{--}3550\text{ cm}^{-1}$) OH stretching motions and (B) structural descriptor based on the preferential orientation within the water network in the binding interfacial layer (BIL)–H/V descriptor. Two macroscopic descriptors: (C) contact angles measured by Sanders et al.⁴² in green and the associated surface adhesion energies in purple (from ref 31 and discussed later in this text).

mixed in a 1:1 ratio (which is expected to result in a decrease by a factor of ~ 2). The reason for this oversuppression of the free OH SFG signal at the mixed SAM/water interface is topological: most of the water molecules in contact with OTS units point their OH groups toward nearby PEG chains, as they can interact with O atoms of PEGs, thus leaving only very few dangling OH groups. This result indicates that the hydrophilic sites dictate the water–surface interactions, and they can suppress molecular hydrophobicity (as measured by free OH groups) around the OTS chains for the uniform mixing of OTS and PEG. In other words, the spatial distribution of the hydrophobic/hydrophilic sites has a large influence on global hydrophobicity, as well documented in previous theoretical works.^{34,44,51–54} As a result of dangling OH suppression for the uniform mixing adopted in our simulation, the tiny free OH peak in the BIL-SFG spectrum of the mixed PEG-OTS/water system is hardly distinguishable from the PEG/water interface, both with an approximately zero free OH signal. It is thus difficult to microscopically quantify the degree of hydrophilicity of the SAM/water interfaces solely by the free OH signature in SFG.

We thus introduce a different descriptor based on the molecular details of the interfacial water structure that are accessible from DFT-MD simulations. Above a hydrophobic surface, water molecules in the BIL form preferential horizontal water–water H-bonds, which leads to a two-dimensional HB network (2DN) structure,^{29,49} whereas vertical ordered structures based on surface water H-bonds are favored above a hydrophilic surface.^{15,43,50} Therefore, apart from the free OH peak in the SFG spectrum, the hydrophilicity can be evaluated by quantifying the number of horizontal water–water H-bonds and the number of vertical surface water H-bonds formed in

the BIL. (See section S6 in the SI for the number of water molecules around each SAM chain.)

To that end, the first approach is to quantify the H-bonds by integrating the SFG signals in the H-bonded region from 3000 cm^{-1} to 3550 cm^{-1} . Since the horizontal H-bonded OH groups are SFG-inactive, the integral of the H-bonded signal in the SFG spectrum solely contains the vertical components of the H-bonded OH groups. However, as shown above, the signals arising from OH groups H-bonded to the O atoms of PEGs are washed out in the total BIL spectra because of the compensation between water-up and water-down contributions due to the same strength of PEG–water and water–water H-bonds. We hence choose to calculate the integral of the deconvolved water-up signal, in order to account for the water-surface H-bonds. The larger this integral, the more vertical H-bonds that are formed between BIL-water and the SAMs and hence the more hydrophilic the surface. The results are shown in Figure 4A in blue bars. As expected, the integrals of these H-bond signals (blue) are anticorrelated with those of the free OH signal (red). Furthermore, the difference in hydrophilicity between the mixed OTS–PEG/water and the PEG/water can now be captured by this molecular descriptor.

We complement such a spectroscopic descriptor with a structural one, namely, the H/V descriptor. To that end, we calculate the number of horizontal HBs (n_H , defined as water–water H-bonds in the BIL, parallel to the instantaneous water surface⁵⁵ in contact with SAMs $\pm 30^\circ$, red lines in Figure 3A) and the number of vertical HBs (n_V , defined as water–surface H-bonds,⁴⁹ blue lines in Figure 3A) from the DFT-MD simulations. We then define the molar fraction x_H of horizontal H-bonds as our H/V descriptor to quantify the hydrophobicity/hydrophilicity of the surface. Note that the choice of the Willard–Chandler instantaneous surface⁵⁵ used in our work allows us to naturally take into account the effects of local morphology in our descriptor.²⁹ The H/V descriptor is thus

$$\text{H/V} = x_H = \frac{n_H}{n_H + n_V} \quad (4)$$

The lower the x_H , the more hydrophilic the interface.

The x_H values calculated for the three SAM/water interfaces, statistically averaged over ~ 400 ps of trajectory for each system, are plotted in Figure 4B. The values are correlated with the integral values of the free OH SFG signal (Figure 4A, red bars) and are anticorrelated with the integral values of the H-bonded OH signal (Figure 4A, blue bars). The x_H values range from 1 for the pure OTS/water interface to 0.63 for the pure PEG/water interface. A value of 0.77 is obtained for the mixed OTS–PEG/water interface in between the two extreme values. The x_H for the mixed OTS–PEG/water interface (0.77) is closer to that of the PEG/water interface (0.63) compared to that of the OTS/water interface (1.0). This confirms the dominating effect of the hydrophilic sites on the global hydrophobicity of the interface implied by the large decrease in the free OH peak intensity compared to the OTS/water interface as discussed above.

Both H/V and SFG descriptors provide a quantitative scale to rank surface hydrophilicity, which can be used in the same way as the contact angle but for molecular hydrophobicity instead of macroscopic hydrophobicity.

We now demonstrate the physics behind the correlation between our structural and spectroscopic descriptors. The BIL-SFG signal, arising from a sum of contributions from each BIL-water OH group, can be deconvolved into the sum of

contributions from BIL–OH groups oriented horizontally (H) and vertically (V) to the surface through the following equation:

$$\chi_{\text{BIL}}^{(2)}(\omega) = \chi_{\text{H}}^{(2)}(\omega) + \chi_{\text{V}}^{(2)}(\omega) \quad (5)$$

The fraction of horizontal and vertical H-bonds in the BIL (i.e., x_{H} and x_{V}) are quantitatively correlated with the SFG signals in the H-bonded frequency range (3000–3550 cm^{-1})

$$\chi_{\text{BIL}}^{(2)}(\omega) = \frac{n_{\text{OH}}}{S}(x_{\text{H}}\bar{\chi}_{\text{H}}^{(2)}(\omega) + x_{\text{V}}\bar{\chi}_{\text{V}}^{(2)}(\omega)) \quad (6)$$

where $x_{\text{V}} = n_{\text{V}} / (n_{\text{H}} + n_{\text{V}}) = 1 - x_{\text{H}}$, $\omega_{\text{HB}} = 3000 - 3550 \text{ cm}^{-1}$, $\frac{n_{\text{OH}}}{S}$ is the density of OH groups per surface area ($/\text{nm}^2$), n_{H} and n_{V} are respectively the number of horizontal (water–water) and vertical (surface water) H-bonds in the BIL, and $\bar{\chi}_{\text{H}}^{(2)}$ and $\bar{\chi}_{\text{V}}^{(2)}$ are averaged SFG signals arising from horizontal and vertical OHs in the BIL, respectively. Noting that $\bar{\chi}_{\text{H}}^{(2)}(\omega) = 0$ due to the horizontal OH groups being perpendicular to the SFG normal vector, the relationship between integrals of SFG signals in the H-bonded OH stretch region and the H/V descriptor can be expressed as

$$I_{\text{BIL}} = I_{\text{V}} = \frac{n_{\text{OH}}}{S}x_{\text{V}}\bar{I}_{\text{V}} = \frac{n_{\text{OH}}}{S}(1 - x_{\text{H}})\bar{I}_{\text{V}} \quad (7)$$

where I_{BIL} and I_{V} are the spectral integrals in the H-bonded region (from 3000 to 3550 cm^{-1}) of the SFG signals arising from all water OHs in the BIL and of the vertical water OH groups in the BIL, respectively. \bar{I}_{V} is the averaged SFG integral of the vertical OH groups.

This linear equation between I_{BIL} and x_{H} quantifies the proportionality between these two molecular descriptors. The linear correlation plot between x_{H} and I_{BIL} calculated based on the DFT-MD simulations of all three SAM/water systems is shown in section S3 of the SI.

The values of the macroscopic hydrophilicity descriptor—contact angle—measured by Sanders et al.⁴² (Figure 4C, green) show in general the same trend as the two DFT-MD and SFG spectral-based microscopic metrics. The same trend is observed for another macroscopic descriptor, namely, surface adhesion energy, which can be derived from the contact angle θ_{WCA} through $^{31}W_{\text{ad}} = \gamma(1 + \cos \theta_{\text{WCA}})$, where γ is the surface tension of water (73 mJ/m^2). The increasing surface adhesion energy plotted in purple in Figure 4C indicates the increase in surface hydrophilicity, where 103.85 mJ/m^2 is the boundary value between hydrophobic and hydrophilic interfaces.³¹ The contact angles range from 114.7° (pure OTS, hydrophobic) to 38.5° (pure PEG, hydrophilic), but the reported value of a mixed OTS–PEG/water sample (80.5°) is closer to that of the pure OTS than to the pure PEG, contrary to our molecular descriptors.

The more hydrophilic behavior of our simulated mixed OTS–PEG/water interface reiterates the relevance of topological patterns over the surface: the homogeneous distribution of OTS and PEG chains imposed in our DFT-MD simulation is an extreme case of the perfectly uniform mixing of the two types of chains. In this scenario, our H/V and spectroscopic descriptors show that the hydrophobicity of OTS chains is masked by their local environment, as they are surrounded by hydrophilic PEG chains and water preferentially interacts with the latter. Our finding nicely agrees with previous results on the effect of the distribution of hydrophobic and hydrophilic functional groups on surface hydrophobicity by means of other

structural descriptors, such as an analysis of water density^{44,51} or orientational³⁶ fluctuations. The effects of local topology and morphology are hence encoded in our H/V-SFG metrics (see also section S4 of the SI). Contrary to the simulated scenario, the mixed SAM surface in the experiments is expected to exhibit a less uniform mixing of OTS and PEG, which must result in an increased hydrophobicity, as confirmed on a macroscopic scale by contact angle and surface adhesion energy measurements.

Local hydrophobicity plays an essential role in the chemical and physical properties of heterogeneous aqueous interfaces. It is usually experimentally probed by SFG spectroscopy following the signature provided by the water dangling OH groups pointing toward a hydrophobic surface. Although powerful and straightforward to use, this fingerprint alone fails to capture the collective structural organization of the water HB network as a function of surface hydrophilicity, which ultimately dictates interfacial properties. In this work, we combined DFT-MD simulations and SFG spectroscopy to go deeper into this molecular complexity. We found that the orientation of the water HB network in the topmost binding interfacial layer (BIL) well describes the collective molecular organization of the liquid water when it adapts to surfaces. In particular, we have shown that such orientation describes well the changes in the HB network when going from a horizontally (H) ordered HB network formed between BIL water molecules in the most hydrophobic environments to a vertical (V) ordering dominated by HBs between a hydrophilic surface and BIL water molecules. These structural changes have a direct signature in the SFG spectra since $\chi^{(2)}(\omega)$ is sensitive to molecular orientation, while horizontally oriented water molecules have negligible activity. Therefore, the SFG intensity of the BIL in the OH stretching region directly measures the degree of H/V ordering and thus local hydrophobicity/hydrophilicity.

However, quantitatively evaluating these changes from the total SFG spectra of charged interfaces, such as the SAM/water interfaces presented here, can be challenging since it requires deconvolution of the spectrum of the BIL from that of the subsequent DL, where water is reoriented by the static field generated by a charged surface. Moreover, contributions from vertically oriented water molecules in the BIL can cancel out if water molecules pointing toward the surface and toward the bulk form HBs with similar strength, as was the case for the PEG/water interface. For these reasons, quantitative estimations of surface hydrophilicity from SFG spectra require theoretical SFG spectroscopy, as it naturally provides the deconvolution of all of these contributions. By combining the experimental and deconvolved theoretical SFG spectra, we showed here that the structural H/V descriptor, the BIL SFG intensity, and macroscopic contact angle measurements were all correlated for the investigated SAM/water interfaces, giving a picture of local surface hydrophobicity/hydrophilicity that extends from the molecular to the macroscopic range.

In cases where theoretical SFG spectroscopy is not available, our work still provides a recipe to experimentally assess molecular hydrophobicity and hydrophilicity qualitatively. That requires $\chi_{\text{BIL}}^{(2)}(\omega)$ to be separated from the total SFG spectrum, which can be achieved by adopting experimental state-of-the-art deconvolution schemes.^{45,56–58} For a given surface, local hydrophobicity and hydrophilicity can be efficiently tuned by playing with parameters such as the pretreatment of the surface, the applied potential (for

electrochemical interfaces),^{6,57,59} the pH,⁵ and the ionic strength.^{56,58} In these cases, changes in the total SFG intensity in the OH stretching region of the BIL spectrum can be directly used to qualitatively identify and interpret changes in surface hydrophilicity. To give a practical example, the addition of salt to the amorphous silica/water interface was shown by DFT-MD simulations to promote the in-plane organization of the BIL water network,⁵ therefore increasing the hydrophobicity of the interface. This was indeed experimentally captured by the decrease in the integrations of the BIL SFG spectra at pH \sim 2 (PZC) with increasing ionic strength.⁵ In the future, our metric that connects the H/V structural descriptor to the integrals of $\chi_{\text{BIL}}^{(2)}(\omega)$ in the H-bonded region can be broadly applied to organic and oxide interfaces. It can be used to detect local hydrophobicity/hydrophilicity from SFG experiments, either qualitatively (experiments alone) or quantitatively by combining with theoretical SFG spectroscopy.

■ COMPUTATIONAL METHODS

Born–Oppenheimer DFT-MD simulations are carried out with the CP2K package,^{60,61} adopting the BLYP^{62,63} electronic representation with Grimme D2 corrections^{64,65} for dispersion, GTH pseudopotentials, and a combined plane wave (400 Ry density cutoff) and DZVP-SR MOLOPT basis set. Newton's equations of motions are integrated through the velocity Verlet algorithm for the nuclei with a 0.4 fs time step. In our previous work, we could show that the chosen approach, BLYP+D2, reproduces with great accuracy the structural and spectroscopic properties in comparison with SFG experiments. We benchmarked on a variety of solid/liquid and liquid/vapor interfaces, encompassing water in contact with air, graphene, BN, quartz, amorphous silica, and alumina.^{11,29,43,46,49,66} The simulation boxes of $13.386 \times 13.286 \times 85.0 \text{ \AA}^3$ are each composed of eight SAM chains (corresponding to the experimental samples in ref 42) anchored to an amorphous silica surface with a hydroxylation degree of 4.5 SiOH/nm^2 , solvated by 120 water molecules. The amorphous silica model is taken from Ugliengo et al.,⁶⁷ and it can accommodate up to 8 SiOH groups at the surface under PZC conditions. See refs 15 and 43. The eight SiOH groups are replaced by eight polymeric chains. Three different systems are built with eight OTS chains (pure OTS/water), alternately mixing four OTS and four PEG chains (mixed OTS–PEG/water) and eight PEG chains (pure PEG/water), respectively. Each box is periodically repeated in the x and y directions and separated by a vacuum layer of 25 \AA from the replica in the vertical z direction. For each system, four individual trajectories are accumulated with initial configurations extracted from different geometry optimization steps. The DFT-MD simulations are composed of 10 ps of equilibration for each trajectory before the production run of 100 ps, both in the NVE ensemble (with velocity rescaling and a target temperature of 300 K in the equilibration part only). Additional DFT-MD simulations were performed after equilibration from classical MD to check that our results are not affected by the chosen initial configurations (sections S4 and S5 in the SI).

The SFG signal, coming from the imaginary component of the total resonant electric dipole nonlinear susceptibility $\chi^{(2)}(\omega)$, calculated following the time-dependent approach of Morita et al.^{40,68} For the water contribution to $\chi^{(2)}(\omega)$, individual molecular dipole moments and polarizability tensors are calculated with the model from ref 69 assuming that only

the OH stretching motions contribute to the spectrum in the high-frequency region ($>3000 \text{ cm}^{-1}$). All details of derivation for the expression of SFG can be found in refs 43 and 46 with the parametrization of APT and Raman tensors of water found in ref 69. Each SFG spectrum is calculated from an average over ~ 400 ps of DFT-MD.

For all structural analyses, we have adopted the H-bond definition proposed by White and co-workers,⁷⁰ with $\text{O}(-\text{H})\cdots\text{O} \leq 3.2 \text{ \AA}$ and the $\text{O}-\text{H}\cdots\text{O}$ angle between 140 and 220° .

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcllett.2c03300>.

Fresnel factor corrections for experimental SFG spectra, SFG signal of SiOH, linear correlation plot of x_{H} and SFG descriptors, spatially resolved H/V descriptor map, additional simulation details and results, and number of BIL water molecules per alkyl chain in each investigated system (PDF)

Transparent Peer Review report available (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Wanlin Chen – *Université Paris-Saclay, Université Evry, CNRS, LAMBE UMR8587, 91025 Evry-Courcouronnes, France*; orcid.org/0000-0003-4933-4530;
Email: wanlin.chen@univ-evry.fr

Marie-Pierre Gaigeot – *Université Paris-Saclay, Université Evry, CNRS, LAMBE UMR8587, 91025 Evry-Courcouronnes, France*; orcid.org/0000-0002-3409-5824; Email: mgaigeot@univ-evry.fr

Authors

Stephanie E. Sanders – *Department of Chemistry and Biochemistry, Ruhr University Bochum, 44801 Bochum, Germany*

Burak Özdamar – *Université Paris-Saclay, Université Evry, CNRS, LAMBE UMR8587, 91025 Evry-Courcouronnes, France*

Dorian Louaas – *Université Paris-Saclay, Université Evry, CNRS, LAMBE UMR8587, 91025 Evry-Courcouronnes, France*

Flavio Siro Brigiano – *Université Paris-Saclay, Université Evry, CNRS, LAMBE UMR8587, 91025 Evry-Courcouronnes, France; Laboratoire de Chimie Théorique, Sorbonne Université, UMR 7616 CNRS, 75005 Paris, France*

Simone Pezzotti – *Université Paris-Saclay, Université Evry, CNRS, LAMBE UMR8587, 91025 Evry-Courcouronnes, France; Department of Physical Chemistry II, Ruhr University Bochum, D-44801 Bochum, Germany*; orcid.org/0000-0003-2023-3648

Poul B. Petersen – *Department of Chemistry and Biochemistry, Ruhr University Bochum, 44801 Bochum, Germany*; orcid.org/0000-0002-9821-700X

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jpcllett.2c03300>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was performed under PRACE (Partnership for Advanced Computing in Europe) under no. 2018194688. We acknowledge PRACE for awarding access to the Fenix infrastructures at GENCI-TGCC-France, which are partially funded from the European Union's Horizon 2020 research and innovation program through the ICEI project under grant agreement no. 800858. We also acknowledge HPC resources from GENCI-France grant 072484 (CINES/IDRIS/TGCC). S.P. acknowledges financial support by European Research Council (ERC) Advanced Grant 695437 THz-Calorimetry. S.P. and P.B.P. acknowledge the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC2033-390677874-RESOLV for funding this study.

REFERENCES

- (1) Zhu, J.; Hu, L.; Zhao, P.; Lee, L. Y. S.; Wong, K.-Y. Recent advances in electrocatalytic hydrogen evolution using nanoparticles. *Chem. Rev.* **2020**, *120*, 851–918.
- (2) Brigiano, F. S.; Gierada, M.; Tielens, F.; Pietrucci, F. Mechanism and free-energy landscape of peptide bond formation at the silica-water interface. *ACS Catal.* **2022**, *12*, 2821–2830.
- (3) Serva, A.; Salanne, M.; Havenith, M.; Pezzotti, S. Size dependence of hydrophobic hydration at electrified gold/water interfaces. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*, e2023867118.
- (4) Alfarano, S. R.; Pezzotti, S.; Stein, C. J.; Lin, Z.; Sebastiani, F.; Funke, S.; Hoberg, C.; Kolling, I.; Ma, C. Y.; Mauelshagen, K. Stripping away ion hydration shells in electrical double-layer formation: Water networks matter. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*, e2108568118.
- (5) Tuladhar, A.; Dewan, S.; Pezzotti, S.; Brigiano, F. S.; Creazzo, F.; Gageot, M.-P.; Borguet, E. Ions tune interfacial water structure and modulate hydrophobic interactions at silica surfaces. *J. Am. Chem. Soc.* **2020**, *142*, 6991–7000.
- (6) Montenegro, A.; Dutta, C.; Mammetkuliev, M.; Shi, H.; Hou, B.; Bhattacharyya, D.; Zhao, B.; Cronin, S. B.; Benderskii, A. V. Asymmetric response of interfacial water to applied electric fields. *Nature* **2021**, *594*, 62–65.
- (7) Qi, C.; Zhu, Z.; Wang, C.; Zheng, Y. Anomalously low dielectric constant of ordered interfacial water. *J. Phys. Chem. Lett.* **2021**, *12*, 931–937.
- (8) DelloStritto, M. J.; Piontek, S. M.; Klein, M. L.; Borguet, E. Effect of functional and electron correlation on the structure and spectroscopy of the Al₂O₃ (001)-H₂O interface. *J. Phys. Chem. Lett.* **2019**, *10*, 2031–2036.
- (9) Piontek, S. M.; Borguet, E. Vibrational dynamics at aqueous-mineral interfaces. *J. Phys. Chem. C* **2022**, *126*, 2307–2324.
- (10) Giberti, F.; Hassanali, A. A. The excess proton at the air-water interface: The role of instantaneous liquid interfaces. *J. Chem. Phys.* **2017**, *146*, 244703.
- (11) Pezzotti, S.; Gageot, M.-P. Spectroscopic BIL-SFG invariance hides the chaotic effect of protons at the air-water interface. *Atmosphere* **2018**, *9*, 396.
- (12) Grosjean, B.; Bocquet, M.-L.; Vuilleumier, R. Versatile electrification of two-dimensional nanomaterials in water. *Nat. Commun.* **2019**, *10*, 1656.
- (13) Monroe, J.; Barry, M.; DeStefano, A.; Aydogan Gokturk, P.; Jiao, S.; Robinson-Brown, D.; Webber, T.; Crumlin, E. J.; Han, S.; Shell, M. S. Water structure and properties at hydrophilic and hydrophobic surfaces. *Annu. Rev. Chem. Biomol.* **2020**, *11*, 523–557.
- (14) Limmer, D. T.; Willard, A. P.; Madden, P.; Chandler, D. Hydration of metal surfaces can be dynamically heterogeneous and hydrophobic. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 4200–4205.
- (15) Cyran, J. D.; Donovan, M. A.; Vollmer, D.; Brigiano, F. S.; Pezzotti, S.; Galimberti, D. R.; Gageot, M.-P.; Bonn, M.; Backus, E. H. Molecular hydrophobicity at a macroscopically hydrophilic surface. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 1520–1525.
- (16) Galli, G. Dissecting hydrophobicity. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 2557–2558.
- (17) Li, J.-L.; Car, R.; Tang, C.; Wingreen, N. S. Hydrophobic interaction and hydrogen-bond network for a methane pair in liquid water. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 2626–2630.
- (18) Rana, K.; Boyd, S. A.; Teppen, B. J.; Li, H.; Liu, C.; Johnston, C. T. Probing the microscopic hydrophobicity of smectite surfaces. A vibrational spectroscopic study of dibenzo-p-dioxin sorption to smectite. *Phys. Chem. Chem. Phys.* **2009**, *11*, 2976–2985.
- (19) Shen, Y. *Fundamentals of Sum-Frequency Spectroscopy*; Cambridge University Press, 2016.
- (20) Du, Q.; Superfine, R.; Freysz, E.; Shen, Y. Vibrational spectroscopy of water at the vapor/water interface. *Phys. Rev. Lett.* **1993**, *70*, 2313.
- (21) Eftekhari-Bafrooei, A.; Borguet, E. Effect of hydrogen-bond strength on the vibrational relaxation of interfacial water. *J. Am. Chem. Soc.* **2010**, *132*, 3756–3761.
- (22) Zhang, Z.; Piatkowski, L.; Bakker, H. J.; Bonn, M. Ultrafast vibrational energy transfer at the water/air interface revealed by two-dimensional surface vibrational spectroscopy. *Nat. Chem.* **2011**, *3*, 888–893.
- (23) Hsieh, C.-S.; Okuno, M.; Hunger, J.; Backus, E. H.; Nagata, Y.; Bonn, M. Aqueous heterogeneity at the air/water interface revealed by 2D-HD-SFG spectroscopy. *Angew. Chem., Int. Ed.* **2014**, *53*, 8146–8149.
- (24) Tian, C.; Shen, Y. Recent progress on sum-frequency spectroscopy. *Surf. Sci. Rep.* **2014**, *69*, 105–131.
- (25) Jubb, A. M.; Hua, W.; Allen, H. C. Environmental chemistry at vapor/water interfaces: insights from vibrational sum frequency generation spectroscopy. *Annu. Rev. Phys. Chem.* **2012**, *63*, 107–130.
- (26) Dalstein, L.; Potapova, E.; Tyrode, E. The elusive silica/water interface: isolated silanols under water as revealed by vibrational sum frequency spectroscopy. *Phys. Chem. Chem. Phys.* **2017**, *19*, 10343–10349.
- (27) Tang, F.; Ohto, T.; Sun, S.; Rouxel, J. R.; Imoto, S.; Backus, E. H.; Mukamel, S.; Bonn, M.; Nagata, Y. Molecular structure and modeling of water-air and ice-air interfaces monitored by sum-frequency generation. *Chem. Rev.* **2020**, *120*, 3633–3667.
- (28) Du, Q.; Freysz, E.; Shen, Y. R. Surface vibrational spectroscopic studies of hydrogen bonding and hydrophobicity. *Science* **1994**, *264*, 826–828.
- (29) Pezzotti, S.; Serva, A.; Sebastiani, F.; Brigiano, F. S.; Galimberti, D. R.; Potier, L.; Alfarano, S.; Schwaab, G.; Havenith, M.; Gageot, M.-P. Molecular fingerprints of hydrophobicity at aqueous interfaces from theory and vibrational spectroscopies. *J. Phys. Chem. Lett.* **2021**, *12*, 3827–3836.
- (30) DelloStritto, M.; Piontek, S. M.; Klein, M. L.; Borguet, E. Relating interfacial order to sum frequency generation with Ab initio simulations of the aqueous Al₂O₃ (0001) and (1120) interfaces. *J. Phys. Chem. C* **2018**, *122*, 21284–21294.
- (31) Kim, D.; Kim, E.; Park, S.; Kim, S.; Min, B. K.; Yoon, H. J.; Kwak, K.; Cho, M. Wettability of graphene and interfacial water structure. *Chem.* **2021**, *7*, 1602–1614.
- (32) Godawat, R.; Jamadagni, S. N.; Garde, S. Characterizing hydrophobicity of interfaces by using cavity formation, solute binding, and water correlations. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 15119–15124.
- (33) Patel, A. J.; Varilly, P.; Chandler, D. Fluctuations of water near extended hydrophobic and hydrophilic surfaces. *J. Phys. Chem. B* **2010**, *114*, 1632–1637.
- (34) Acharya, H.; Vembanur, S.; Jamadagni, S. N.; Garde, S. Mapping hydrophobicity at the nanoscale: Applications to heterogeneous surfaces and proteins. *Faraday Discuss.* **2010**, *146*, 353–365.
- (35) Gómez, S.; Rojas-Valencia, N.; Gómez, S. A.; Cappelli, C.; Merino, G.; Restrepo, A. A molecular twist on hydrophobicity. *Chem. Sci.* **2021**, *12*, 9233–9245.

- (36) Shin, S.; Willard, A. P. Characterizing hydration properties based on the orientational structure of interfacial water molecules. *J. Chem. Theory Comput.* **2018**, *14*, 461–465.
- (37) Monroe, J. I.; Shell, M. S. Computational discovery of chemically patterned surfaces that effect unique hydration water dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 8093–8098.
- (38) Dallin, B. C.; Van Lehn, R. C. Spatially heterogeneous water properties at disordered surfaces decrease the hydrophobicity of nonpolar self-assembled monolayers. *J. Phys. Chem. Lett.* **2019**, *10*, 3991–3997.
- (39) Tian, C.; Shen, Y. Structure and charging of hydrophobic material/water interfaces studied by phase-sensitive sum-frequency vibrational spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 15148–15153.
- (40) Morita, A.; Ishiyama, T. Recent progress in theoretical analysis of vibrational sum frequency generation spectroscopy. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5801–5816.
- (41) Ye, S.; Nihonyanagi, S.; Uosaki, K. Sum frequency generation (SFG) study of the pH-dependent water structure on a fused quartz surface modified by an octadecyltrichlorosilane (OTS) monolayer. *Phys. Chem. Chem. Phys.* **2001**, *3*, 3463–3469.
- (42) Sanders, S. E.; Petersen, P. B. Heterodyne-detected sum frequency generation of water at surfaces with varying hydrophobicity. *J. Chem. Phys.* **2019**, *150*, 204708.
- (43) Pezzotti, S.; Galimberti, D. R.; Gaigeot, M.-P. Deconvolution of BIL-SFG and DL-SFG spectroscopic signals reveals order/disorder of water at the elusive aqueous silica interface. *Phys. Chem. Chem. Phys.* **2019**, *21*, 22188–22202.
- (44) Jamadagni, S. N.; Godawat, R.; Garde, S. Hydrophobicity of proteins and interfaces: Insights from density fluctuations. *Annu. Rev. Chem. Biomol.* **2011**, *2*, 147–171.
- (45) Wen, Y.-C.; Zha, S.; Liu, X.; Yang, S.; Guo, P.; Shi, G.; Fang, H.; Shen, Y. R.; Tian, C. Unveiling microscopic structures of charged water interfaces by surface-specific vibrational spectroscopy. *Phys. Rev. Lett.* **2016**, *116*, 016101.
- (46) Pezzotti, S.; Galimberti, D. R.; Shen, Y. R.; Gaigeot, M.-P. Structural definition of the BIL and DL: A new universal methodology to rationalize non-linear $\chi^{(2)}(\omega)$ SFG signals at charged interfaces, including $\chi^{(3)}(\omega)$ contributions. *Phys. Chem. Chem. Phys.* **2018**, *20*, 5190–5199.
- (47) Dove, P. M. The dissolution kinetics of quartz in sodium chloride solutions at 25 degrees to 300 degrees C. *Am. J. Sci.* **1994**, *294*, 665–712.
- (48) Joutsuka, T.; Hirano, T.; Sprik, M.; Morita, A. Effects of third-order susceptibility in sum frequency generation spectra: a molecular dynamics study in liquid water. *Phys. Chem. Chem. Phys.* **2018**, *20*, 3040–3053.
- (49) Pezzotti, S.; Galimberti, D. R.; Gaigeot, M.-P. 2D H-bond network as the topmost skin to the air-water interface. *J. Phys. Chem. Lett.* **2017**, *8*, 3133–3141.
- (50) Sulpizi, M.; Gaigeot, M.-P.; Sprik, M. The silica–water interface: how the silanols determine the surface acidity and modulate the water properties. *J. Comput. Chem.* **2012**, *8*, 1037–1047.
- (51) Rego, N. B.; Xi, E.; Patel, A. J. Identifying hydrophobic protein patches to inform protein interaction interfaces. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*, e2018234118.
- (52) Adams, E. M.; Pezzotti, S.; Ahlers, J.; Rüttermann, M.; Levin, M.; Goldenzweig, A.; Peleg, Y.; Fleishman, S. J.; Sagi, I.; Havenith, M. Local mutations can serve as a game changer for global protein solvent interaction. *JACS Au* **2021**, *1*, 1076–1085.
- (53) Xi, E.; Venkateshwaran, V.; Li, L.; Rego, N.; Patel, A. J.; Garde, S. Hydrophobicity of proteins and nanostructured solutes is governed by topographical and chemical context. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 13345–13350.
- (54) Giovambattista, N.; Debenedetti, P. G.; Rosky, P. J. Hydration behavior under confinement by nanoscale surfaces with patterned hydrophobicity and hydrophilicity. *J. Phys. Chem. C* **2007**, *111*, 1323–1332.
- (55) Willard, A. P.; Chandler, D. Instantaneous liquid interfaces. *J. Phys. Chem. B* **2010**, *114*, 1954–1958.
- (56) Urashima, S.-h.; Myalitsin, A.; Nihonyanagi, S.; Tahara, T. The topmost water structure at a charged silica/aqueous interface revealed by heterodyne-detected vibrational sum frequency generation spectroscopy. *J. Phys. Chem. Lett.* **2018**, *9*, 4109–4114.
- (57) Wang, H.; Xu, Q.; Liu, Z.; Tang, Y.; Wei, G.; Shen, Y. R.; Liu, W.-T. Gate-controlled sum-frequency vibrational spectroscopy for probing charged oxide/water interfaces. *J. Phys. Chem. Lett.* **2019**, *10*, 5943–5948.
- (58) Backus, E. H.; Schaefer, J.; Bonn, M. Probing the mineral-water interface with nonlinear optical spectroscopy. *Angew. Chem., Int. Ed.* **2021**, *60*, 10482–10501.
- (59) Zhang, Y.; de Aguiar, H. B.; Hynes, J. T.; Laage, D. Water structure, dynamics, and sum-frequency generation spectra at electrified graphene interfaces. *J. Phys. Chem. Lett.* **2020**, *11*, 624–631.
- (60) Hutter, J.; Iannuzzi, M.; Schiffrmann, F.; VandeVondele, J. cp2k: atomistic simulations of condensed matter systems. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 15–25.
- (61) 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.
- (62) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098.
- (63) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785.
- (64) Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **2004**, *25*, 1463–1473.
- (65) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (66) Pezzotti, S.; Galimberti, D. R.; Shen, Y. R.; Gaigeot, M.-P. What the diffuse layer (DL) reveals in non-linear SFG spectroscopy. *Minerals* **2018**, *8*, 305.
- (67) Ugliengo, P.; Sodupe, M.; Musso, F.; Bush, I.; Orlando, R.; Dovesi, R. Realistic models of hydroxylated amorphous silica surfaces and MCM-41 mesoporous material simulated by large-scale periodic B3LYP calculations. *Adv. Mater.* **2008**, *20*, 4579–4583.
- (68) Morita, A.; Hynes, J. T. A theoretical analysis of the sum frequency generation spectrum of the water surface. II. Time-dependent approach. *J. Phys. Chem. B* **2002**, *106*, 673–685.
- (69) Khatib, R.; Backus, E. H.; Bonn, M.; Perez-Haro, M.-J.; Gaigeot, M.-P.; Sulpizi, M. Water orientation and hydrogen-bond structure at the fluorite/water interface. *Sci. Rep.* **2016**, *6*, 24287–1.
- (70) White, J. A.; Schwegler, E.; Galli, G.; Gygi, F. The solvation of Na⁺ in water: First-principles simulations. *J. Chem. Phys.* **2000**, *113*, 4668–4673.

Supporting Information for: On the Trail of Molecular Hydrophilicity and Hydrophobicity at Aqueous Interfaces

Wanlin Chen,[†] Stephanie E. Sanders,[‡] Burak Özdamar,[†] Dorian Louaas,[†] Flavio
Siro Brigiano,^{¶,†} Simone Pezzotti,^{§,†} Poul B. Petersen,[‡] and Marie-Pierre
Gaijeot^{*,†}

[†]*Université Paris-Saclay, Univ Evry, CNRS, LAMBE UMR8587, 91025
Evry-Courcouronnes, France*

[‡]*Department of Chemistry and Biochemistry, Ruhr University Bochum, 44801 Bochum,
Germany*

[¶]*Laboratoire de Chimie Théorique, Sorbonne Université, UMR 7616 CNRS, 4 Place
Jussieu, 75005 Paris, France*

[§]*Department of Physical Chemistry II, Ruhr University Bochum, D-44801 Bochum,
Germany*

E-mail: wanlin.chen@univ-evry.fr, mgaigeot@univ-evry.fr

S1. Fresnel Factors Corrections

In order to compare theoretical SFG spectra of the SAM/water interfaces with experimental SFG spectra measured by Sanders *et. al.* in Ref.,¹ the Fresnel factors were calculated and corrected for the experimental SFG spectra presented in Fig. 2A in the main text. We adopted the same approach for the calculation of Fresnel coefficients as in ref.² with the same system of investigation (OTS/water interface with silica substrate) based on the work of Shen and Heinz.^{3,4} The Fresnel factors were calculated by:

$$L_{xx}(\Omega) = \frac{2n_1(\Omega) \cos \gamma}{n_1(\Omega) \cos \gamma + n_2(\Omega) \cos \beta} \quad (1)$$

$$L_{YY}(\Omega) = \frac{2n_1(\Omega) \cos \beta}{n_1(\Omega) \cos \beta + n_2(\Omega) \cos \gamma} \quad (2)$$

$$L_{zz}(\Omega) = \frac{2n_2(\Omega) \cos \beta}{n_1(\Omega) \cos \gamma + n_2(\Omega) \cos \beta} \left(\frac{n_1(\Omega)}{n'(\Omega)} \right)^2 \quad (3)$$

$$\chi_{eff,SSP}^{(2)} = L_{YY}(w_{SF}) L_{YY}(w_{vis}) L_{ZZ}(w_{IR}) \sin(\beta_{IR}) \chi_{YYZ}^{(2)} \quad (4)$$

The experimental SFG spectra of OTS/water interface before and after Fresnel correction are depicted in Fig. 1. All the experimental SFG spectra presented in Fig. 2A in the main text were corrected.

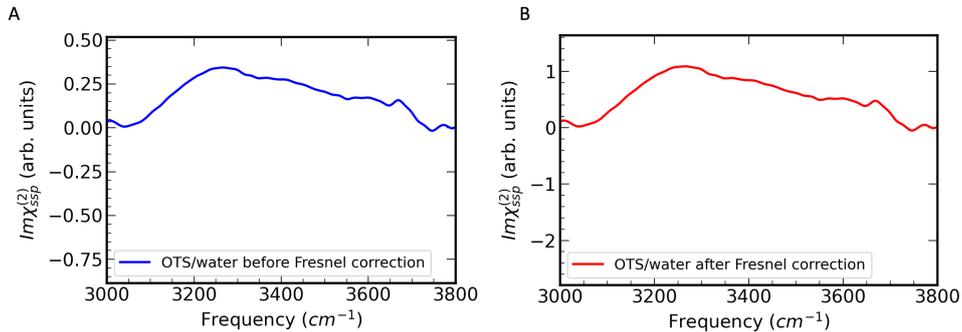


Figure 1: Experimental SFG spectra of OTS/water interface measured by Sanders *et al.*¹(A) before correction of Fresnel factor (scaled) and (B) after correction of Fresnel factor.

S2. SFG Signal of SiOH

Apart from water, SiOH groups on silica also contribute to SFG spectra in the O-H stretch range. The theoretical SFG signature of SiO-H of an amorphous silica/water interface presents a negative peak at low frequency (3000–3200 cm^{-1}), corresponding to surface SiO-H groups H-bonded to water, as reported in ref. 5,6. However, in our systems, silica is not directly in contact with water since it is covered by SAMs, and no water is observed to penetrate the monolayer and reach silica in the simulation. The spectral contributions of SiOH groups are therefore not taken into account in the theoretical SFG spectra of SAM/water interfaces presented in the main text. We present in Fig. 2 the calculated SFG signal of SiO-H stretch at the amorphous silica/water interface, in refs. 5,6.

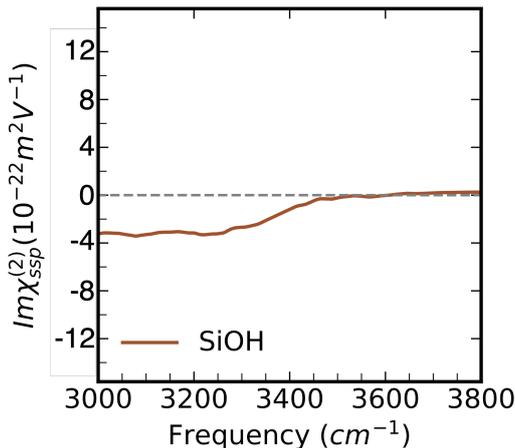
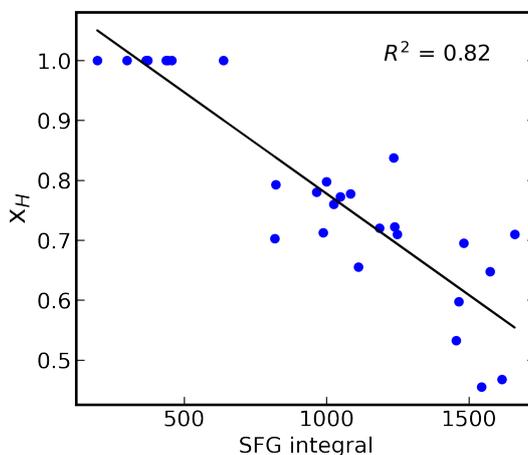


Figure 2: SFG signal arising from SiOH groups on our model silica surface put in contact with liquid water, as previously calculated from DFT-MD in refs. 5,6.

However, although the water penetration is not observed in the simulations, where the monolayers are densely packed, it is still possible to have water penetrating the monolayer and H-bonds to SiOH in experiments, due to the possible less dense packing and disorder of monolayers generated during the deposition procedures. As shown in Fig. 2A of the main text, our DFT-MD calculated SFG spectra of SAM/water interfaces are in good agreement with the experimental ones¹ at above 3200 cm^{-1} . The main discrepancy between the the-

oretical and experimental SFG spectra lies in the low-frequency region (3000–3200 cm^{-1}), where theoretical SFG spectra present positive tails that are not observed in the experimental ones. In this case, the negative peak in the low-frequency range of SiOH groups is the possible contribution to the experimental SFG spectrum, compensating for the positive tail in the water OH SFG spectrum.

S3. Linear Correlation Plot of x_H and SFG Descriptors



S4. H/V Descriptor Captures Local Heterogeneity in Surface Hydrophobicity: Spatially-Resolved x_H Map

While the x_H presented in this manuscript is statistically averaged over the whole simulation box, it does include the effect of local environment. This can be better seen by a spatially-resolved x_H map. We show such a map on a classical MD simulation of the PEG/water interface. For this classical MD, the choice of force field for different parts of the system is based on the setup in ref. 7: CLAYFF force field⁸ for the silica substrate, AMBER force field⁹ coupled with the RESP method¹⁰ for SAMs and SPC-Martini¹¹ for water. An equilibration run of 10 ns in the NVT ensemble has been carried out, followed by an NVT production run of 100 ns, both using a Nose-Hoover thermostat and with a time-step of 0.1 fs. The coordinates and velocities are saved at every 200 fs. Box-size of $40 \times 40 \times 300$ Å has been chosen in our classical MD simulations, and periodic boundary conditions have been applied in all three spatial directions.

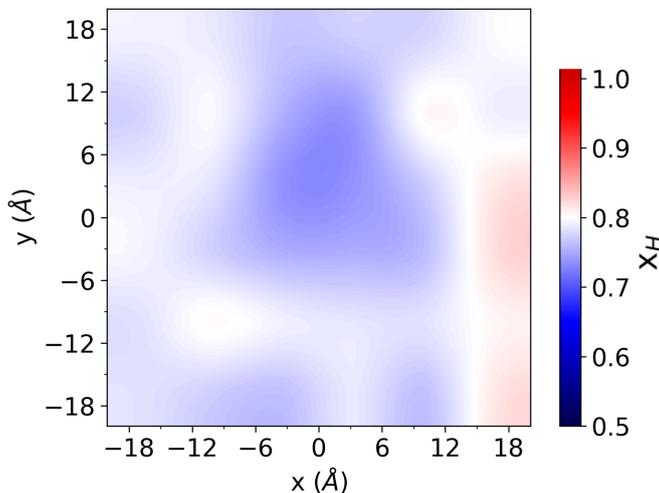


Figure 4: Spatially-resolved x_H map of a PEG/water interface.

The spatially-resolved x_H map of PEG/water interface (Fig. 4) is obtained by dividing the simulation box into grids of 6.6×6.6 Å, over which the density of horizontal H-bonds in the BIL and that of water-surface H-bonds are calculated. The spatially-resolved x_H map of

the PEG/water interface nicely shows that our structural descriptor captures the small local variations in hydrophilicity (from -0.70 to -0.82), with the recognition of a local hydrophobic patch (in light red).

S5. Structural and Spectral Results of DFT-MD Simulations Following Equilibration by Classical MD

The results presented in the main text are obtained from DFT-MD simulations without pre-equilibration from classical MDs. However, for the PEG/water interface, we have also carried out an additional 20 ns classical MD simulations in NVT (same setup as the one introduced in the previous section) followed by a 40 ps DFT-MD simulation (same setup as the ones in the main text).

The density of horizontal and vertical HBs deduced from this 40 ps simulation are $5.4/\text{nm}^2$ and $3.7/\text{nm}^2$, respectively, resulting in a x_H of 0.59, similar to the value (0.63) calculated from 400 ps of DFT-MD simulations without the pre-equilibration of classical MD simulations. Moreover, the total BIL-SFG spectra as well as the up and down deconvolved spectra (see the main text for details) calculated from this simulation are in nice agreement with that from the 400 ps DFT-MD simulations, as shown in Fig. 5.

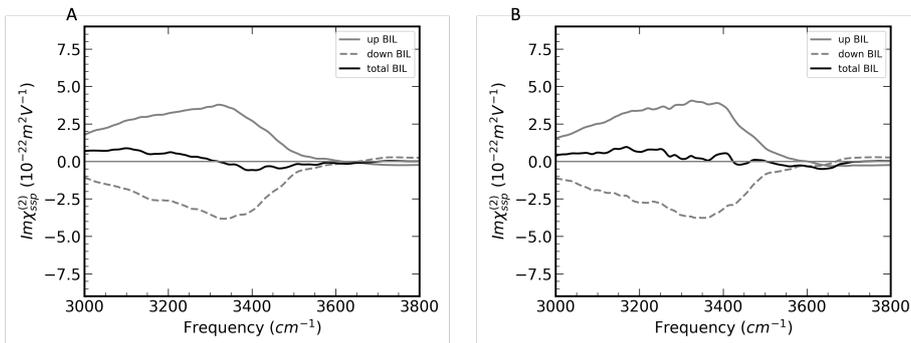


Figure 5: Deconvolved SFG spectra of PEG/water interface calculated from (A) 400 ps of DFT-MD simulations without the pre-equilibration of classical MD simulations, presented in the main text, and (B) from a 40 ps DFT-MD simulation following a 20 ns classical MD simulation.

S6. Number of BIL Water Molecules per Alkyl Chain in Each SAM/Water System

To count the numbers of BIL water molecules per alkyl chain in each of the SAM/water systems, we used cut-off distance values based on the radial distribution functions of SAM atoms (C and O) with O of water extracted from the DFT-MD trajectories. We find cut-off distances of water around C atoms to be 4.5 Å (for both OTSs and PEGs) and that around O of PEGs to be 3.5 Å. By counting the number of water molecules surrounding SAMs chain by chain, we find that at the OTS/water interface, there are on average 4 ± 2 water molecules in the BIL around each OTS chain. At mixed OTS-PEG/water interface, there are 5 ± 3 water per alkyl chain and at PEG/water interface there are 6 ± 3 water/PEG chain.

References

- (1) Sanders, S. E.; Petersen, P. B. Heterodyne-detected sum frequency generation of water at surfaces with varying hydrophobicity. *J. Chem. Phys.* **2019**, *150*, 204708.
- (2) Liljeblad, J. F.; Tyrode, E. Vibrational sum frequency spectroscopy studies at solid/liquid interfaces: Influence of the experimental geometry in the spectral shape and enhancement. *J. Phys. Chem. C.* **2012**, *116*, 22893–22903.
- (3) Shen, Y. Optical second harmonic generation at interfaces. *Annu. Rev. Phys. Chem.* **1989**, *40*, 327–350.
- (4) Heinz, T. F. *Modern Problems in Condensed Matter Sciences*; Elsevier, 1991; Vol. 29; pp 353–416.
- (5) Cyran, J. D.; Donovan, M. A.; Vollmer, D.; Brigiano, F. S.; Pezzotti, S.; Galimberti, D. R.; Gaigeot, M.-P.; Bonn, M.; Backus, E. H. Molecular hydrophobicity at a macroscopically hydrophilic surface. *Proc. Natl. Acad. Sci.* **2019**, *116*, 1520–1525.

- (6) Pezzotti, S.; Galimberti, D. R.; Gaigeot, M.-P. Deconvolution of BIL-SFG and DL-SFG spectroscopic signals reveals order/disorder of water at the elusive aqueous silica interface. *Phys. Chem. Chem. Phys.* **2019**, *21*, 22188–22202.
- (7) Roscioni, O. M.; Muccioli, L.; Mityashin, A.; Cornil, J.; Zannoni, C. Structural characterization of alkylsilane and fluoroalkylsilane self-assembled monolayers on SiO₂ by molecular dynamics simulations. *J. Phys. Chem. C* **2016**, *120*, 14652–14662.
- (8) Cygan, R. T.; Liang, J.-J.; Kalinichev, A. G. Molecular models of hydroxide, oxyhydroxide, and clay phases and the development of a general force field. *J. Phys. Chem. B* **2004**, *108*, 1255–1266.
- (9) 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.
- (10) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
- (11) Marti, J.; Padró, J.; Guardia, E. Molecular dynamics calculation of the infrared spectra in liquid H₂O-D₂O mixtures. *J. Mol. Liq.* **1994**, *62*, 17–31.

3.2 Interfacial organization of both SAMs and water

In section 3.1.2, we introduced a new metric for measuring the microscopic hydrophilicity/hydrophobicity of an interface combining the structural H/V descriptor and spectroscopic quantities. However, the actual structural organization of water and of the polymer part, i.e. the densely packed but flexible SAM polymer chains, remains “mysterious”. To uncover the structures and the dynamics of the flexible surfaces interfaced with water, a series of analyses have been realized and are reported in this section.

3.2.1 Statistically averaged density maps of interfacial atoms of both the surface (SAMs) and water

The amorphous silica substrate, where the SAM chains are anchored, has eight silanol sites at the surface in the model used for the simulations, corresponding to a silanol density of 4.5 SiOH/nm² [163, 164]. To anchor the SAMs, each of the eight silanol sites is replaced by a silane-based SAM chain. Hence, the SAM monolayer of each SAM/water system is composed of eight densely packed SAM chains in our simulation box, i.e. 4.5 SAM chains/nm².

To understand the behavior of the SAM chains in contact with water, the most intuitive way is to visualize how these chains are distributed in space in the simulation boxes during the trajectories. Are they homogeneously distributed and well ordered thanks to the dense packing, or are they disordered due to their interactions with water?

To answer this question, we calculated density maps of the projection onto the xy-plane (parallel to the solid surface) of the coordinates of the atoms belonging to the SAM chains (both carbon atoms and oxygen atoms) specifically located at the interface (denoted as interfacial SAM atoms), which are defined as the SAM atoms located within 3 Å below the Willard and Chandler (WC) interface [109]. The density maps of oxygen atoms of the interfacial water molecules in the BIL (within 3 Å to the WC interface) are also plotted to check the correlation between locations of the interfacial SAM atoms and that of the water molecules in the BIL. Note that the physical meaning of the density is the probability of finding an interfacial atom at a given position in the (x,y) plane (at a given vertical z position). Positions of high-density spots (from greenish to reddish patches, in our scales in the figures) thus correspond to the most probable locations of the interfacial atoms during the trajectory. The maps are indeed calculated as an average over the whole trajectory (for the time being).

Fig. 3.10 displays the density maps of the interfacial atoms at the OTS/water and PEG/water interfaces; both averaged over the ~100 ps of DFT-MD. In the averaged density map of carbon atoms of the OTS polymers (Fig. 3.10B), there are eight localized spots of high density (from greenish to

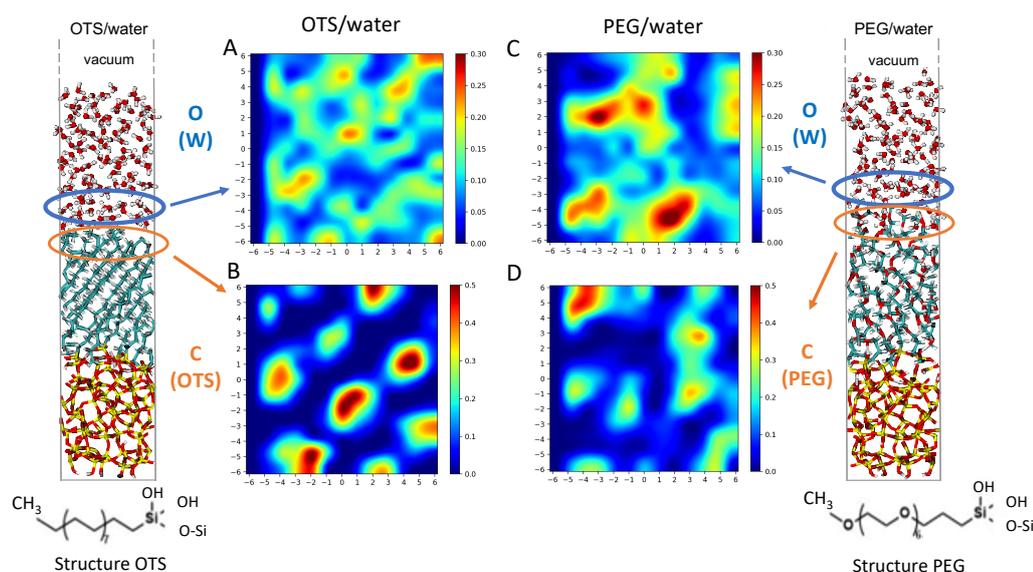


Figure 3.10: Density maps in the (x,y) plane averaged over ~ 100 ps of DFT-MD for the interfacial atoms located within 3 \AA of the Willard and Chandler interface: (A) oxygen atoms of water at the OTS/water interface, (B) carbon atoms of the OTS at the OTS/water interface, (C) oxygen atoms of water at the PEG/water interface, (D) carbon atoms of PEG at the PEG/water interface. The snapshots of OTS/water and PEG/water interfaces from the DFT-MD simulations are presented to illustrate the atoms plotted in each density map.

reddish), corresponding to the most probable positions of the interfacial carbon atoms, distributed into well defined spots in the xy -plane. Note that one has to take into account the replica through periodic boundary conditions. The map indicates that the eight OTS polymer chains are orderly packed near the interface and pointing straight toward the water slab. As regards the carbon map of the PEG/water interface (see Fig. 3.10D), the positions of the high-density spots are more randomly distributed in space, showing that the organization of PEG polymer chains are more disordered at the interface compared to the OTS/water interface, and also are more fluctuating than at the OTS surface (broad spots are observed). The reason for the disordered organization of PEG chains at the interface will be explained later in this section.

Let us now have a look at the averaged density maps of oxygen atoms of interfacial water molecules at the two interfaces and see how the ordered/disordered organization of the SAM polymer chains influence the water organization (and vice versa). At both OTS/water and PEG/water interfaces, the maps of oxygen atoms of the interfacial water molecules in the BIL (Fig. 3.10A and Fig. 3.10C) are anti-correlated with the carbon maps of the SAMs. Indeed, the water molecules seem to occupy interstitial locations in the (x,y) plane, i.e. ‘between’ carbon atoms, showing that water molecules escape from the carbon atoms of the SAM chains. The water map of the OTS/water interface represents rather homogeneously distributed water molecules interconnected with each other, as seen by the highly continuous green colors. This is in line with the discovery of the 2-dimensional H-bond

network (2DN) formed by water in the BIL indicated by the H/V value of the hydrophobicity ($H/V = 1$), see section 3.1.2. The map of water oxygen atoms above the PEGs shows a bit less interconnected patches. This is in line with the more vertical ordered water organization at this hydrophilic interface, correlated with the H/V value of 0.63 (see section 3.1.2).

The behavior of SAMs in contact with water is hence revealed by the density map of carbon atoms located at the interface. While OTS chains are orderly organized at the interface with water (Fig. 3.10B), PEG chains at the interface are more randomly distributed as shown by the more heterogeneously distributed high-density spots of both carbon and oxygen atoms in Fig. 3.11.

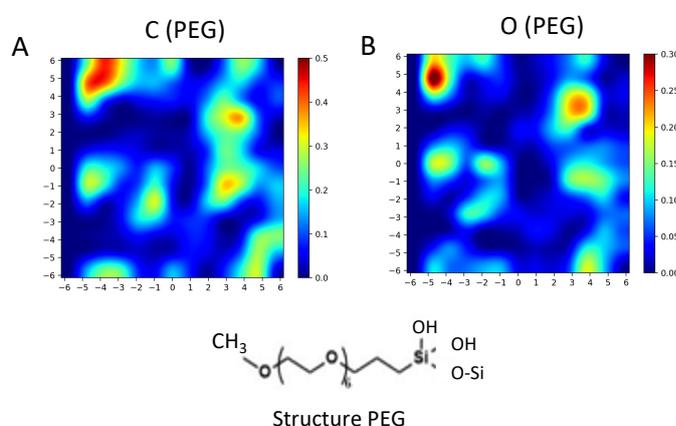


Figure 3.11: Density maps in the (x,y) plane averaged over ~ 100 ps of DFT-MD of the PEG/water interface for the interfacial atoms located within 3 \AA of the Willard and Chandler interface: (A) carbon atoms of PEG at the PEG/water interface. (B) oxygen atoms of PEG at the PEG/water interface.

Morphologically, the heterogeneous distribution of PEG atoms at the interface with water can be due to either aggregation of different PEG chains or bending of certain PEG chains at the interface, which consequently exposes more carbon and oxygen atoms along the bending chains. On the other hand, it might also be due to the dynamical behavior of the PEG chains since the map is calculated as an average over the whole trajectory of 100 ps.

Based on these, we came up with three hypothetical scenarios to explain the non-homogeneous distribution of interfacial PEG atoms: (1) PEG chains are bending at the interface or aggregating; (2) PEG chains are not bending or aggregating but being highly dynamical during the trajectory and therefore they expose various atoms at the surface with water; and (3) PEG chains are dynamically bending or aggregating during the trajectory.

To investigate which of the three is the real scenario that is responsible for the non-homogeneous distribution of interfacial PEG atoms, supplementary analyses have been realized and reported in the following sections. We first develop time-resolved density maps to investigate the dynamics of the SAM chains. We then quantify the number of SAM atoms exposed to water, chain

by chain, to uncover the morphology of the SAMs at the interface with water.

3.2.2 Time-resolved density maps of interfacial atoms at OTS/water and PEG/water interfaces

In this section, we investigate the dynamics of the OTS and PEG chains interfaced with water to answer to the questions raised in the previous section. To that end, we cut each trajectory of 100 ps into pieces of 10 ps and calculate density maps for each 10 ps piece. The time-resolved density maps of interfacial atoms are hence calculated and the representative ones for an OTS/water interface are depicted in Fig. 3.12 (carbon maps) and those of the PEG/water interface in Fig. 3.13 (carbon maps) and in Fig. 3.14 (oxygen maps).

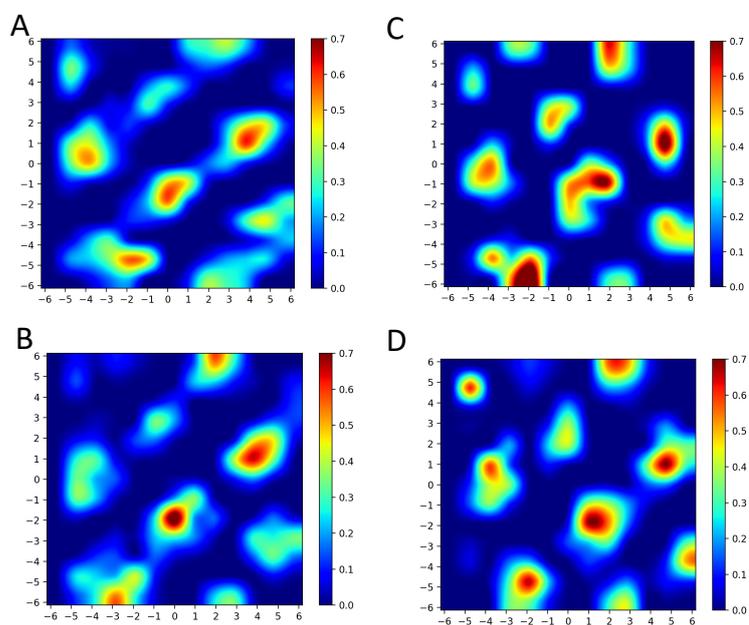


Figure 3.12: Chosen time-resolved density maps of carbon atoms of the OTS chains located within 3 Å to the WC interface of the OTS/water interface averaged (A) from 10 to 20 ps, (B) from 30 to 40 ps, (C) from 50 to 60 ps, (D) from 70 to 80 ps.

From 0 to 80 ps, the time-resolved density maps of carbon atoms at the OTS/water interface (Fig. 3.12) are barely changed and all resemble the averaged one in Fig. 3.10B. The 8 clear density spots equally distributed in the xy-plane show the continuous ordered packing of the OTS chains over the 80 ps trajectory. Moreover, the fact that they are located at almost the same positions during the simulation demonstrates the non-dynamical behavior of the OTS monolayer while in contact with water, i.e. OTS chains are rather static over time.

However, the PEG chains behave differently according to their carbon density maps shown in Fig. 3.13 and their oxygen density maps shown in Fig. 3.14. From one map to another, we indeed observe the changes

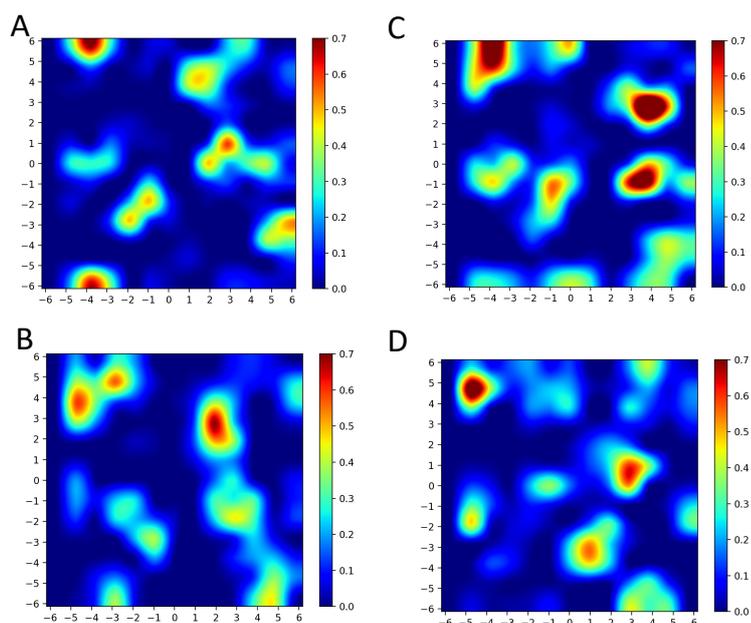


Figure 3.13: Chosen time-resolved density maps of carbon atoms of the PEG chains located within 3 Å to the WC interface of the PEG/water interface averaged (A) from 10 to 20 ps, (B) from 30 to 40 ps, (C) from 50 to 60 ps, (D) from 70 to 80 ps.

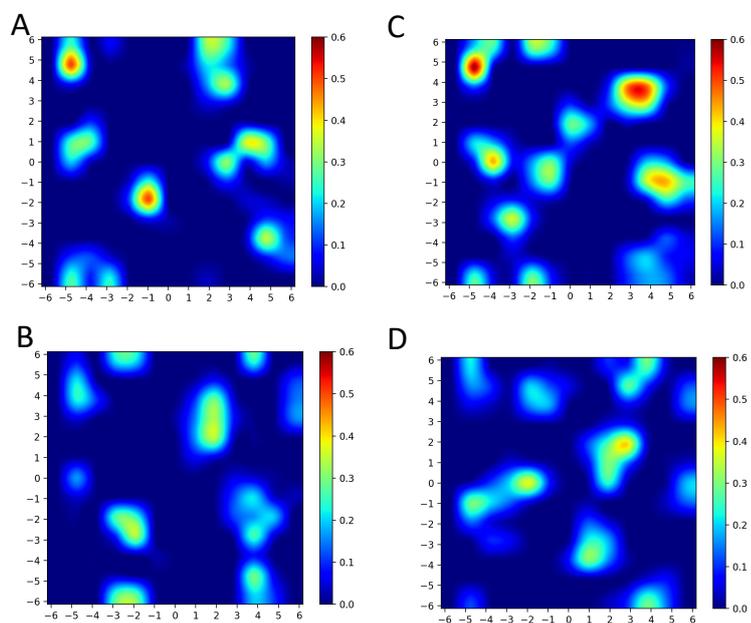


Figure 3.14: Chosen time-resolved density maps of oxygen atoms of the PEG chains located within 3 Å to the WC interface of the PEG/water interface averaged (A) from 10 to 20 ps, (B) from 30 to 40 ps, (C) from 50 to 60 ps, (D) from 70 to 80 ps.

in location of the density spots in both carbon and oxygen maps of the PEG chains at the PEG/water interface. Moreover, the disappearance and reappearance of some of the density spots from time to time illustrate that the PEG chains are highly dynamical so that a certain PEG chain can be

buried (i.e. away from the WC interface) at time t and re-exposed to the interface at another time.

Here comes a question: why are the PEG chains dynamical in contact with water while the OTS chains are not? To answer this, we need to review the main difference between the OTS and PEG chains: the existence of oxygen atoms in the PEG chains. While facing water, the goal of the oxygen atoms of the PEG chains is certainly to form H-bonds with water, which consequently makes the PEG/water interface hydrophilic, as we have shown in section 3.1.2. However, our PEG chains (same as the ones used in ref. [139] for the experimental SFG measurements), should not directly expose oxygen atoms to water. According to its chemical formula $\text{CH}_3(-\text{O}-\text{CH}_2-\text{CH}_2)_7-\text{CH}_2-\text{SiCl}_3$, the group on top of the chain within contact with water should actually be CH_3 and the oxygen atom is only the subsequent part. Therefore, one reasonable hypothesis for the dynamics of PEG chains is that they keep moving in order to maximize the number of oxygen atoms exposed to and H-bonded to water. In another word, PEG chains are in pursuit of water, without having a stable and “static” organization at the interface with water.

3.2.3 Time-resolved PEG-water HBs density and H/V ratio: Oxygens of PEGs in pursuit of water

To check the dynamics of PEG chains at the interface, the most direct way is to keep track of the time evolution of the number of H-bonds formed between the PEG chains and water. Since the number of PEG-water H-bonds is incorporated in our hydrophilicity descriptor, the H/V descriptor (see section 3.1.2 for detailed definition), the time-resolved H/V values are now plotted.

In Fig. 3.15, the evolution as a function of time of the H/V value for one PEG/water DFT-MD trajectory of 100 ps (the same trajectory as the one based on which the density maps presented in Fig. 3.13 and Fig. 3.14 are calculated) is on the upper panel of Fig. 3.15. We further deconvolve the Horizontal and Vertical contributions of the H/V descriptor, i.e. the density of horizontal (H) HBs (water-water) and vertical (V) HBs between water and the PEG surface, and plot them in the middle panel. The dynamics of the H (in red) and V (in blue) contributions are observed and captured in the evolution of the H/V.

The H/V descriptor fluctuates between 0.64 (at 10 ps) and 0.38 (at 60 ps, minimum value), and increases back to ~ 0.6 at 100 ps. This illustrates nicely the dynamics of the H/V descriptor and how it captures subtle fluctuations over time.

The values of the density of PEG-water (vertical) H-bonds (blue bars in middle panel in Fig. 3.15) oscillate between 3 HBs/nm² and 6 HBs/nm². This value can go beyond the density of PEG chains (4.5 nm²). Such high values will be commented and analyzed further in section 3.2.4.

We combine the density maps of the oxygen atoms exposed to the interface

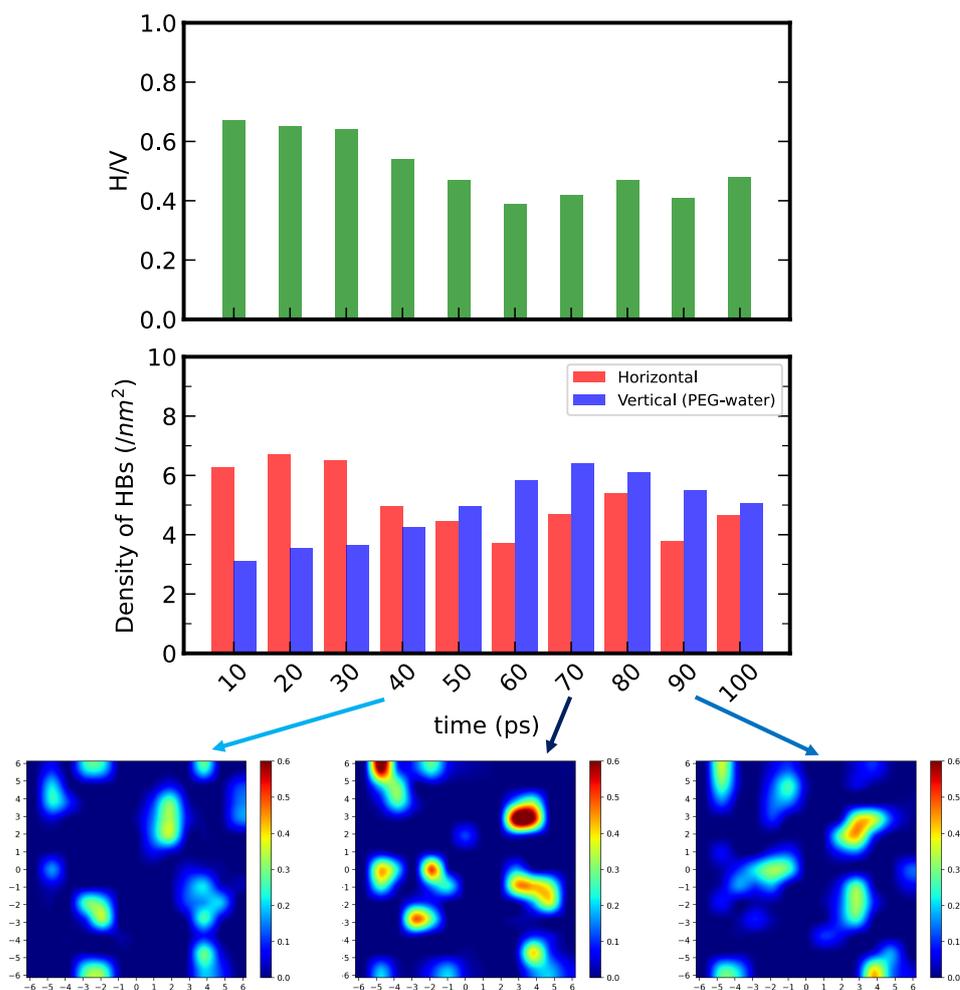


Figure 3.15: Time evolution of H/V descriptor (upper panel) and density of H-bonds being formed horizontally or vertically (middle panel) within one trajectory of the PEG/water interface. The corresponding density maps of oxygen atoms of PEGs at representative moments are plotted at the bottom.

with the evolution of the PEG-water H-bonds density in Fig. 3.15. Three representative density maps of oxygen atoms of PEGs are depicted at times 40 ps, 70 ps and 90 ps. At time 30–40 ps, the density of PEG-water H-bonds is ~ 4 HBs/nm². The corresponding interfacial oxygen density map at time 30–40 ps shows 6 greenish spots. While the number of H-bonds increasing from 40 ps to 70 ps, from 4 HBs/nm² to 6 HBs/nm², the density of the number of oxygen atoms of PEGs exposed to the interface augments (7 reddish spots with the same color scaling). From 70 ps to 100 ps, the number of PEG-water vertical H-bonds decreases, corresponding to the change from 7 reddish spots to 6 greenish spots in the oxygen density map.

The dynamics of the PEG chains engenders more or less exposed PEG atoms at the interface with water at different instants, which consequently causes the fluctuation of the surface-water H-bond number. This correlation nicely demonstrates that oxygen atoms of the PEG chains are willing to be exposed to the water molecules in order to form H-bonds with water, and this

is the driving force for the dynamics of the PEG chains.

3.2.4 Number of hydrated PEG atoms counted chain by chain

The dynamics of the PEG/water interface, leading to the fluctuation of the degree of hydrophilicity at the interface, has just been shown by time-resolved density maps and validated by the correlation with the fluctuation of the number of surface-water H-bonds and of the resulting H/V value.

The density maps of interfacial PEG atoms averaged over a trajectory of ~ 100 ps displayed in section 3.2.1 revealed the heterogeneous distribution of PEG chains, that is possibly due to dynamics or bending of PEG chains at the surface. The further time-resolved density maps shown in section 3.2.2 proved the dynamics of the PEG/water interfaces, partially explaining the heterogeneous distribution of PEG chains at the interface.

The next step is to uncover the morphology of this interface in order to fully answer to the question regarding the cause of the inhomogeneously distributed interfacial PEG atoms raised in section 3.2.1.

As indicated by the chemical formula of PEG ($\text{CH}_3(-\text{O}-\text{CH}_2-\text{CH}_2)_7-\text{CH}_2-\text{SiCl}_3$), the end group (CH_3 group), which is in direct contact with water, is followed by seven repeat units of $(-\text{O}-\text{CH}_2-\text{CH}_2)$. Supposedly, for a densely packed SAM monolayer (like in the present case), in which the chains would be oriented vertically, only the terminal oxygen atom of the PEG chain should be close enough to the interface and should be exposed and H-bonded to water. The second oxygen atom of the PEG chain is located two CH_2 groups away from the first one. However, the fact that the density of surface-water H-bonds plotted in section 3.2.3 can exceed the 4.5 HBs/nm² (in case where all first oxygen atom of 8 PEG chains are exposed to and H-bonded to water) during the simulation indicates that on average more than one H-bond could be made per PEG chain, and indirectly shows that on average more than one oxygen atom can be exposed to water. However, no direct proof has been provided, yet.

To unveil the morphology of PEG chains at the interface, e.g. how many PEG atoms are exposed to water, and are the PEG chains bending at the surface, we quantify the number of carbon and oxygen atoms that are being hydrated and that belong to each PEG chain (in total eight chains in each simulation box). The numbers (chain by chain) are averaged over a trajectory of ~ 100 ps in Fig. 3.16.

The criterion to define if a PEG atom (C or O) is hydrated is based on the radius of the first solvation shell of C atom (4.5 Å) and O atom (3.5 Å), respectively. These values are determined by plotting their radial distribution functions $g(r)$ over the trajectory, and keeping the r value at the minimum intensity of the first peak of the $g(r)$ as the cut-off criterion. Therefore, a carbon atom of PEG is defined to be hydrated when the closest water

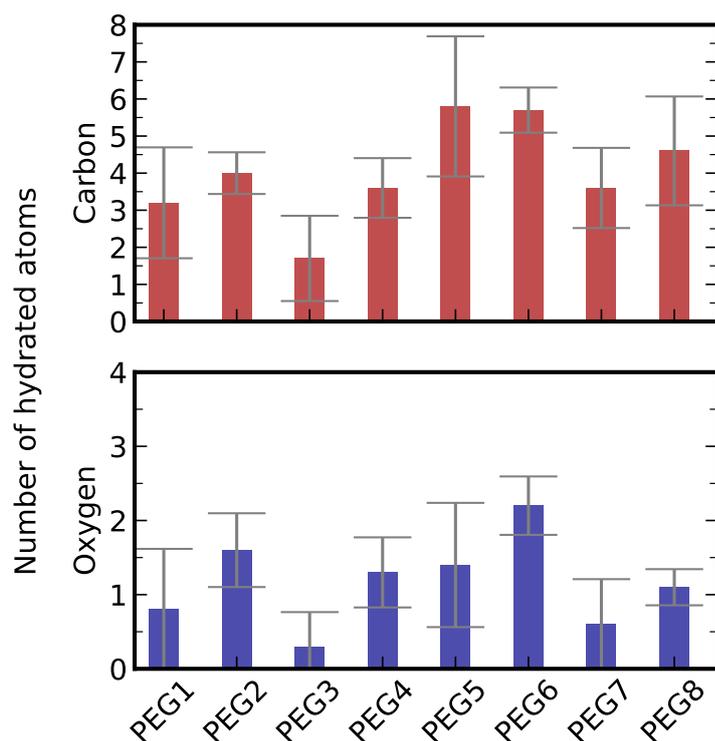


Figure 3.16: Number of hydrated atoms of PEGs chain by chain (in total eight PEG chains in the DFT-MD simulations): Carbon atoms on the upper panel and Oxygen atoms on the lower panel. Carbon atoms are defined as hydrated when the distance between the closest water molecule (O atom) and the Carbon atom is less than 4.5\AA , while Oxygen is considered hydrated with the closest water molecule located within 3.5\AA .

molecule (oxygen atom) is located less than 4.5\AA to it, whereas an oxygen atom of PEG is considered hydrated when the distance between the closest water molecule and this oxygen is less than 3.5\AA .

According to Fig. 3.16, the numbers of hydrated atoms of the 8 PEG chains (both C and O atoms) are highly diversified from chain to chain. The numbers of hydrated Carbon and Oxygen atoms of one chain is linearly correlated. A more “exposed” chain with more hydrated Carbon atoms (e.g. PEG6) simultaneously has more hydrated Oxygen atoms.

As discussed above, for a well ordered and densely packed monolayer, one oxygen per PEG (the terminal top O atom of each chain) is expected to be exposed to water. However, this is not what the simulations show. In the present case, some of the PEG chains have on average more than one hydrated oxygen atom, while some of them have less than one hydrated oxygen atom, revealing the high heterogeneity of the interface.

The heterogeneity and the absolute numbers of hydrated atoms depict the following scenario: while some PEG chains are highly exposed to water by bending themselves to maximize the occupied space at the interface, or more like to maximize their number of oxygen atoms H-bonded to water, some others are miserably buried.

A common known cause for the disorder and heterogeneity of SAM mono-

layers is that the polymer chains have different chain lengths [165, 166], which is not our case since all the PEG chains in the present simulations have the same number of repeat unit, thus same length. Another possible reason for disorder is the loose packing, that intrinsically gives more flexibility for the SAM polymer chains to move, which is not our case either as all the presented systems are composed of fully packed SAMs by construction of the simulation box. Hence, it is interesting to see that despite the same chain length and the dense packing of PEG chains in our simulations, we still observe a disordered and heterogeneous organization of the PEG chains at the surface. This has to be related to the intrinsic interactions with the liquid water at the interface.

Apparently, oxygen atoms being willing to be exposed and H-bonded to water is the driving force for both dynamics and morphological bending of PEGs. This driving force is so strong that the PEG chains with equal starting points can land up with unequal endings, which is concurrently apprehensible: when massive population of polymer chains are fighting for limited resources (water), an uneven distribution of resources is inevitable. Note that these data have not been published yet.

Now we need classical MD simulations with both the same box size and larger box sizes to demonstrate that all these data obtained on small DFT-MD boxes and at short time-scales (100' ps) are correct.

3.3 Dynamics of PEG/water interfaces at larger time- and size-scales: from DFT-MD to Classical FF-MD

In section 3.2.2, we revealed the dynamical behavior of the PEG/water interface and the fluctuation of hydrophilicity as measured by our H/V descriptor within the DFT-MD time scale (tens of ps) and length scale ($13 \times 13 \times 85 \text{ \AA}^3$).

This raises the following questions: (1) Does this fluctuation also exist in a larger box? and (2) Is this the only time-scale for hydrophobicity fluctuation or is there a larger time-scale that cannot be captured by short DFT-MD simulations?

To answer these two questions, we have carried out a series of classical MD simulations (denoted FF-MD for Force Field MD in the text) of the PEG/water interface with different box sizes at different time-scales, see section 1.5.5 for simulation details. These simulations were done in collaboration with Prof. Predota and Dr. Ondrej Kroutil of the University of South Bohemia in Czech Republic, all the classical MD simulations presented in this manuscript have been carried out using the GROMACS package.

In a nutshell, 4 FF-MD simulations of the PEG/water interface on nanosecond-time-scale were carried out, as summarized in the following table. Two box-sizes were chosen in the FF-MD simulations: $13 \times 13 \times 85 \text{ \AA}^3$ (same as the DFT-MD system, denoted 1 \times 1 in the following sections) and $40 \times 40 \times 300 \text{ \AA}^3$ (9 times the DFT-MD size, denoted 3 \times 3).

Table 3.1: List of FF-MD simulations for the PEG/water interface

No.	Size (\AA^3)	Length (ns)	Saved each (ps)	Ensemble
1	$13 \times 13 \times 85$	1	1	NVT
2	$13 \times 13 \times 85$	80	50	NVT
3	$40 \times 40 \times 300$	1	1	NVT
4	$40 \times 40 \times 300$	80	50	NVT

3.3.1 Average structural and spectroscopic properties

Fig. 3.17 (upper panel) shows the averaged H/V (x_H) values calculated from the FF-MD simulations of the PEG/water interface at different time- and size-scales. In general, x_H values of FF-MD simulations are all within the hydrophilic range (less than 0.8), slightly higher than that of the DFT-MD simulation ($x_H = 0.63$, see also section 3.1.2).

The densities of both horizontal H-bonds (H part in the H/V) and vertical H-bonds (V part in the H/V) are plotted in Fig. 3.17 (lower panel). The number of H-bonds in the 1×1 boxes (both H and V) is overall underestimated compared to the values in the DFT-MD simulations. The values in the 3×3 boxes are closer to the reference values in the DFT-MD simulations.

x_H values also show slight size-scale dependency in the FF-MD simulations, ranging from 0.75 for the 1×1 simulations (same size as the DFT-MD system), to 0.68 for the 3×3 box simulations. The latter is closer to the DFT-MD value (0.63), indicating that the FF-MD simulations somehow require a larger simulation box to accurately reproduce the interfacial H-bond network of the PEG/water interface.

However, generally speaking, all these FF-MD simulations overall reproduce well the structural properties of the PEG/water interface, regardless of their size- and time-scales.

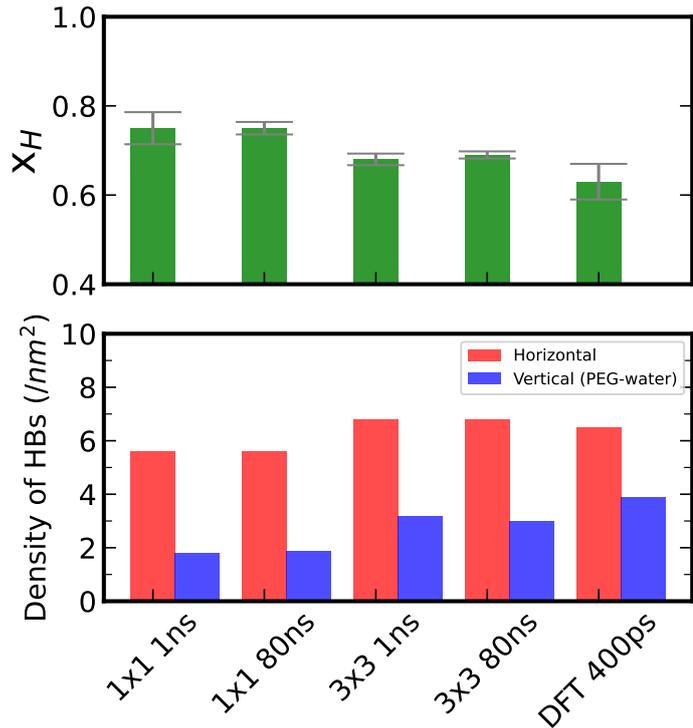


Figure 3.17: Averaged H/V values (x_H) obtained in FF-MD simulations of the PEG/water interface of different time- and length-scales (error bars in grey) compared to x_H of the DFT-MD simulations (400 ps) of the PEG/water interface. Bottom: densities of Horizontal and Vertical H-bonds of the H/V descriptor for all systems.

The SFG spectrum of the BIL-water at the PEG/water interface for the 3×3 box FF-MD simulation of 80 ns is now calculated using our new approach, i.e. the “pop model” that will be detailed in chapter 4. In a nutshell, we do not calculate our theoretical spectrum through the time correlation function between water dipole moments and polarizability tensors [114], nor by calculating the velocity-velocity correlation function (VVCF) using parameterized atomic polar tensors (APT) and Raman tensors as detailed in section 2.3. Instead, thanks to the SFG fingerprints of certain pre-identified interfacial water populations at various aqueous interfaces obtained from DFT-MD calculated spectra, we construct the SFG spectrum by summing up the weight averaged spectrum of each population. The weighting is done by the corresponding abundance of each population observed in the MD simulation (here

the FF-MD). Here, we deduce the abundance from the 3×3 box FF-MD simulation of 80 ns. With this method, we can calculate SFG spectra over a classical FF-MD trajectory at the precision of a DFT-MD trajectory. Furthermore, the spectra calculation is done over an NVT simulation, which is not permitted in the VVCF method.

Using the “pop model” approach, the spectrum of the PEG/water interface based on the 80 ns FF-MD (NVT) simulation is displayed in Fig. 3.18, lower panel. The nice agreement with the DFT-MD calculated spectrum using the VVCF based method (Fig. 3.18, upper panel) demonstrates that our FF-MD simulations accurately reproduce the structural and spectroscopic properties of water at the PEG/water interface. Additionally, this agreement illustrates the success of our “pop model”, which will be examined in detail in chapter 4, with more applications of this approach on various interfaces, ranging from organic to oxides aqueous interfaces, on both water and the surface sides.

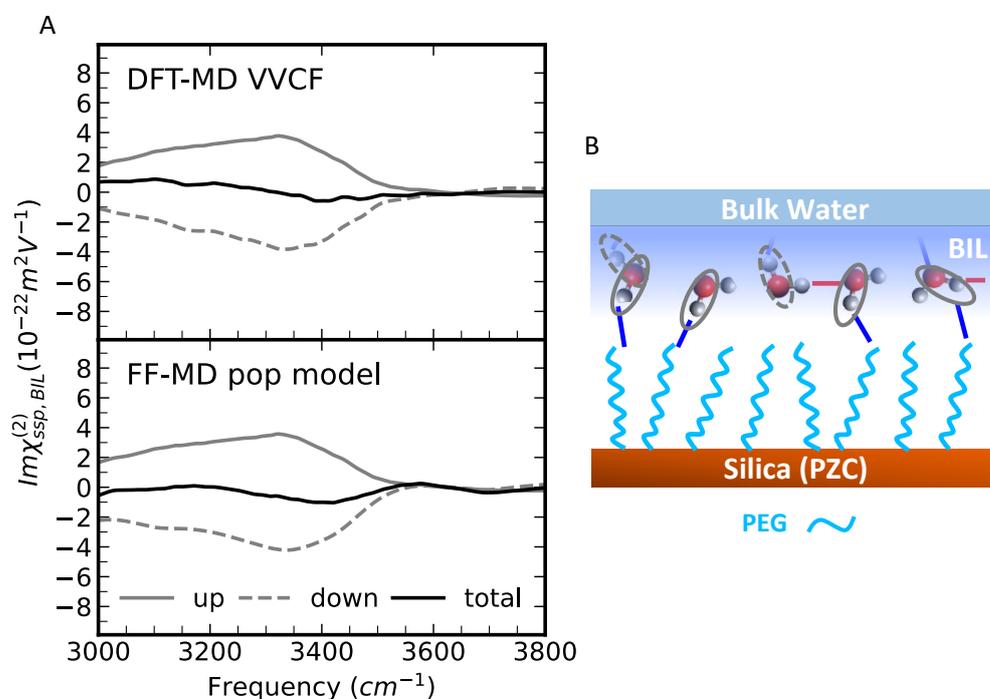


Figure 3.18: (A) BIL-water SFG spectrum of the PEG/water interface. Upper panel: calculation based on DFT-MD simulations of 400 ps using the VVCF method [33, 36] (see section 2.3 for details) Lower panel: based on FF-MD simulations of the 3×3 box of 80 ns using the “pop model” approach (see chapter 4 for details). Up (solid grey) and down (dashed grey) spectra respectively show the deconvolution of the (total) spectrum in terms of water OH groups oriented towards the surface (up) and outwards the surface (down), as sketched in (B) by solid grey and dashed grey circles, respectively.

3.3.2 Dynamics of hydrophilicity at the PEG/water interface

We further investigate the fluctuation in hydrophilicity of the PEG/water interface, and its dependence on size- and time-scales through the time-resolved

$x_H(t)$.

This fluctuation was first captured in our DFT-MD simulations by plotting the time-resolved $x_H(t)$ evolution during 100 ps of trajectory, with a frequency of each 10 ps, as was presented in section 3.2.2. To investigate whether this 10' picoseconds time-scale fluctuation also exists in the FF-MD simulations at different size scales, we plot time-resolved $x_H(t)$ evolution of each 10 ps, for both 1×1 and 3×3 boxes. Fig. 3.19 shows the representative time-resolved $x_H(t)$ plots.

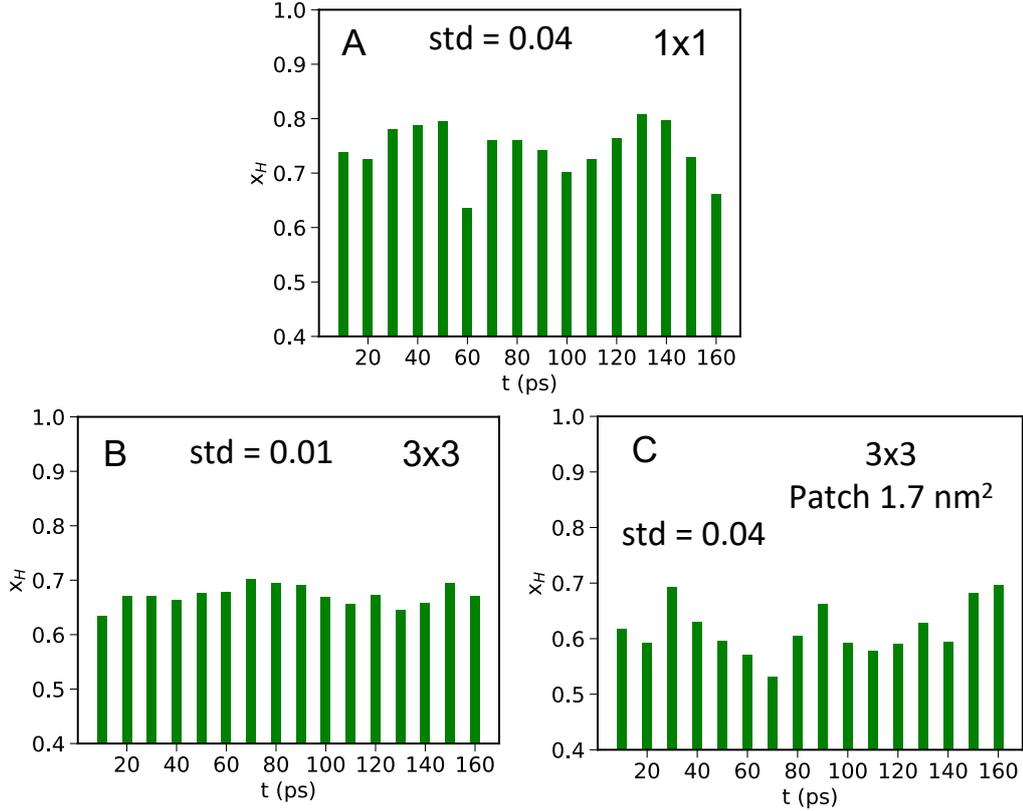


Figure 3.19: Time resolved $x_H(t)$ values for FF-MD simulations of the PEG/water interface of different time- and length-scales: (A) 1×1 box size, each bar corresponds to an average over 10 ps, for a total 160 ps trajectory, (B) 3×3 box size, each bar corresponds to an average over 10 ps, for a total 160 ps trajectory (C) 3×3 box size, each bar corresponds to an average over 10 ps, for a total 160 ps trajectory, averaged over a chosen patch of 1.7 nm^2 at the center of the 3×3 box. The choice of 1.7 nm^2 for the patch size corresponds to the same size as the 1×1 box. The standard deviation (std) values of the $x_H(t)$ fluctuations, that are used to quantify the dynamics of hydrophilicity/hydrophobicity, are marked in all the plots.

The dynamics of hydrophilicity at the PEG/water interface at a time-scale of 10' ps is again captured in the 1×1 box FF-MD simulation, as shown by the fluctuation of x_H values in the course of time (Fig. 3.19A). We use the standard deviation (std) value of the $x_H(t)$ to quantify this dynamics. The std value of the $x_H(t)$ in this simulation is 0.04, same as in the DFT-MD simulations. However, this 10' ps time-scale fluctuation is washed out when we plot the $x_H(t)$ for the 3×3 box system, indicated by a well reduced std

value (0.01) compared to the 1×1 box, see Fig. 3.19B. This shows an overall less pronounced dynamics in hydrophilicity of the PEG/water interface in the 3×3 box than in the 1×1 box. The fact that the 10' ps time-scale dynamics only exists in a small box (1×1) seems to suggest that this dynamics is size-dependent.

To prove this, we plot the time evolution of $x_H(t)$ values that are calculated over a portion of the whole 3×3 box, in a patch of 1.7 nm^2 ($13\times 13\text{ \AA}^2$, same size as the 1×1 box). The plot is shown in Fig. 3.19C. Interestingly, the 10' ps time-scale dynamics observed in the 1×1 box (in both DFT-MD and FF-MD simulations) is recovered in a patch of 1.7 nm^2 in the big 3×3 box. This nicely demonstrates that the dynamics captured in the 1×1 box is actually a local effect with a certain size-scale, that is hence not seen in a big 3×3 box where the dynamics are statistically washed out over the space.

The local dynamics of the hydrophobicity/hydrophilicity with certain size-scales will consequently lead to an interface with spatially heterogeneous hydrophobicity/hydrophilicity. For some physical and chemical processes that take place at the interface, the local properties of the interface are highly relevant. The quantification of the local hydrophobicity/hydrophilicity is thus essential. However, until now, our H/V descriptor has only been used to deduce average hydrophobicity/hydrophilicity over the space. It is time to extend the application of our H/V descriptor to resolve local hydrophobicity/hydrophilicity. A spatially-resolved x_H map of local hydrophobicity/hydrophilicity is hence developed, answering to the call from its destiny.

3.3.3 Spatially-resolved x_H maps

x_H is intrinsically suitable to be developed spatially-resolved, as it probes local quantities by construction. It quantifies vertical and horizontal H-bonds that are local in essence. It is thus easy to develop a x_H map at an angstrom resolution (single-water resolution). This is an advantage that other quantities do not possess, e.g. density fluctuations [8, 30, 111], that cannot reach single-water resolution.

An x_H map is obtained by dividing the simulation box into grids. For PEG/water (3×3 box), we divide grids of $6.6\times 6.6\text{ \AA}^2$, the whole simulation box is thus made of in total 36 patches. Over each patch, the density of horizontal H-bonds in the BIL and that of water-surface H-bonds are calculated. The x_H value of each patch is hence obtained.

Spatially-resolved x_H maps of the PEG/water interface for the FF-MD simulation of the 3×3 box averaged over different time-lengths (1 ns and 80 ns) are plotted respectively in Fig. 3.20A and B.

In the x_H maps of fig. 3.20, blueish patches ($x_H < 0.8$) correspond to local hydrophilic patches, while reddish patches ($x_H > 0.8$) are hydrophobic patches. The x_H map calculated over 1 ns (Fig. 3.20A) nicely captures the heterogeneity in the interface hydrophilicity, with coexistence of well-defined

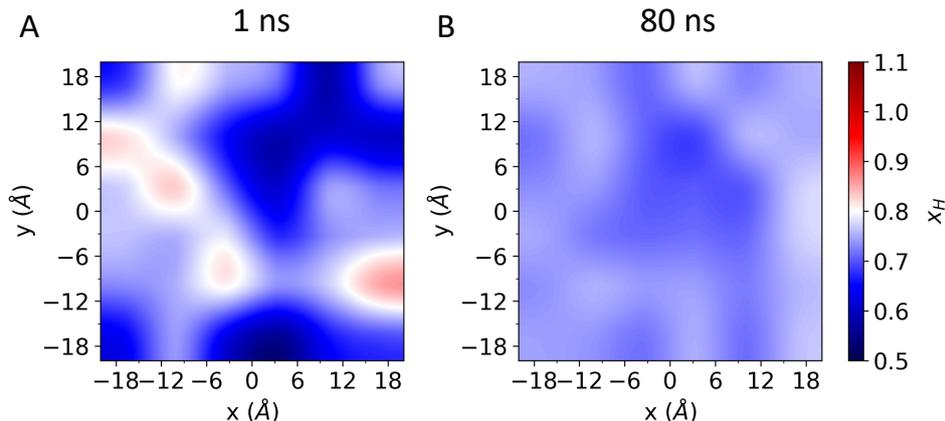


Figure 3.20: Spatially-resolved x_H map of the PEG/water interface in the 3×3 box FF-MD trajectory, calculated over (A) 1 ns trajectory and (B) 80 ns trajectory. In x_H maps, blueish patches ($x_H < 0.8$) correspond to local hydrophilic patches, reddish patches ($x_H > 0.8$) are hydrophobic.

hydrophilic (blueish) and hydrophobic (reddish) patches, despite the fact that the averaged x_H value is hydrophilic (0.68, see section 3.3.2). However, this heterogeneity disappears when the x_H map is calculated over a longer time length. Fig. 3.20B shows the spatially-resolved x_H map averaged over 80 ns, which is now homogeneously hydrophilic (blueish).

Hence, the existence of hydrophobic patches at a hydrophilic interface (on average) is local, not only in space, but also in time. It is thus necessary to quantify the relevant time-scale of the heterogeneity in surface hydrophilicity.

3.3.4 Quantification of the time-scale of the heterogeneity in hydrophilicity/hydrophobicity at the PEG/water interface

With the spatially resolved x_H maps, we have a tool that captures heterogeneity in hydrophilicity/hydrophobicity at the PEG/water interface, as shown in section 3.3.3. Moreover, we discovered that this heterogeneity exists within a certain time-scale, since the x_H map averaged over a relatively long trajectory (80 ns) becomes homogeneously hydrophilic.

The next goal is to deduce the time-scale of the heterogeneity in hydrophilicity/hydrophobicity at the PEG/water interface.

Apparently, the spatial heterogeneity decreases when we increase the time lengths over which the x_H maps are calculated. To obtain the time-scale of this heterogeneity, we calculate x_H maps averaged over increasing time lengths, and evaluate what is the time-scale for the heterogeneity in hydrophilicity to disappear, i.e. for obtaining a homogeneously hydrophilic x_H

map.

We denote the x_H maps as $M_{x_H}(t)$, where t is the time length over which the map is calculated. The heterogeneity in hydrophilicity can be quantified by calculating the spatial variance of these $M_{x_H}(t)$ maps (i.e. the x_H value of each patch in the map) with respect to the average x_H value over the whole box and over the total trajectory of 80 ns (denoted $\bar{x}_H = 0.69$).

The spatial variance as a function of time length (t) over which the map is calculated is denoted $C(t)$:

$$C(t) = \langle [M_{x_H}\{x, y\}(t) - \bar{x}_H]^2 \rangle \quad (3.4)$$

where $M_{x_H}\{x, y\}(t)$ corresponds to the x_H value of the patch at $\{x, y\}$ coordinate in the map calculated over time t , and $\langle \dots \rangle$ denotes averages over all grid points of the map.

To quantify the timescales for relaxing this heterogeneity in hydrophilicity, we measure the decay of $C(t)/C(t_0)$, normalized by the variance value of the chosen initial configuration. This is shown in Fig. 3.21A. The decay plot is

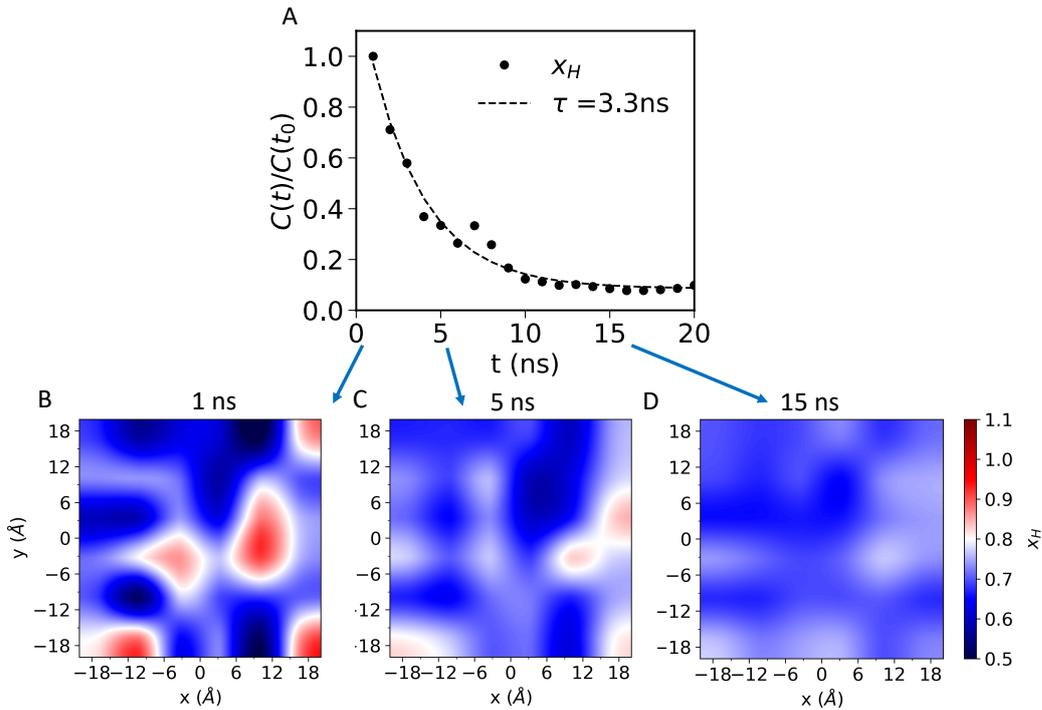


Figure 3.21: Relaxation of the spatial heterogeneity in hydrophilicity at the PEG/water interface. (A) Decay plot of the spatial heterogeneity. The heterogeneity is quantified by the spatial variance ($C(t)/C(t_0)$) of the $M_{x_H}(t)$ maps calculated using eq. 3.4. The fitted exponential function of the decay plot is represented in dashed black, with a hence deduced relaxation time of 3.3 ns. The representative $M_{x_H}(t)$ maps averaged over (B) 1 ns, (C) 5 ns and (D) 15 ns of the trajectory.

fitted by an exponential function:

$$C(t)/C(t_0) = A * \exp\left(-\frac{t}{\tau}\right) + B \quad (3.5)$$

where τ is the relaxation time deduced from the decay plot, i.e. the time-scale of the heterogeneity in hydrophilicity at the PEG/water interface. The fitted exponential function yields a relaxation time of 3.3 ns (obtained for one trajectory). This could be calculated over a statistically relevant number of trajectories for a statistical average.

The representative x_H maps averaged over 1 ns, 5 ns and 15 ns, are shown in Fig. 3.21B, C and D respectively, below the decay plot. These x_H maps illustrate the evolution of surface hydrophilicity from being heterogeneous to homogeneous with increasing average time. This nicely demonstrates that the PEG/water interface, despite being hydrophilic, is locally and dynamically hydrophobic at a nanometer/nanosecond scale.

The next question is, what is the origin of this dynamics of the hydrophilicity at the PEG/water interface? In section 3.2, we qualitatively ascribed the picosecond dynamics of the surface hydrophilicity, that was captured by the evolution of the H/V values, to the dynamics of the PEG chains in pursuit of the oxygens of water in the BIL, that was revealed by the density maps of interfacial PEG atoms. Is this nanoscale dynamics in hydrophilicity also related to the movement of the PEG chains at the interface?

To answer this question, the quantification of the time-scale of PEG dynamics is necessary. To this end, we calculate the same decay plot for the density maps of PEG atoms (both C and O) at the interface, see Fig. 3.22B. The relaxation time for the PEG density maps to be homogeneous is also on the order of ~ 3 ns. This nicely demonstrates that the dynamics of hydrophobicity captured by the x_H maps is induced by the movements of the PEG chains at the interface.

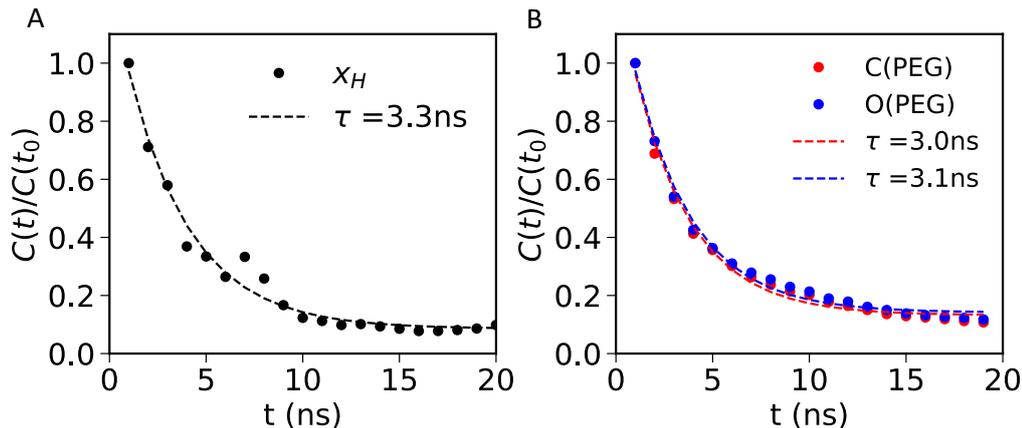


Figure 3.22: Decay plots. (A) time evolution of x_H map variance of water in the BIL and (B) time evolution of variance of density maps of interfacial atoms of PEG (both C and O).

In summary, the dynamics of the PEG chains at the interface triggers the dynamics of the surface hydrophilicity at two different time scales.

The shorter picosecond time-scale is triggered by the topmost part of the PEG chains that expose more O atoms to water when bended, while they mostly expose the topmost C atom when straight. This in turn leads to the intermittent formation of local vertical PEG-water HBs.

The nanosecond time-scale dynamics is related to large-scale (nanometers) fluctuation of the structure of the monolayer, that induces the formation of nanometric patches with increased or reduced hydrophilicity.

This nanoscale dynamics in hydrophilicity is highly relevant for $< \sim$ ns chemical and physical processes that happen at the interface, such as adsorption of ions. For the processes faster than ns time scale, what matters is no longer average properties that can be measured by the experiments, but local and instantaneous properties at a given time.

This discovery, combined with other results presented in this chapter, well uncovered the complexity and subtleties of the interface hydrophobicity/hydrophilicity. As a short summary:

- Molecular hydrophobicity/hydrophilicity is a local property, that can not be fully captured by macroscopic descriptors, such as contact angle.

This is the reason why microscopic descriptors for hydrophobicity/hydrophilicity are needed. To answer to that need, the structural H/V descriptor for hydrophobicity and hydrophilicity, that is for the first time connected to spectroscopic quantities, was developed and presented in section 3.1.2 in this chapter.

- Molecular hydrophobicity/hydrophilicity depends on the topology of the interface [8, 9], that can not be naively deduced by adding up assigned intrinsic hydrophobicity and hydrophilicity values to various chemical functional groups [7].

This is in line with our results in section 3.1.2 and reported in more details in our paper [*On the trail of molecular hydrophilicity and hydrophobicity at aqueous interfaces, J. Phys. Chem. Lett. 2023, 14, 1301-1309*].

We found that the homogeneously mixed OTS-PEG/water interface modeled in our DFT-MD simulation does not have a middle hydrophilic/hydrophobicity in between the OTS/water and PEG/water interface. Instead, its hydrophilic/hydrophobicity is closer to the hydrophilic PEG/water interface, because of the homogeneous distribution of hydrophobic and hydrophilic functional groups that somehow keeps the dynamics of the PEG chains.

- Molecular hydrophobicity/hydrophilicity is not always a static property as we commonly assume. It can manifest dynamically in time with a nanosecond time scale for the PEG/water interface investigated in the work.

As demonstrated in this section, the dynamics of the interface induces the heterogeneity of hydrophobicity/hydrophilicity. The timescale of this heterogeneity is correlated to that of the surface dynamics. This dynamics-triggered heterogeneity will affect the chemical and physical processes at the interface, that are faster than the corresponding timescales.

All the subtleties mentioned above were successfully captured by our H/V descriptor. As a local property (quantifying H-bonds at the interface), H/V is suitable to be developed spatially-resolved at a single-water resolution, that probes the spatial heterogeneity in hydrophobicity/hydrophilicity. By calculating the H/V values evolution as a function of time based on MD simulations, the dynamics of hydrophobicity/hydrophilicity is revealed.

As a perspective, by playing with the distribution of different functional groups in the simulations, the topology effect for hydrophobicity/hydrophilicity will be captured by the H/V descriptor (both time-resolved and spatially-resolved). We expect to see future utilities of this descriptor showing its full potential.

These data and conclusions are currently assembled for a paper.

Chapter 4

Pop model: a simplified approach for theoretical SFG calculation and interpretation

In this chapter, we introduce a new concept for theoretical SFG spectroscopy.

In chapter 2, we introduced the principles of SFG spectroscopy and different approaches for theoretical SFG calculation. In these approaches, the second order susceptibility $\chi^{(2)}(\omega)$ is calculated either based on the correlation function between molecular dipole moments and polarizability tensors [34, 35, 37–39] (section 2.2) or based on the velocity-velocity correlation function (VVCf) [41, 42] (section 2.3). Despite the fact that the VVCf-based method substantially reduces the computational cost [41, 42] of an SFG spectrum, the time-correlation function between velocities still remains computationally costly. This indeed requires a certain length of trajectory for the correlation to converge. See ref. [43] for a discussion showing that.

This approach works with high quality MD simulations, as shown in many previous works in the group, where good to excellent agreements with experiments was obtained for a large variety of aqueous interfaces, ranging from hydrophobic to hydrophilic, from organic to oxide interfaces [27, 32, 33, 43, 110], including the spectra of SAM/water interfaces presented in chapter 3 [164].

However, the need to evaluate time correlation functions limits the applicability to atomistic molecular dynamics simulations. This excludes theoretical spectroscopy from e.g. Monte Carlo simulation, coarse-grain approaches, classical DFT approaches. Moreover, even for classical FF-MD, the evaluation of time correlation functions still requires the use of sufficiently sophisticated flexible force fields FFs, which are difficult to design and more computationally expensive.

Of equal importance, the interpretation of the SFG spectroscopy is not straightforward in most cases. The SFG spectrum of a given aqueous interface comprises spectral contributions from both water and the surface. Different

orientations of OH groups result in positive or negative peaks that compensate in cases where the peaks are located at the same frequencies. For instance, as shown in chapter 3, section 3.1.1, the SFG spectrum of the BIL-water at the PEG/water interface is zero due to the compensation of the positive and negative bands because of the same strength of PEG-water and water-water H-bonds [164]. This can not be anticipated.

Such compensation plus the overlap of SFG signals from different water layers on top of the surface contributions impede us an immediate interpretation of the experimental SFG features in terms of the molecular structures at the interface [33].

Therefore, up to now, full interpretation of a SFG spectrum has been based on the deconvolution of the total spectrum and assignment of the different spectroscopic fingerprints to the corresponding structural populations of both water in the BIL and surface [33, 43, 164]. This procedure is not straightforward, is system dependent, there is no automatized protocol, and it requires chemical and physical insights.

To overcome these limits, in this chapter, we propose a new approach for simultaneously calculating and interpreting SFG spectra, named the “**pop model**”. This approach allows us to construct and interpret a SFG spectrum from a database of spectroscopic fingerprints of water and surface populations as long as their abundances are known. The abundances of the interface populations can be obtained either through simulations or from existing theoretical models.

A large database of theoretical and experimental SFG spectra have been established at various aqueous interfaces in the previous works in the group [27, 32, 33, 43, 110]. Different water and surface populations at the interface have been identified based on their local H-bond coordination and water OH group orientation in the DFT-MD simulations. The fingerprints arising from different populations at aqueous interfaces have been extracted. For instance, see ref. [43] where silica-water motifs have been found with their individual SFG signatures in the 3000 to 4000 cm^{-1} spectral range.

The “pop model” approach is detailed in this chapter in the form of a paper that is still in its working form [*A simplified method for theoretical sum-frequency generation spectroscopy calculation and interpretation: the “pop model”*].

The “pop model” approach circumvents the need to sample vibrational dynamics and to calculate the velocity-velocity correlation functions [33, 36]. This expands the applicability of theoretical spectroscopy to other computational approaches. Moreover, for MD it allows the use of simple FFs (and even coarse grain models) as it needs only the knowledge of the statistics of populations at the interface.

In this chapter, this method has been applied to calculate SFG spectra in both the O-H stretch region (to probe water and surface functional OH groups) and the phonon range (to probe the surface structure) as long as the database of population spectra is established. It thus increases the size- and time-scale samplings of interfaces at the precision of the DFT-MD without the need to design a FF of spectroscopic accuracy. We believe this constitutes an advance in the domain of computational spectroscopy.

Even more importantly, the “pop model” provides a way to calculate theoretical SFG spectra even in cases where DFT-MD or FF-MD simulations are not available, as long as the abundance of interfacial populations can be deduced from existing theoretical models.

A successful application is shown in chapter 5, section 5.3.4. Using the “pop model”, by defining the surface populations and deducing the abundances through theoretical models, we managed to construct pH-dependent SFG spectra without performing DFT-MD simulations under pH conditions.

We now report our paper [A simplified method for theoretical sum-frequency generation spectroscopy calculation and interpretation: the “pop model”] where the development of the pop model and a few applications are presented.

A simplified method for theoretical sum-frequency generation spectroscopy calculation and interpretation: the “pop model”

Wanlin Chen,¹ Simone Pezzotti,² Dorian Louaas,¹ Flavio Siro Brigiano,³ and Marie-Pierre Gaigeot¹

¹*Université Paris-Saclay, Univ Evry, CNRS, LAMBE UMR8587, 91025 Evry-Courcouronnes, France*

²*Department of Physical Chemistry II, Ruhr University Bochum, D-44801 Bochum, Germany*

**Corresponding author: Marie-Pierre Gaigeot, mgaigeot@univ-evry.fr, Wanlin Chen, wanlin.chen@univ-evry.fr*

³*Laboratoire de Chimie Théorique, Sorbonne Université, UMR 7616 CNRS, 4 Place Jussieu, 75005 Paris, France*

INTRODUCTION

Sum Frequency Generation (SFG) spectroscopy, first reported by Shen *et al.* in 1987^{1,2}, is a powerful tool for interface characterization. As a non-linear vibrational spectroscopy technique, SFG depends on the second-order susceptibility $\chi^{(2)}(\omega)$, which is zero in centrosymmetric media (e.g. bulk of liquids and many solids). Thanks to its surface specificity, SFG spectroscopy has been widely utilized to investigate aqueous interfaces to unravel their structural, vibrational and dynamical properties at a molecular level^{3–10}.

Interpreting SFG spectra in terms of microscopic surface structures requires theoretical analyses. Extensive research has been dedicated for theoretical $\chi^{(2)}(\omega)$ calculation via molecular dynamics (MD) simulations, which are based on correlation functions between molecular dipole moment and polarizability tensor^{11–15} or velocity-velocity correlation functions (VVCf)^{16–22}. However, the interpretation of SFG spectra remains challenging due to the overload of information contained in the signals.

In the 3000–4000 cm^{-1} spectral domain, the first problem of the existing methods on Heterodyne-detected (HD)-SFG signal interpretation is the compensation of positive and negative peaks due to different orientations of the OH groups. Both water OHs and surface OH groups (e.g. silanols on silica surface) contribute to the SFG signal, which is hard to be dissected. Moreover, the SFG signals arising from water have spectral contributions from different water layers with distinct structural properties. Full interpretation of SFG spectrum thus requires deconvolution of the total spectrum and assignment of different spectroscopic fingerprints to the corresponding structural populations of both water and surface.

Based on distinct structural properties, an aqueous interface can be separated into three layers^{18,19,23}: the Binding Interfacial Layer (BIL), i.e. the water layer directly in contact with the surface together with the top-surface contribution, the Diffuse Layer (DL), i.e. the subsequent water layer with bulk-like water reoriented by the surface charges²³, and liquid bulk water. Only BIL and DL are SFG active with broken centrosymmetry. By calculating separately the SFG signals arising

from water and top-surface OHs in the BIL and DL, the total SFG spectrum can be deconvolved:¹⁹

$$\chi^{(2)}(\omega) = \chi_{BIL}^{(2)}(\omega) + \chi_{DL}^{(2)}(\omega) \quad (1)$$

where $\chi_{BIL}^{(2)}(\omega)$ and $\chi_{DL}^{(2)}(\omega)$ are spectral contributions from BIL and DL.

The physics behind $\chi_{BIL}^{(2)}(\omega)$ and $\chi_{DL}^{(2)}(\omega)$ are different. $\chi_{DL}^{(2)}(\omega)$ arises solely from a population of bulk-like water molecules reoriented by the electric field generated by the surface charges. $\chi_{DL}^{(2)}(\omega)$ is thus a universal two-band signal centered at 3200 cm^{-1} and 3400 cm^{-1} , which sign (in HD-SFG) and amplitude depend on the surface potential^{19,23}. $\chi_{BIL}^{(2)}(\omega)$ is composed of contributions from the surface functional groups (e.g. OH terminations) and the water molecules (denoted BIL-water) interacting with both the surface and the subsequent water layer (DL). The interactions between BIL-water and surface depend on the types and local distribution of the surface functional groups, the surface morphology, the topology, etc., resulting in a variety of structural populations providing surface-specific fingerprints in the SFG spectrum.

Because of these complexities, just computing theoretical SFG spectra is usually far from sufficient in order to interpret them. This requires different deconvolution steps (to dissect BIL, DL signals, as well as the contributions from distinct BIL populations) and input of structural knowledge extracted from MD simulations. More importantly, this is limited to systems that can be simulated with good enough accuracy for spectral calculation, e.g. at the DFT-MD level in the case of the oxide interfaces (where water-surface interactions are hard to describe with sufficient accuracy for spectroscopy by classical force fields). Making things even worse, many relevant interfacial conditions can not even be realistically simulated, such as variations in pH conditions or electrolyte solutions with low ionic concentrations. It is thus highly challenging to accurately calculate and interpret the SFG spectra of interfaces in the wide range of conditions that they experience in nature and common applications.

In the present work, we hereby propose a simplified method, the “pop model”, that constructs SFG spectra of aqueous interfaces as a sum of spectral contributions from

populations present in the different layers. This method consists in rewriting the standard equations to compute theoretical spectra into a sum of partial contributions from the different interfacial populations, weighted by the abundance of these populations at the interface. With this, we will show here that it is possible to parameterize the SFG signatures from each population in a minimum dataset of reference partial spectra and predict with high accuracy (comparable to DFT-MD) the SFG spectrum of a given interface only from the knowledge of the population statistics at that interface. Such knowledge can be derived with the simplest classical force fields, which substantially enlarge the complexity of interfaces over which theoretical spectra can be deduced from simulations. More importantly, the ‘‘pop model’’ provides a way to calculate theoretical SFG spectra even in cases where simulations are not available, as long as the abundance of interfacial populations can be deduced from existing analytical models. We provide a proof of principle for this exciting perspective by predicting pH-dependent surface SFG spectra of silica/water interfaces as a function of pH in excellent quantitative agreement with experiments, from the knowledge of pKa values of surface SiOHs and Henderson–Hasselbalch equations, only.

METHOD AND APPLICATIONS

In the following, we introduce the equations for the ‘‘pop model’’ in order of increasing complexity. We start from the DL, which can be described by a single water population, and then extend the approach to the BIL, where the signals of multiple water and surface populations overlap.

A. SFG Signal in the Diffuse Layer (DL)

As shown in refs.^{19,23–25}, the SFG $\chi_{DL}^{(2)}(\omega)$ response from water in the DL region of the interface can be rewritten in terms of the bulk liquid water third order susceptibility $\chi_{Bulk}^{(3)}(\omega)$ contribution, hence providing the generic two bands at 3200 and 3400 cm^{-1} of any DL-SFG spectrum, weighted in sign and magnitude by the potential difference $\Delta\phi_{DL}$ across the DL:

$$\chi_{DL}^{(2)}(\omega) = \chi_{Bulk}^{(3)}(\omega) \cdot \Delta\phi_{DL} \quad (2)$$

where $\chi_{Bulk}^{(3)}(\omega)$ is the third order susceptibility of liquid water and $\Delta\phi_{DL}$ is:

$$\Delta\phi_{DL} = \int_{z_a}^{\infty} dz \cdot E_{DC}(z) e^{i\Delta k_z z} \quad (3)$$

with z_a the vertical (i.e. perpendicular to the surface) boundary between the BIL and DL and ∞ is bulk water, $E_{DC}(z)$ the electrostatic field, and Δk_z a phase factor

that takes into account interferences between the emitted light at different depths from the surface.

When theoretical DL-SFG spectra are calculated from MD simulations using the approach we have developed and validated in ref. 19, the knowledge of $\chi_{Bulk}^{(3)}(\omega)$ and $\Delta\phi_{DL}$ is not required anymore. $\chi_{DL}^{(2)}(\omega)$ is simply obtained from the Fourier transform of the dipole-polarizability correlation function (i.e. the standard time-dependent method introduced by Morita *et al.*,^{11,12}) and by considering only the contribution of the DL-water molecules (identified based on a clear definition of the evolution of water density as a function of the vertical distance from the instantaneous surface²⁶, as detailed in ref. 18,19), so that:

$$\chi_{DL}^{(2)}(\omega) = \frac{i\omega}{k_B T} \sum_{i=1}^{N_{DL}} \left(\sum_{j=1}^{N_W} \int_0^{\infty} dt e^{i\omega t} \langle \alpha_{xx}^j(t) \mu_z^i(0) \rangle \right) \quad (4)$$

where all the possible cross-correlation terms between the N_{DL} water molecules in the DL layer and all the N_W water molecules in the system ($N_W = N_{BIL} + N_{DL} + N_{Bulk}$) are explicitly included. The xxz component of the $\chi_{DL}^{(2)}(\omega)$ tensor is here considered, corresponding to *ssp* polarization of the SFG lasers. $\mu_z^i(0)$ and $\alpha_{xx}^j(t)$ respectively refer to the individual dipole and polarizability of the i -th and j -th water molecules, j and i are indices running over the total number of water molecules in the simulation box (N_W) and over the number of water molecules in the DL (N_{DL}) respectively. We note here that in common applications from DFT-MD the SFG signal is calculated with the VVCF approach, which use a different formulation than eq. 4 (furthermore often including self-correlation terms only)¹⁷. Hereafter, the demonstration of the ‘‘pop model’’ is derived from the eq. 4.

The $\frac{i\omega}{k_B T} \sum_{j=1}^{N_W} \int_0^{\infty} dt \exp(i\omega t) \langle \alpha_{xx}^j(t) \mu_z^i(0) \rangle$ term in the equation represents the contribution to the total SFG spectrum of the i -th water molecule in the DL. This is labeled $\beta_{DL}^i(\omega)$ hereafter.

We can thus rewrite Eq. 4 as:

$$\chi_{DL}^{(2)}(\omega) = \sum_{i=1}^{N_{DL}} \beta_{DL}^i(\omega) \quad (5)$$

As demonstrated in refs.^{19,23,24,27}, the SFG activity of the DL is due to a pure second order effect, dominated by bulk-water reorientation induced by the static-field arising from the charged surface. Any difference between DL-SFG spectra of interfaces with different surface charges can thus be related to the difference in the average orientation of the DL-water molecules.

In order to explicitly show the dependence of the DL-SFG signal on the water orientation in Eq. 5, $\beta_{DL}^i(\omega)$ can be divided and multiplied by the orientation of each DL-molecule, denoted $\cos\theta_i$. Specifically, $\cos\theta_i$ is the cosine of the angle formed between the dipole direction of the i -th water molecule and the normal to the surface (defined from the liquid to the solid/vapor phase).

Assuming that differences between the SFG-activity of the DL-water molecules only arise from their orientation and that $\beta_{DL}^i(\omega)$ linearly depends on $\cos\theta_i$, so that $\frac{\beta_{DL,xxz}^i(\omega)}{\cos\theta_i} = \beta_{DL,xxz}^{eff}(\omega)$ is a constant for each given ω , we obtain:

$$\chi_{DL}^{(2)}(\omega) = \sum_{i=1}^{N_{DL}} \frac{\beta_{DL}^i(\omega)}{\cos\theta_i} \cdot \cos\theta_i \quad (6)$$

$$= \beta_{DL}^{eff}(\omega) \cdot \sum_{i=1}^{N_{DL}} \cos\theta_i \quad (7)$$

where $\frac{\beta_{DL}^i(\omega)}{\cos\theta_i} = \beta_{DL}^{eff}(\omega)$ is defined as the effective hyperpolarizability of the i -th DL-water molecule, constant for all DL-water molecules for each given ω .

This equation allows separating the orientation part of any DL-SFG spectrum (due to the surface charge), i.e. $\sum_{i=1}^{N_{DL}} \cos\theta_i$, from the $\beta_{DL}^{eff}(\omega)$ term, which is the contribution to the SFG spectrum of a DL water molecule per unitary orientation. This latter is identical for all DL-water molecules at the charged interface. It is important to stress here that Eq. 7 and Eq.2 are equivalent under the assumption that orientation dominates the DL-SFG activity (which is indeed the usual assumption). The $\sum_{i=1}^{N_{DL}} \cos\theta_i$ orientation term in Eq. 7 is hence the counterpart of $\Delta\phi_{DL}$ in Eq.2, at the origin of the sign and intensity of DL-SFG spectra, while $\beta_{DL}^{eff}(\omega)$ is the counterpart of $\chi_{bulk}^{(3)}(\omega)$, providing the two bands structure of any DL-SFG spectrum.

In order to prove that $\beta_{DL}^{eff}(\omega)$ is, as $\chi_{bulk}^{(3)}(\omega)$, a universal term at any aqueous interface, and therefore that Eq. 2 and Eq. 7 are equivalent formulations, we now consider the same DFT-MD simulations of charged aqueous interfaces as the ones used in refs. 19 and 28, which allowed us to obtain the theoretical $\chi_{bulk}^{(3)}(\omega)$ using Eq. 2 in order to calculate $\beta_{DL}^{eff}(\omega)$ from the theoretical $\chi_{DL}^{(2)}(\omega)$ spectra via Eq. 7. The systems considered include both neat and electrolytic silica-water interfaces with different surface charge values and ions concentrations. The neat interfaces are: the fully protonated (0001)- α Quartz-water (QW) interface, QW with 3% of surface SiOH sites being deprotonated, and QW with 12% of deprotonated sites. As regards the three electrolytic interfaces, they are all obtained by adding different concentrations of KCl ion-pairs in the simulation box of the fully hydroxylated QW interface. The resulting excess concentrations of K^+ in the BIL, $[K^+]_{BIL}$, is 1.6 M, 4.2 M and 7.2 M for the investigated three systems. We refer to ref. 28 for more details on these DFT-MD simulations.

As shown in Fig. 1, similar $\beta_{DL}^{eff}(\omega)$ spectra are obtained for all considered systems. They have the same shape as the $\chi_{Bulk}^{(3)}(\omega)$ spectrum,^{19,28} with the same two positive bands at 3200 and 3400 cm^{-1} . With this universal $\beta_{DL}^{eff}(\omega)$ parameterized based on these DFT-MD simulations, we are now ready to calculate DL-SFG spectra

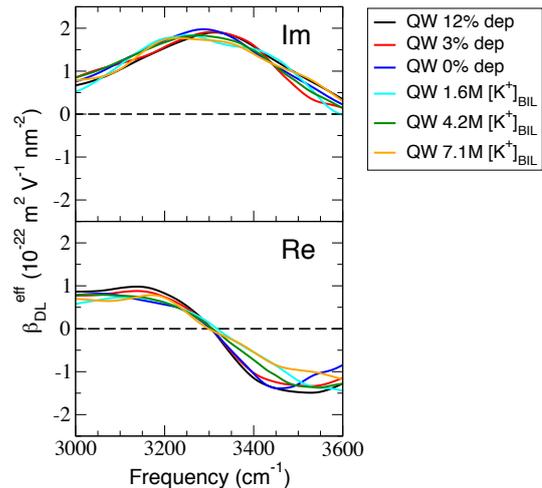


FIG. 1. $\beta_{DL}^{eff}(\omega)$ spectra (imaginary, Im and real, Re components) calculated using Eq.7 for the DFT-MD simulations already used to parametrize $\chi^3(\omega)$ in refs.^{19,28}: Quartz-water interfaces with different degrees of surface deprotonation (0%, 3% and 12%) and ions (KCl) concentrations, reported as excess cation (K^+) concentration in the BIL (1.6, 4.2 and 7.1 M).

from any MD simulation (DFT-MD or classical MD) or even from Monte Carlo simulations by multiplying the constant $\beta_{DL,xxz}^{eff}(\omega)$ spectrum (Fig. 1) by the orientation term of the water molecules in the DL ($\sum_{i=1}^{N_{DL}} \cos\theta_i$) obtained from simulations without the need to evaluate correlation functions.

B. SFG Signal in the Binding Interfacial Layer (BIL)

Compared to the DL, the water structural organization in the Binding Interfacial Layer (BIL) is more complex and cannot be simply described by the orientation of the water molecules. The reason is that several water populations contribute to the BIL spectrum, which are characterized by distinct local coordination and orientation and results from the balance between water-water and water-surface interactions. We hereby define such populations based on the water OH groups being either H-bonded to water, or to the surface, or non H-bonded and on their orientations with respect to the surface normal. These criteria are easily obtained as output of simulations, making the classification of BIL populations and their statistical analysis straightforward. The types and relative abundances of such BIL-populations are surface specific as they depend on their interaction strength with surface functional groups as well as on the local surface morphology, topology, hydrophilicity, charge state, etc. However, we will show in the following that a small number of spectroscopically relevant populations is sufficient to describe various families of aqueous interfaces (e.g. hydrophobic interfaces, or interfaces between water and or-

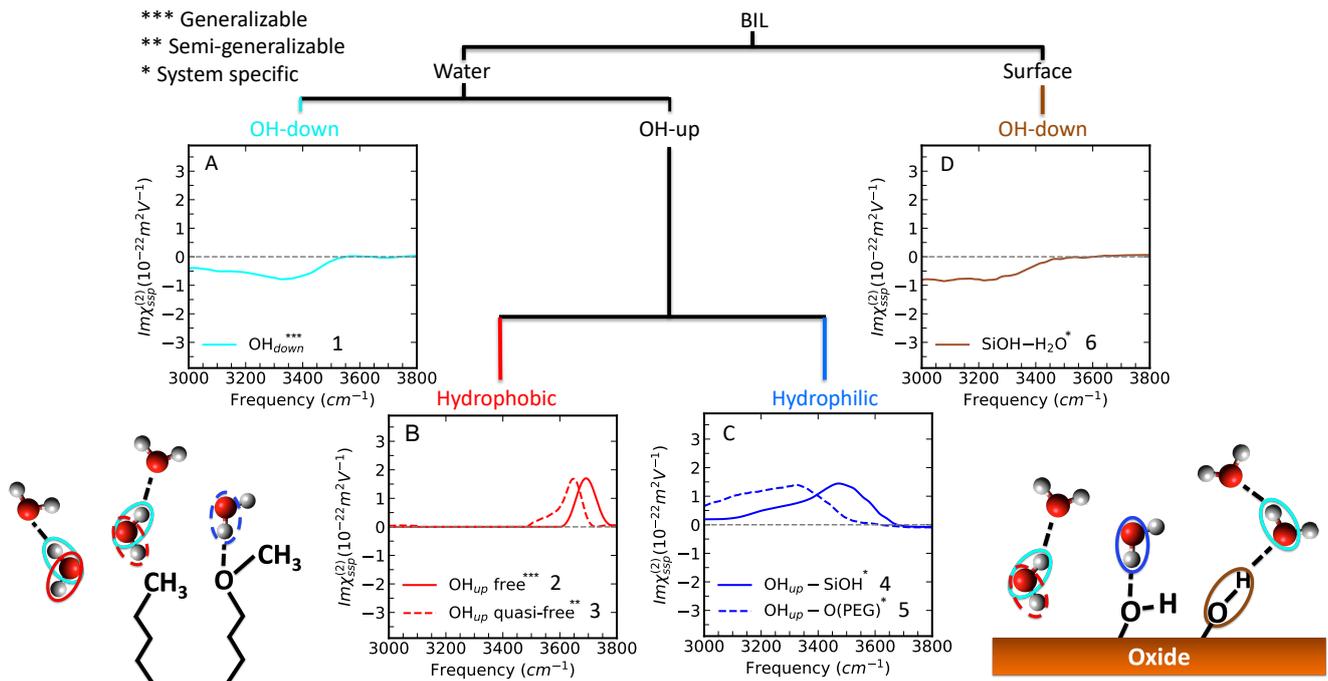


FIG. 2. Spectroscopic fingerprints $\beta^{eff}(\omega)$ of BIL-water OH groups and surface SiOH groups in the OH stretch frequency range, parameterized from various aqueous interfaces including the air/water, amorphous silica/water, self-assembled monolayer/water interfaces. Parameterization on DFT-MD trajectories and VVCF SFG calculations (see text, section B).

ganic monolayers). As for the DL, we need to extract the $\beta^{eff}(\omega)$ of each defined population in the BIL, which requires building a database for each family of surfaces. In this work, the SFG fingerprints for the BIL-populations have been obtained thanks to extensive studies of various aqueous interfaces based on DFT-MD simulations and on calculated SFG spectra of the BIL using the VVCF method^{18–22,28}.

These fingerprints of BIL-water populations were extracted from aqueous interfaces ranging from pure hydrophobic to hydrophilic, with in between the interfaces having coexistent local hydrophobic and hydrophilic patches. They are summarized in Fig. 2. These fingerprints serve as building blocks for the “pop model” of SFG calculation discussed in the following.

BIL-water OH groups are first categorized into (1) OH groups pointing towards the subsequent water layer (opposite to the conventional SFG normal vector, denoted “OH-down”, $\cos\theta < 0$ with θ being the angle between an OH group and the SFG normal vector, i.e. perpendicular to the surface, pointing from water to the surface), resulting in a negative SFG band, and (2) these pointing towards the surface (along the conventional SFG normal vector, denoted “OH-up”, $\cos\theta > 0$), leading to a positive SFG band. The spectroscopic fingerprint of “OH-down” groups (Fig. 2A, solid cyan line) is generalizable for any investigated aqueous interface as these groups are systematically HB-donors to bulk (or DL) water molecules, as demonstrated in ref. 22. This population is denoted

pop 1 in the following. However, the “OH-up” fingerprints are system dependent due to the strength of interaction between BIL-water and the top-surface sites.

At hydrophobic interfaces, the “OH-up” groups are dangling, i.e. pointing to the surface without forming H-bonds, resulting in a free-OH peak located at above 3600 cm^{-1} in SFG spectrum. Note that the free-OH peak is usually recognized as a molecular proof of hydrophobicity^{22,29,30}. In the case of the air/water interface, the prototype hydrophobic interface, the dangling OH groups pointing towards the air is free, corresponding to a positive SFG peak at 3700 cm^{-1} , denoted pop 2 (Fig. 2B, solid red line).

At other hydrophobic interfaces, e.g. graphene, or hydrophilic interfaces with local hydrophobic patches, like the amorphous silica with low hydroxylation degree, the dangling OH groups point towards the hydrophobic functional groups (i.e. siloxanes) with weak interactions (e.g. van der Waals force), leading to a red-shifted free OH peak (denoted quasi-free, Fig. 2B, dashed red line) compared to that of the air/water interface²². The representative fingerprint depicted in Fig. 2B with the dashed red line is parameterized from SFG fingerprint arising from OH groups pointing to hydrophobic patches of amorphous silica surface with a hydroxylation degree of 4.5 SiOH/nm².^{20,22} Note that the quasi-free OH peak frequency shows very little variation between hydrophobic interfaces ($\pm 15\text{ cm}^{-1}$) investigated in ref. 22, which makes the representative fingerprint semi-transferable to

other hydrophobic interfaces. This quasi-free OH population is denoted pop 3.

In the case of hydrophilic interfaces, “OH-up” groups are H-bonded to surface hydrophilic functional groups, corresponding to a positive band in the H-bonded frequency range (below 3600 cm^{-1}). The band shape and frequency center of this peak is system specific, depending on the types of surface hydrophilic functional groups. Here, we present two examples of H-bonded OH-up fingerprints, one parameterized from the water OH groups forming donor H-bonds with surface SiOHs at an amorphous silica/water interface with a hydroxylation degree of $4.5\text{ SiOH/nm}^{2,22}$ (denoted pop 4, Fig. 2C, dashed blue), and the other from water OH groups forming donor H-bonds with oxygen atoms of the organic self-assembled monolayer (SAM) composed of polyethylene glycol (PEG, $\text{CH}_3(\text{OCH}_2\text{CH}_2)_7\text{CH}_2\text{SiCl}_3$), denoted pop 5, see Fig. 2C, solid blue. While being system dependent, these fingerprints can be utilized to reconstruct the SFG spectra of interfaces within the same family but with different parameters, e.g. different degrees of hydroxylation and crystallinity of silica/water interfaces, different ratios of PEG polymer chains involved in the SAM monolayer composition. These examples will be demonstrated in the following section of applications.

One last population is hereby defined for SFG signals in the OH stretch frequency range arising from surface functional groups. This population, same as pop 4 and 5, is again system dependent. Among the investigated interfaces in our database, the only SFG-active surface group in the OH stretch range arises from surface SiOHs^{20,21,31}, which is hence defined as pop 6 in the present work (Fig. 2D, solid brown line).

With the $\beta^{eff}(\omega)$ of different populations being parameterized from DFT-MD simulations of various aqueous interfaces through the VVCF-based SFG calculation method and shown in Fig. 2, we are now able to construct the SFG spectrum of any given interface within the family of the database by determining the statistics of each population from MD simulations (either DFT-MD or standard classical MD) in analogy to Eq. 7 for the DL but now summing over all possible i -th populations:

$$\chi_{BIL}^{(2)}(\omega) = \sum_{i=1}^{N_{pop}} \beta_i^{eff}(\omega) \cdot \left(\sum_{j=1}^{n_{pop_i}} \cos\theta_{ij} \right) \quad (8)$$

Note that unlike in the DL where water orientation dominates the SFG activity, the orientation of the water OH groups is dictated by the water-surface or water-water interactions, which is intrinsically taken into account in the definition of the populations.

C. Applications

1. Hydrophobic Interfaces

We start our demonstration of “pop model” for SFG calculations by hydrophobic aqueous interfaces. The air/water interface has been extensively studied through SFG spectroscopy^{4,30,33–35} during the last decades. The $\Im(\chi_{BIL}^{(2)}(\omega))$ of the air/water interface is composed of a positive peak located at $\sim 3680\text{ cm}^{-1}$ (in the free OH region) and a negative band centered at $\sim 3400\text{ cm}^{-1}$ (in the H-bonded region). Apart from the free OH peak, a two-dimensional H-bond extended network (2DN)¹⁸ composed of interconnected interfacial water molecules parallel to the water surface was uncovered. The 2DN is another molecular hydrophobic descriptor, as shown in refs. 18,22,36, which is however SFG inactive in *ssp* or *ppp* polarization due to the orientation of OH groups of water perpendicular to the SFG normal vector. The theoretical $\Im(\chi_{BIL}^{(2)}(\omega))$ of the air/water interface calculated from DFT-MD simulation using the VVCF based method^{17,18} was shown in good agreement with experimental SFG spectrum of air/water interface¹⁸.

With these structural and spectroscopic insights on the air/water interface, the SFG signal of this interface can then be reconstructed by using two of our defined populations: OH_{down} population (pop 1, Fig. 8A for $\beta_{effective}(\omega)$) and OH_{up} free (pop 2, Fig. 8B for $\beta_{effective}(\omega)$) as building blocks. By calculating each population density from a classical MD simulation using the SPC/E force field³² with the same box size and number of water molecules as in the DFT-MD simulation used for the SFG calculation in ref. 18, but larger time scale (20 ns), (see ref. 37 for simulation details), we then obtain the SFG spectrum of the air/water interface constructed from our “pop model” in Eq. 8, shown in Fig. 3A, upper panel. The excellent agreement between the obtained SFG spectrum over a classical FF-MD trajectory and the DFT-MD calculated spectrum through VVCF approach validates the “pop model”, which makes it possible to construct a SFG spectrum in good quality using classical MD simulations without a specialized force field and without calculating any kinds of correlation functions.

The SFG spectra of other hydrophobic interfaces, e.g. graphene/water, are similar to that of air/water interface as demonstrated in ref. 22. In graphene SFG spectrum calculated from DFT-MD simulation using VVCF method, we recognize a similar positive free OH peak and a negative H-bonded peak as the ones in air/water SFG spectrum. Hence, we use the same two-population model to construct the SFG spectrum by calculating the density of each population from the DFT-MD simulation (see ref. 22 for simulation details). However, the free OH peak in this case is red shifted compared to that of the air/water interface due to the weak interactions between water and the surface. This is well captured by the OH_{up} quasi-free population (pop 3), while the negative

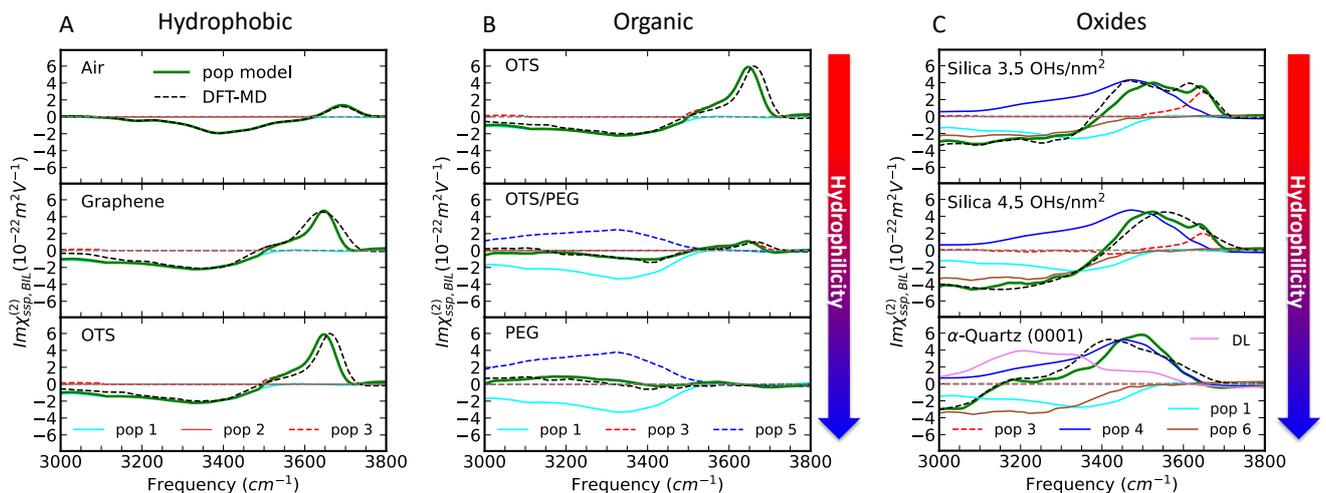


FIG. 3. BIL-SFG $\Im(\chi_{BIL}^{(2)}(\omega))$ spectra of various hydrophobic/hydrophilic interfaces calculated using “pop model” approach in solid green lines. (A) Hydrophobic interfaces. Top: the air/water interface calculated over a classical MD simulation of 20 ns (SPCE³² water); Middle: graphene/water interface calculated over a DFT-MD simulation of 50 ps. Bottom: OTS/water interface calculated over a DFT-MD simulation of 400 ps. (B) Organic SAM/water interfaces SFG spectra calculated over DFT-MD simulations of 400 ps using the “pop model”. Top: OTS/water interface; Middle: PEG/water interface; and Bottom: homogeneously (1:1) mixed OTS-PEG/water interface. (C) Silica/water interfaces with different degrees of crystallinity and hydroxylation. Top: amorphous silica/water interface with a hydroxylation degree of 3.5 OHs/nm²; Middle, amorphous silica/water interface with a hydroxylation degree of 4.5 OHs/nm²; Bottom: α -Quartz (0001)/water interface with added DL contribution (in pink) for a direct comparison with the DFT-MD calculated spectrum reported in ref. 21. The “pop model” SFG spectra are calculated over DFT-MD trajectories, respectively of 20 ps (top), 20 ps (middle) and 15 ps (bottom). These trajectories can be found in refs. 20,21. The spectral contributions from each population are depicted for each system using the same color code and line style as in Fig. 2. In all plots, the VVCF based DFT-MD calculated spectra are plotted in dashed black for comparison with the “pop model” spectra in green.

band arises from the same OH_{down} (pop 1). The SFG spectrum deduced from our “pop model” (Fig. 3A, middle panel) well matches the one calculated from DFT-MD simulation in terms of frequency and intensity.

2. Organic Interfaces

To increase the complexity of interfaces, we investigate a set of aqueous interfaces composed of self-assembled monolayers (SAMs) with well-controlled and flexibly tunable hydrophilicity by playing with the ratio of different functional groups. These SAM/water interfaces have attracted considerable attention as a model platform to study hydrophobicity and hydrophilicity both theoretically^{38–40} and experimentally^{12,41–43}. Water SFG spectra³⁶ of three different silane-based SAM/water interfaces have been calculated from DFT-MD simulations of: (1) pure octadecyltrichlorosilane (OTS, $\text{CH}_3(\text{--CH}_2)_{17}\text{--SiCl}_3$) monolayer (intrinsically hydrophobic), (2) pure polyethylene glycol (PEG, $\text{CH}_3(\text{OCH}_2\text{CH}_2)_7\text{--CH}_2\text{SiCl}_3$) monolayer (intrinsically hydrophilic with oxygens at the interface), and (3) homogeneously mixed OTS-PEG monolayer (1:1 mixing ratio).

We start from the OTS/water interface, which in principle possesses the same two populations as other

hydrophobic interfaces. The water SFG spectrum of OTS/water interface is reconstructed by the “pop model” and displayed in Fig. 3A, lower panel. Same as for graphene/water interface, the frequency shift of OH_{up} with respect to that of air/water is taken into account by the OH_{up} quasi-free (pop 3) fingerprint. The good agreement with DFT-MD calculated SFG spectrum once again validates our choice of the two populations for hydrophobic interfaces. We reiterate that the discrepancy in frequency of the OH_{up} between the DFT-MD calculated and “pop model” constructed SFG spectrum (within 15 cm^{-1}) does not prevent us from correctly interpreting the spectrum.

Turning now to a hydrophilic interface, the $\Im(\chi_{BIL}^{(2)}(\omega))$ of pure PEG/water interface calculated from DFT-MD simulations using VVCF-based method surprisingly shows almost zero signal (Fig. 3B, middle panel, dashed black). This is however due to the compensation between the contributions of 2 populations: OH_{down} (pop 1) and the specific $\text{OH}_{up}\text{-O(PEG)}$ (pop 5) for aqueous interfaces with the presence of PEGs. The spectral contributions arising from each population are also depicted in this plot using the same color code and line style as in Fig. 2, with pop 1 in solid cyan and pop 5 in dashed blue. Compensation of the two contributions is due to: (1) the $\text{OH}_{up}\text{-O(PEG)}$ and OH_{down} spectra

have the same center frequency, implying the same strength of the hydrogen bond between water-water and water-O(PEG) and (2) they have equivalent density of population. The zero SFG signal of PEG/water interface is thus due to the compensation between the spectral contributions from the two populations.

The homogeneously mixed OTS-PEG/water interface, as the combination of OTS/water and PEG/water, has both $\text{OH}_{up}\text{-O(PEG)}$ (pop 5) and OH_{up} quasi-free (pop 3) as OH-up populations, depending on whether the OH-up is in proximity and H-bonded to a PEG or pointing to an OTS and remaining dangling. Three populations (pop 1, 3, 5) contribute to the SFG spectrum. After evaluating each corresponding population density from the DFT-MD simulation, the water SFG spectrum of the mixed OTS-PEG/water interface is hence obtained using Eq. 8, in perfect agreement with the one calculated from DFT-MD simulation using VVCF method, see Fig. 3B, lower panel.

The resultant SFG spectrum has a tiny free OH peak corresponding to dangling OH groups pointing towards OTS and a well reduced negative peak in the H-bonded region compared to that of other hydrophobic interfaces due to the compensation of signals arising from up OHs H-bonded to PEGs and down OHs H-bonded to subsequent water layer, as indicated by the separated population contributions from pop 1 (solid cyan) and pop 5 (dashed blue).

3. Aqueous Silica

For the systems demonstrated above, the SFG signals in the OH stretch region solely arise from the O-H stretch of water molecules. However, for some aqueous interfaces with the presence of surface hydroxylated functional groups (e.g. silica/water interface), the surface OHs also contribute to the SFG spectra, whose signals are often overlapped with contributions from water OHs in the H-bonded region.

This is the case of silica-water interfaces. The organization of interfacial water molecules and the corresponding SFG spectra varies from system to system as a function of degrees of hydroxylation and crystallinity of the silica surface^{19–21,27,44,45}. We show in Fig. 3C that these variations are fully captured by the “pop model” and rationalized in changes in the relative abundance of hydrophilic and hydrophobic populations. Though being macroscopically hydrophilic, local hydrophobic patches consisting of siloxane bridge (Si-O-Si) coexist with hydrophilic functional groups (SiOH). Similar to mixed OTS-PEG/water interface, we thus find that three BIL-water populations are responsible for the SFG spectrum of a set of silica/water interfaces, i.e. OH_{down} (pop 1), OH_{up} quasi-free (pop 3) and $\text{OH}_{up}\text{-SiOH}$ (pop 4). In addition to BIL-water populations, surface silanol groups perpendicular to the surface pointing out-of-plane to water (pop 6) also contributes to SFG spectrum in the O-H

stretch frequency range.

Three silica/water interfaces are investigated: two amorphous silica surfaces with a hydroxylation degree of 3.5 OHs/nm² and 4.5 OHs/nm², together with one crystalline surface of α -quartz (0001), all fully hydroxylated (corresponding to pH < 4 conditions). The corresponding SFG spectra are calculated by “pop model” and depicted in Fig. 3C, with each population contribution illustrated in the corresponding color code and line style as shown in Fig. 2. The SFG spectra of amorphous silica/water interfaces of both 3.5 and 4.5 OHs/nm² are in good agreement with that from VVCF DFT-MD method. The individual spectral contributions from each population manage to capture the decrease of ratio between spectral intensities of OH_{up} quasi-free (pop 3) and $\text{OH}_{up}\text{-SiOH}$ (pop 4), with increasing degree of hydroxylation (from 3.5 to 4.5 OHs/nm²), informing that more BIL-water OHs are H-bonded to SiOHs and less are pointing to the siloxane bridge being quasi-free when the hydroxylation degree increases.

In case of quartz/water interface, unlike the neutral amorphous silica/water interfaces, aqueous α -quartz (0001) surface is slightly charged despite under PZC condition, resulting from the dipole moment triggered by alternatively distributed in-plane and out-of-plane silanols²¹. This results in a $\Im(\chi_{DL}^{(2)}(\omega))$ contribution to the total $\Im(\chi^{(2)}(\omega))$ compared to the other two amorphous silica/water interfaces, where $\Im(\chi_{DL}^{(2)}(\omega))$ equals zero. The DL-water contribution calculated with the “pop model” from Eq. 7 is added for direct comparison with VVCF DFT-MD calculated spectrum. Note that the “pop model” confirms that SFG-DL equals zero for the amorphous surfaces. The “pop model” calculated SFG spectrum has a blue shifted positive peak compared to that of VVCF DFT-MD calculated. Since the $\text{OH}_{up}\text{-SiOH}$ is parameterized from amorphous silica/water interface, this frequency shift due to stronger H-bonds between BIL and SiOH of crystalline silica surface can thus not be captured. However, this does not impede us from interpreting spectrum by attributing each spectral contribution to the corresponding population and capturing the population variation of interface components under different conditions, e.g. pH, ionic strength, etc.

4. The “Pop Model” Applied to Surface Chemistry in the Phonon Range

Up to now, as regards surfaces, we only recognize one population (pop 6) corresponding to out-of-plane silanol groups pointing towards and H-bonded to water as a SFG-active population in the O-H stretch range. This population does not inform on the surface structure organization of silica. To unravel the surface chemistry of silica in contact with water, we now go from the high O-H stretch range (3000 – 3800 cm⁻¹) to the low phonon frequency range, corresponding to the Si-O stretch region (750 – 1100 cm⁻¹).

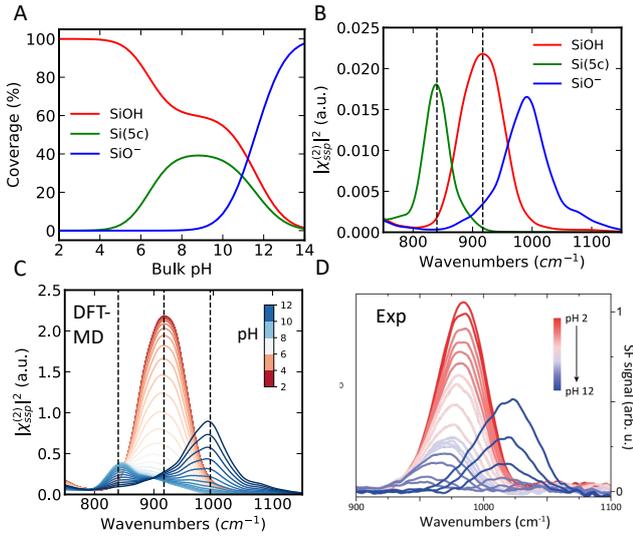


FIG. 4. Theoretical pH-dependent SFG calculation: (A) Surface coverages of each population as a function of pH deduced by Henderson-Hasselbalch equations based on the pKa values of surface silanols from metadynamics simulations. (B) SFG fingerprints in the Si-O stretch frequency range of each surface population calculated from DFT-MD simulations based on VVCF method. (C) Theoretical pH-dependent SFG spectra of amorphous silica surface in contact with water in the Si-O stretch frequency range constructed by “pop model” through combination of (A) and (B). (D) Experimental pH-dependent SFG spectra of amorphous silica surface in contact with water in the Si-O stretch frequency range probed by W. Liu *et al.*

The molecular understandings on surface chemistry of silica surface in contact with water have long been a hard nut to crack due to difficulties in probing buried oxide surfaces⁴⁶. In 1992, Eissenthal and coworkers⁴⁷ first employed second harmonic generation (SHG) technique to monitor the silica/water interface under different pH conditions and discovered two pKa values associated to surface silanol deprotonation, giving rise to the historical bimodal titration behavior, which have been extensively discussed in the literature. A substantial amount of both experimental^{48–50} and theoretical works^{51–54} have been dedicated to investigating the microscopic origins of the bimodal behavior of silica surface acidities.

Despite extensive works on the water side, only a limited amount of research has been conducted on the oxide side due to the difficulties of experimentally probing buried surfaces of insulating solids. Hence, the surface chemistry behind the bimodal or trimodal behavior has been highly debatable without the understanding of surface reconstruction triggered by deprotonation of surface silanols under different pH conditions.

In order to fully uncover the surface chemistry of silica in contact with water, we carry out DFT-MD simulations and aim to calculate SFG spectra in the Si-O stretch frequency range to elucidate the silica surface structure under different pH conditions.

The results from metadynamics and DFT-MD simulations revealed surface chemistry on the amorphous aqueous silica surface under different pH conditions. In the pH range between pH~2 (Point of zero charge for silica) to pK1, the surface is fully hydroxylated. At pK1 < pH < pK2, the acidic silanols start to be deprotonated, each deprotonated oxygen attacks and covalent bonds to a neighbouring Si-O(H), forming a five-coordinated Si species [denoted Si(5c)]. When pH rises above pK2, basic SiOH groups are deprotonated to SiO⁻ and Si(5c) species become unstable, split up and form SiO⁻.

With this picture, we hence deduce the surface coverages of each surface species/population (SiOH, Si(5c) and SiO⁻) as a function of pH through Henderson-Hasselbalch equations. The determined coverage of each surface population (corresponding to population density) as a function of pH is depicted in Fig. 4A. The SFG fingerprints (corresponding to $\beta^{eff}(\omega)$) of each surface population is calculated using the same VVCF method as for water in the O-H stretch region from unbiased DFT-MD simulations and displayed in Fig. 4B. (detailed in Section computational methodology).

The pH-dependent theoretical SFG spectra of silica surface in the Si-O stretch region are thus constructed using the “pop model”, as for water O-H spectra, without requiring calculations for each DFT-MD simulation at a certain pH. In Fig. 4C, we observe that at pH 2, where the silica surface is fully hydroxylated and covered by SiOH, the spectrum showed a single peak arising from Si-OH stretching vibration at ~ 920 cm⁻¹. As pH increases, silica surface starts to be deprotonated, resulting in a decrease in the intensity of the Si-OH peak, accompanied by a redshift to ~ 840 cm⁻¹ before disappearing at \sim pH 10, corresponding to the signal of Si(5c). Above pH 10, a new mode at ~ 1000 cm⁻¹ emerged, which is attributed to the Si-O⁻ stretching mode.

An experimental scheme enabling in situ vibrational spectroscopy of oxide surfaces in liquid water were developed by W. Liu *et al.* and applied on amorphous silica surface in water to record the pH-dependent SFG spectra of silica in the Si-O stretch region (Fig. 4D). The nice agreement between experimental pH-dependent SFG spectra and theoretical SFG spectra calculated through “pop model” validate at the same time the “pop model” and the newly discovered Si(5c) species.

CONCLUSIONS

Interpreting experimental SFG spectrum of aqueous interfaces has long been non-trivial due to the overload information contained in overlapped signals arising from different interfacial components. Calculating theoretical SFG spectroscopy based on time-dependent formalism¹¹ from MD simulations allows us to elucidate different spectral contributions through spectrum deconvolution. However, the high computational cost for step-by-step dipole moment and polarizability evaluation us-

ing DFT-MD simulations and the difficulty in designing system-specific force fields for classical simulations limited the transferability of this calculation method.

In the present work, we introduce a new approach for SFG spectroscopy calculation of aqueous interfaces—“pop model”—that constructs the SFG spectra by recognizing different interface populations from either DFT-MD or classical MD simulations. The SFG fingerprints of defined populations, served as building blocks for the “pop model” calculation, are parameterized from DFT-MD simulations using VVCF-based calculated SFG spectra of numerous aqueous interfaces. Due to different physics behind each part of the interface, the choice of populations is treated separately for water diffuse layer (DL), water binding interfacial layer (BIL) and surface.

For water in the diffuse layer, known as bulk-like water reoriented by static electric field generated by surface charge²³, only one orientation-dependent population is defined, parameterized from a set of quartz/water interfaces with different surface charges, and demonstrated to be universal and transferable to other aqueous interfaces. Beyond the conventional 1-D model using the Gouy-Chapman model to predict DL-SFG intensities (through the electrostatic potential) at charged aqueous interfaces^{23,25,55}, the DL spectrum deduced by the orientation-based “pop model” manage to capture the non-negligible effect of lateral distribution of charges at the surface (homogeneously or heterogeneously) on the SFG response of DL water.

As regards water in the binding interfacial layer, the population definition is in principle more system dependent. However, with a few parameterized SFG signatures, we manage to construct and interpret the SFG spectra of water in contact with diversified surfaces (from hydrophobic to hydrophilic and from organic to oxide surfaces) using “pop model”. The in general good agreement between “pop model” calculated and VVCF-MD calculated spectra validated the new SFG calculation approach for interfacial water.

Finally, in terms of surface characterization, we manage to construct the pH-dependent theoretical SFG spectra in Si-O stretch region using the same “pop model” principle without requiring tuning proton concentration to set the pHs in the simulations. We first calculate pKa values of each surface silanol group through metadynamics simulations and discovered two populations of silanols with acidic and basic pKa values, respectively, in line with the long discussed bimodal behavior of silica surface acidities first uncovered by Eisenthal and coworkers⁴⁷. Based on these pKa values, we deduce the surface coverages (population densities in “pop model”) of each species at given pH through Henderson-Hasselbalch equations. After calculating the SFG fingerprints arising from each surface population (SiOH, Si(5c) and SiO⁻), we construct the theoretical SFG spectra at a given pH by summing up SFG signals from each surface species, knowing the density of each surface population at given pH, in nice agreement with experimental pH-dependent

SFG spectra of aqueous amorphous silica surface measured by W. Liu et al.

Summing up the spectroscopic contributions from all interfacial populations, we hence obtain a complete theoretical SFG spectrum of a chosen aqueous interface through the “pop model” calculation. In addition to the computational expediency, the presented SFG calculation approach allows for a straightforward interpretation of the SFG spectra with a clear recognition of the spectroscopic contribution of each interfacial population, thus providing an in-depth understanding of the interfacial structures of aqueous interface.

COMPUTATIONAL METHODOLOGY

The details of DFT-MD and classical MD simulations with SPC/E force field of air/water interface are presented in ref. 37. The DFT-MD simulation details of graphene/water interface and silica/water interface with a hydroxylation degree of 3.5OHs/nm² are introduced in ref. 22. We refer to ref. 21 for the DFT-MD simulation details of silica/water interface with a hydroxylation degree of 4.5OHs/nm² and of α -quartz (0001)/water interface. The simulation details for SAM/water interface are presented in ref. 36. We refer to ref. 56 for simulation details of metadynamics for pKa calculations of surface silanols.

All the VVCF-based DFT-MD calculated SFG spectra are calculated following the approach introduced in ref. 17. For the water contribution to $\chi^{(2)}(\omega)$, individual molecular dipole moments and polarizability tensors are calculated with the model from ref 17, supposing that only the OH stretching motions contribute to the spectrum in the high frequency region ($> 3000 \text{ cm}^{-1}$). All details of derivation for the expression of SFG can be found in refs 19,21 with the parameterization of APT and Raman tensors of water found in ref 17. For the out-of-plane SiOH contribution to the SFG spectrum in the OH stretch region, the parameterization of APT and Raman tensors are adopted from ref. 21. For SFG spectra calculated in the Si-O stretch region, the parameterization of APT and Raman tensors are set to unity since we are more interested in the influence on the SFG intensity as the result of a change in group population (as it happens when changing the pH) than in the change in the SFG activity of the group. For a more accurate description of the activity, the APT and Raman tensors must be rigorously evaluated from ab initio calculations.

For all structural analyses, we have adopted the H-bond definition proposed by White and co-workers⁵⁷, with $O(-H)\cdots O \leq 3.2 \text{ \AA}$ and the $O-H\cdots O$ angle between 140° and 220° .

ACKNOWLEDGMENTS

AUTHOR DECLARATIONS

Conflict of interest

The authors have no conflicts to disclose.

DATA AVAILABILITY STATEMENT

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

- ¹J. Hunt, P. Guyot-Sionnest, and Y. Shen, *cpl*, 1987, **133**(3), 189–192.
- ²X. Zhu, H. Suhr, and Y. Shen, *prb*, 1987, **35**(6), 3047.
- ³Y. Shen, *Fundamentals of sum-frequency spectroscopy*, Cambridge University Press, 2016.
- ⁴Q. Du, R. Superfine, E. Freysz, and Y. Shen, *Phys. Rev. Lett.*, 1993, **70**(15), 2313.
- ⁵A. Eftekhari-Bafrooei and E. Borguet, *J. Am. Chem. Soc.*, 2010, **132**(11), 3756–3761.
- ⁶Z. Zhang, L. Piatkowski, H. J. Bakker, and M. Bonn, *Nat. Chem.*, 2011, **3**(11), 888–893.
- ⁷C.-S. Hsieh, M. Okuno, J. Hunger, E. H. Backus, Y. Nagata, and M. Bonn, *Angew. Chem. Int. Ed.*, 2014, **53**(31), 8146–8149.
- ⁸C. Tian and Y. Shen, *Surf. Sci. Rep.*, 2014, **69**(2-3), 105–131.
- ⁹A. M. Jubb, W. Hua, and H. C. Allen, *Annu. Rev. Phys. Chem.*, 2012, **63**, 107–130.
- ¹⁰L. Dalstein, E. Potapova, and E. Tyrode, *Phys. Chem. Chem. Phys.*, 2017, **19**(16), 10343–10349.
- ¹¹A. Morita and J. T. Hynes, *J. Phys. Chem. B.*, 2002, **106**, 673–685.
- ¹²A. Morita and T. Ishiyama, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5801–5816.
- ¹³Y. Ni and J. L. Skinner, *J. Chem. Phys.*, 2016, **145**(3), 031103.
- ¹⁴T. Ishiyama, T. Imamura, and A. Morita, *Chem. Rev.*, 2014, **114**(17), 8447–8470.
- ¹⁵T. Ishiyama and A. Morita, *Ann. Rev. Phys. Chem.*, 2017, **68**, 355–377.
- ¹⁶T. Ohto, K. Usui, T. Hasegawa, M. Bonn, and Y. Nagata, *jcp*, 2015, **143**(12), 124702.
- ¹⁷R. Khatib, E. H. Backus, M. Bonn, M.-J. Perez-Haro, M.-P. Gaigeot, and M. Sulpizi, *Sci. Reports*, 2016, **6**(1), 1–10.
- ¹⁸S. Pezzotti, D. R. Galimberti, and M.-P. Gaigeot, *J. Phys. Chem. Lett.*, 2017, **8**(13), 3133–3141.
- ¹⁹S. Pezzotti, D. R. Galimberti, Y. R. Shen, and M.-P. Gaigeot, *Phys. Chem. Chem. Phys.*, 2018, **20**, 5190–5199.
- ²⁰J. D. Cyran, M. A. Donovan, D. Vollmer, F. S. Brigiano, S. Pezzotti, D. R. Galimberti, M.-P. Gaigeot, M. Bonn, and E. H. Backus, *Proc. Natl. Acad. Sci.*, 2019, **116**(5), 1520–1525.
- ²¹S. Pezzotti, D. R. Galimberti, and M.-P. Gaigeot, *Phys. Chem. Chem. Phys.*, 2019, **21**(40), 22188–22202.
- ²²S. Pezzotti, A. Serva, F. Sebastiani, F. S. Brigiano, D. R. Galimberti, L. Potier, S. Alfarano, G. Schwaab, M. Havenith, and M.-P. Gaigeot, *J. Phys. Chem. Lett.*, 2021, **12**(15), 3827–3836.
- ²³Y.-C. Wen, S. Zha, X. Liu, S. Yang, P. Guo, G. Shi, H. Fang, Y. R. Shen, and C. Tian, *Phys. Rev. Lett.*, 2016, **116**(1), 016101.
- ²⁴T. Joutsuka, T. Hirano, M. Sprik, and A. Morita, *Phys. Chem. Chem. Phys.*, 2018, **20**, 3040–3053.
- ²⁵G. Gonella, C. Lutgebaucks, A. G. F. de Beer, and S. Roke, *J. Phys. Chem. C.*, 2016, **120**, 9165–9173.
- ²⁶A. P. Willard and D. Chandler, *J. Phys. Chem. B.*, 2010, **114**(5), 1954–1958.
- ²⁷S.-h. Urashima, A. Myalitsin, S. Nihonyanagi, and T. Tahara, *J. Phys. Chem. Lett.*, 2018, **9**(14), 4109–4114.
- ²⁸S. Pezzotti, D. R. Galimberti, Y. R. Shen, and M.-P. Gaigeot, *Minerals*, 2018, **8**(7), 305.
- ²⁹F. Tang, T. Ohto, S. Sun, J. R. Rouxel, S. Imoto, E. H. Backus, S. Mukamel, M. Bonn, and Y. Nagata, *Chem. Rev.*, 2020, **120**(8), 3633–3667.
- ³⁰Q. Du, E. Freysz, and Y. R. Shen, *Science*, 1994, **264**(5160), 826–828.
- ³¹A. Tuladhar, S. Dewan, S. Pezzotti, F. S. Brigiano, F. Creazzo, M.-P. Gaigeot, and E. Borguet, *J. Am. Chem. Soc.*, 2020, **142**(15), 6991–7000.
- ³²H. J. Berendsen, J. v. Postma, W. F. Van Gunsteren, A. DiNola, and J. R. Haak, *The Journal of chemical physics*, 1984, **81**(8), 3684–3690.
- ³³C. Tainter, Y. Ni, L. a. Shi, and J. Skinner, *J. Phys. Chem. Lett.*, 2013, **4**(1), 12–17.
- ³⁴Y. Ni and J. Skinner, *J. Chem. Phys.*, 2016, **145**(3), 031103.
- ³⁵G. R. Medders and F. Paesani, *J. Am. Chem. Soc.*, 2016, **138**(11), 3912–3919.
- ³⁶W. Chen, S. E. Sanders, B. Ozdamar, D. Louaas, F. S. Brigiano, S. Pezzotti, P. B. Petersen, and M.-P. Gaigeot, *J. Phys. Chem. Lett.*, 2023, **14**, 1301–1309.
- ³⁷A. Serva, S. Pezzotti, S. Bougueroua, D. R. Galimberti, and M.-P. Gaigeot, *J. Mol. Struct.*, 2018, **1165**, 71–78.
- ³⁸J. I. Monroe and M. S. Shell, *Proc. Natl. Acad. Sci.*, 2018, **115**(32), 8093–8098.
- ³⁹B. C. Dallin and R. C. Van Lehn, *J. Phys. Chem. Lett.*, 2019, **10**(14), 3991–3997.
- ⁴⁰J. Monroe, M. Barry, A. DeStefano, P. Aydogan Gokturk, S. Jiao, D. Robinson-Brown, T. Webber, E. J. Crumlin, S. Han, and M. S. Shell, *Annu. Rev. Chem. Biomol.*, 2020, **11**, 523–557.
- ⁴¹C. Tian and Y. Shen, *Proc. Natl. Acad. Sci.*, 2009, **106**(36), 15148–15153.
- ⁴²S. Ye, S. Nihonyanagi, and K. Uosaki, *Phys. Chem. Chem. Phys.*, 2001, **3**(16), 3463–3469.
- ⁴³S. E. Sanders and P. B. Petersen, *J. Chem. Phys.*, 2019, **150**(20), 204708.
- ⁴⁴Q. Du, E. Freysz, and Y. R. Shen, *Phys. Rev. Lett.*, 1994, **72**, 238.
- ⁴⁵V. Ostroverkhov, G. A. Waychunas, and Y. Shen, *prl*, 2005, **94**(4), 046102.
- ⁴⁶U. Diebold, S.-C. Li, and M. Schmid, *Annu. Rev. Phys. Chem.*, 2010, **61**, 129–148.
- ⁴⁷S. Ong, X. Zhao, and K. B. Eisenthal, *cpl*, 1992, **191**(3-4), 327–335.
- ⁴⁸V. Ostroverkhov, G. A. Waychunas, and Y. Shen, *prl*, 2005, **94**(4), 046102.
- ⁴⁹J. D. Fisk, R. Batten, G. Jones, J. P. O’Reill, and A. M. Shaw, *jpcb*, 2005, **109**(30), 14475–14480.
- ⁵⁰A. M. Darlington and J. M. Gibbs-Davis, *jpcc*, 2015, **119**(29), 16560–16567.
- ⁵¹M. Sulpizi, M.-P. Gaigeot, and M. Sprik, *J. Chem. Theory Comput.*, 2012, **8**(3), 1037–1047.
- ⁵²M. Pfeiffer-Laplaud, D. Costa, F. Tielens, M.-P. Gaigeot, and M. Sulpizi, *jpcc*, 2015, **119**(49), 27354–27362.
- ⁵³F. Tielens, M. Gierada, J. Handzlik, and M. Calatayud, *Catal*, 2020, **354**, 3–18.
- ⁵⁴M. Gierada, F. De Proft, M. Sulpizi, and F. Tielens, *jpcc*, 2019, **123**(28), 17343–17352.
- ⁵⁵P. E. Ohno, S. A. Saslow, H. fei Wang, F. M. Geiger, and K. B. Eisenthal, *Nat. Comm.*, 2016, **7**, 13587–13591.
- ⁵⁶F. Siro Brigiano *Reactivity at aqueous interfaces by DFT-MD simulations* PhD thesis, Université Paris-Saclay, 2020.
- ⁵⁷J. A. White, E. Schwegler, G. Galli, and F. Gygi, *J. Chem. Phys.*, 2000, **113**(11), 4668–4673.

Chapter 5

Surface chemistry at the amorphous silica/water interface under different pH conditions

The results presented in this chapter have been obtained in collaboration with one of our experimental SFG partners: Prof. Wei-Tao Liu's group at Fudan University in Shanghai, China, in collaboration with Prof. Y.Ron Shen at the University of California, Berkeley, USA.

Note that the results regarding the pKa calculations by metadynamics were obtained during the PhD of Dr. Flavio Siro Brigiano, and detailed in his PhD manuscript, NNT: 2020UPASF025.

5.1 Historical but debated bimodal behavior of pKa values on silica/water interfaces

Aqueous interfaces of oxides not only play a significant role in the ecosphere on our planet [167], but also exploited in many modern technological applications, e.g. by hosting important reactions, such as the oxygen evolution reaction in renewable energy schemes [168]. However, the molecular insights on the chemistry of oxide surfaces in contact with water are still lacking due to difficulties in probing buried oxide surfaces. Even the most fundamental chemical reaction, the deprotonation equilibrium of silica surfaces as a function of water pH, is still hotly debated [46–48, 50–52]. Several interpretations have been proposed in the last decades, which are all based on the same elementary reaction $\text{SiOH} \leftrightarrow \text{SiO}^- + \text{H}^+$, but differ in terms of how many different populations of silanols with distinct pKa values exist on the surface, of the molecular origin of their distinct pKas, of the dependence of the pKa values on the way the surface is prepared.

On the experimental side, most progresses in our understanding have been driven by the non linear optical techniques, i.e. sum frequency generation (SFG) and second harmonic generation (SHG), which are surface specific techniques and have been widely applied for interface characterization, as de-

tailed in chapter 2. The first application of these techniques on aqueous silica surface was achieved by Eissenthal and coworkers in 1992 [46] by employing the SHG to monitor the silica/water interface under different pH conditions. This pioneering work showed a bimodal titration behavior of the silica surface, i.e. with 2 populations of silanols at the surface, one with an acidic pKa and the other population with a basic pKa. This result has long been considered as the standard behaviour for silica surface deprotonation processes followed by a substantial amount of works. [47, 48, 50–52].

A similar bimodal behavior has been observed by Shen *et al.* [47] on the aqueous (0001) α -quartz surface through SFG spectroscopy and by Shaw *et al.* [48] using evanescent wave cavity ring-down spectroscopy experiments. However, the acidic and basic pKa values of the surface silanols deduced by different works are highly variable depending on the way the surface is prepared, the crystallinity of the surface, etc. Beyond the bimodal behavior, interestingly, Gibbs-Davis *et al.* [49] discovered that depending on the starting pH value in their experimental SHG, three instead of two different populations of silanols with 3 pKa values could be observed. They hence proposed a trimodal behavior for the silica surface deprotonation process.

All the above studies investigated the response of interfacial water in order to indirectly derive the charge state of the silica surface. However, no research has been conducted yet on the oxide side once in contact with liquid water at ambient conditions, due to the difficulties of experimentally probing buried surfaces of insulating solids. Hence, the direct spectroscopic signatures that could resolve the debate around silica surface chemistry were still missing at the time we started our investigation.

On the theory side, thanks to the growing computational power, a substantial amount of computational works [50–53] have attempted to microscopically identify the nature of the acidic and non acidic populations of silanols. By performing DFT-MD simulations on (0001) α -quartz surface, Sulpizi *et al.* [50] ascribed the bimodal acid-base behavior to the surface geometry, in particular to the different O-H orientations of silanols. Using the same method, Pfeiffer *et al.* [51] further confirmed the influence of the surface geometry and of the geometry of solvation of the conjugate base on the acidities of silanols on amorphous aqueous silica surfaces.

All these experimental and theoretical works on the investigation of the bimodal behavior naturally assumed that the deprotonation of silanols lead to the formation of SiO^- , actually without direct proof. In order to fully understand the silica surface chemistry triggered by the deprotonation of silanols under different pH conditions, a technique which is able to directly probe the evolution of the silica surface exposed to increasing pH conditions is mandatory.

To that end, an experimental scheme enabling in situ vibrational spectroscopy of oxide surfaces in liquid water was developed by W.Liu *et al.* and

5.2 In situ experimental SFG spectra of silica surface at the interface with water from Y.R.Shen, W.Liu et al.: a missing band in the intermediate pH

applied on amorphous silica surface in water to record the pH-dependent SFG spectra of silica in the Si-O stretch region (phonon region). This technique allows us, for the first time, to directly probe signatures of Si-O vibrations of the aqueous silica surface, and therefore to unveil the knowledge on the surface physics and chemistry related to silanol deprotonation processes, as will be discussed in section 5.2.

Combined with our unbiased and biased metadynamics DFT-MD simulations as well as theoretical Si-O stretch SFG calculations, an unconventional surface species in addition to SiOH and SiO⁻, was surprisingly discovered, which is formed due to a surface reconstruction triggered by deprotonation processes. This discovery allowed us to reconcile different observations previously reported in the literature and interpret the silica surface chemistry under a new light. All of these results are now presented.

5.2 In situ experimental SFG spectra of silica surface at the interface with water from Y.R.Shen, W.Liu et al.: a missing band in the intermediate pH

All the experimental results presented in this section were obtained from the experiments performed by our collaborator Wei-Tao Liu et. al. and Y. Ron Shen. The details of the experimental setup of the in situ experimental SFG spectroscopy applied on aqueous silica surface at varying pH together with the sample preparation and characterization are presented in the method section of our submitted paper [*“Unveiling structural evolution and reconstruction of oxide surface in water”*] reported at the end of this chapter.

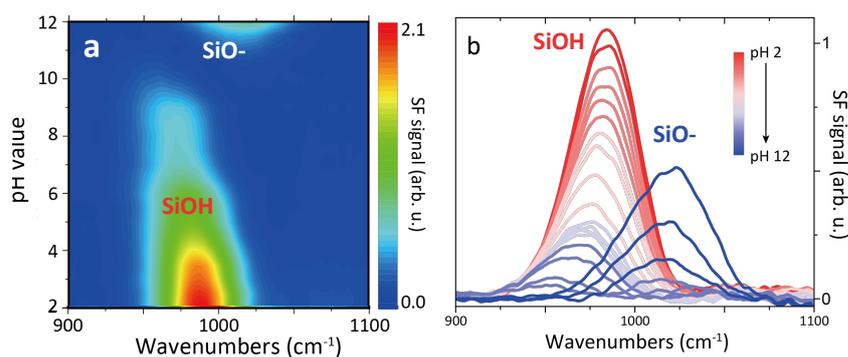


Figure 5.1: Experimental ssp SFG spectra of the amorphous silica-water interface from pH 2 to 12 with 10mM NaCl, from W.Liu et al., Fudan University.

A series of SFG spectra (in the ssp polarization) in the Si-O stretch range for the silica/water interface, as a function of pH with NaCl buffer of 10 mM were acquired, see Fig. 5.1. Two surface species, SiOH and SiO⁻ are normally considered on silica reacting with water. At pH 2 (point of zero charge (PZC) condition for silica), the silica surface is fully hydroxylated and covered by

silanol groups, resulting in a single band at 980 cm^{-1} . This band was already assigned in previous studies [153] to the Si-O(H) stretch. As pH increases, the intensity of the Si-O(H) band decreases, accompanied by a red-shift to 940 cm^{-1} , and disappears at $\sim\text{pH } 10$, where a new band at $\sim 1020\text{ cm}^{-1}$ appears. This new mode matches with the expected signature from the Si-O⁻ stretch of the silanolate groups [169, 170], the commonly assumed end product of the deprotonation of silanols.

Interestingly, this set of SFG spectra of the Si-O surface stretch revealed an unexpected relationship between SiOH and SiO⁻. In all existing studies (both experimental and theoretical) of silica/water interfaces, the two species were believed to convert to each other, following a 1-to-1 correspondence during the deprotonation processes and were believed to coexist on the surface before all SiOHs are deprotonated. However, Fig. 5.1 shows that the SiO⁻ peak only emerges when the SiOH band almost disappears.

This highlights a non 1-to-1 conversion between SiOH and SiO⁻, indicating the possible emergence of a new surface species at the 4–11 pH range. Remarkably, this experimental result questions the basic assumption that underlines any existing model of oxide surface chemistry in water, i.e. the direct interconversion between SiOH to SiO⁻. In order to investigate this exciting new aspect of oxides surface chemistry, theoretical interpretation is needed.

5.3 Silica surface chemistry as a function of pH revealed by DFT-MD biased metadynamics simulations and theoretical Si-O stretch SFG calculations

On the theory side, we aim to uncover the surface chemistry of silica in contact with water at different pHs and unveil possible new surface species that can explain the experimental SFG Si-O stretch spectra discussed in the previous section. To that end, we carry out DFT-MD simulations of aqueous silica interfaces and calculate pH-dependent theoretical SFG spectra in the Si-O stretch frequency range.

However, it is not possible to realistically simulate pH conditions with DFT-MD, as well as atomistic classical MD, due to the high computation cost. For this reason, pH-dependent vibrational spectra (in the full pH range) were never obtained so far. We develop a strategy to overcome this limit and predict pH-dependent SFG spectra of aqueous silica surface from few well-chosen simulations, by combining metadynamics, theoretically SFG spectroscopy, and Henderson-Hasselbalch equations.

Most notably, this approach allows to predict SFG spectra for conditions that can not be simulated with state-of-the-art techniques. It consists of the following steps:

1. We performed metadynamics simulations in DFT-MD framework to investigate possible surface species at different pHs and deduced pKa

value of each silanol. (section 5.3.1)

2. We then used Henderson-Hasselbalch equations to calculate the coverage of each surface species at a given pH, with the knowledge of surface species and the pKa values of silanols from metadynamics. (sections 5.3.2 and 5.3.3)
3. To compute pH-dependent theoretical Si-O stretch SFG spectra, we calculated the group-specific SFG signals arising from each surface species and we constructed the theoretical SFG spectra at a given pH by summing up SFG contributions from each surface species, knowing their coverage as a function of pH. (section 5.3.4)

In the following sub sections, the above steps will be detailed one by one.

5.3.1 DFT-MD metadynamics simulations: discovery of a new surface species

The pKa values of all the surface silanol groups of the surface model and their corresponding end products (conjugate bases) were obtained from the metadynamics simulations by Dr. Flavio Siro Brigiano during his PhD. The details of the metadynamics simulations are presented in his PhD manuscript, NNT: 2020UPASF025. In the following, I will briefly present the essential parts which are the prerequisite for my follow-up work.

The amorphous silica model with a silanol density of 4.5 OH/nm² developed by Ugliengo [163] was adopted for our DFT-MD simulations, which was shown in previous studies to correctly reproduce the structural and spectroscopic properties of experimental amorphous surfaces in contact with both vacuum [27] and water [110].

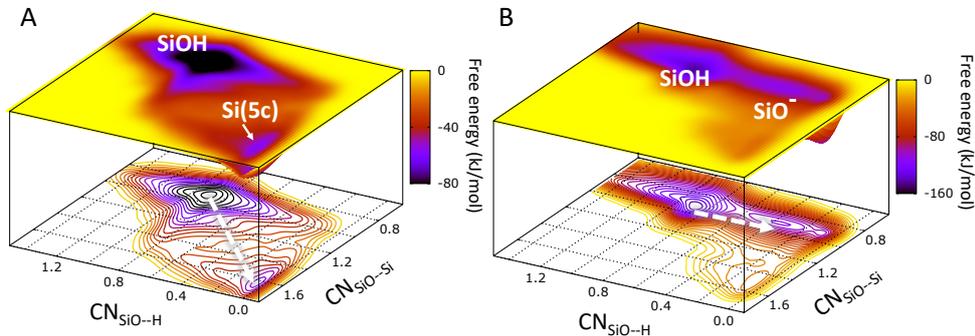


Figure 5.2: Representative free energy landscapes associated with (A) the deprotonation of an acidic SiOH, leading to the unexpected formation of Si(5c)⁻ as conjugate base, see all details in the text. Two free energy minima can be seen, associated with stable surface species of SiOH and Si(5c), respectively, and (B) the deprotonation of a basic SiOH leading to the conventional formation of SiO⁻, with two corresponding free energy minima for SiOH and SiO⁻.

Metadynamics simulations of the deprotonation of each surface silanol (8 in total in the surface model, therefore 8 metadynamics simulations) in contact with water (115 water molecules in total) were performed and the corresponding free energy landscapes (FEL) of deprotonation were obtained. In our metadynamics simulations, we use two reaction coordinates: (1) the coordination number of a chosen silanol oxygen with respect to all hydrogen atoms at the surface (CN_{SiO-H}), and (2) the coordination number of a chosen silanol oxygen with respect to all silicon atoms (CN_{SiO-Si}). The first coordinate aims to investigate the deprotonation of silanols while the second one allows us to trace possible surface reconstructions by monitoring breakage/formation of a Si-O bond. Both coordination numbers were defined by a function of inter-atomic distances introduced in Ref. [171].

Amazingly, two different reaction pathways of the deprotonation of silanol groups were revealed by the free energy landscapes. Two representative FELs were depicted in Fig. 5.2. In one case (Fig 5.2B), we observed two free energy minima corresponding to the standard deprotonation process of SiOH, forming SiO^- as the end product. However, in the other case shown in Fig 5.2A, no minimum for SiO^- was found, showing that SiO^- is not the most stable conjugate base for these silanols. Instead, a minimum for a five-coordinated Si species surprisingly emerged, where the deprotonated oxygen of the given silanol is further covalently bonded to a neighboring Si, leading to a five-coordinated Si species denoted $Si(5c)^-$. This process is sketched in Fig. 5.3.

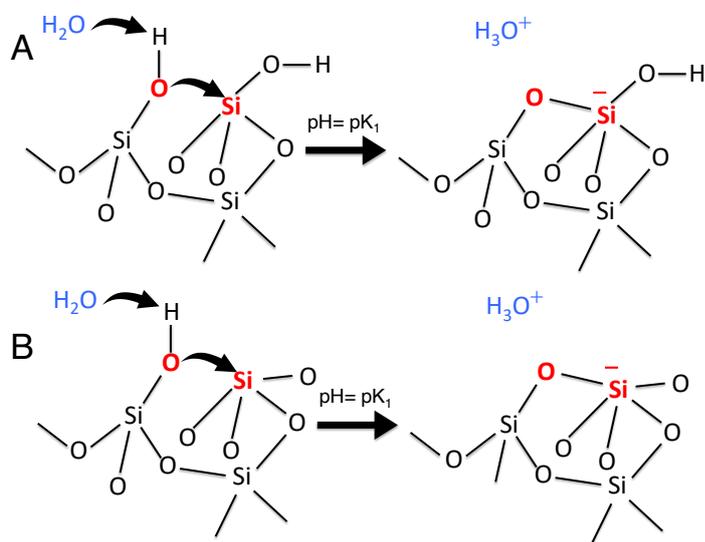


Figure 5.3: (A) Illustrations of the formation of $Si(5c)OH^-$ species and (B) the formation of SiO_5^- species.

Furthermore, two kinds of $Si(5c)^-$ species were observed as end products of the deprotonation of the SiOH groups. In most cases ($\sim 67\%$), the silicon under attack was the host of a silanol group, the resultant five-fold Si species is thus a $Si(5c)OH^-$ species, as shown in Fig. 5.3A. In other cases ($\sim 33\%$),

5.3 Silica surface chemistry as a function of pH revealed by DFT-MD biased metadynamics simulations and theoretical Si-O stretch SFG calculations

the four-coordinated silicon under attack had no hydrogen attached, and the reaction led to a SiO_5^- species (Fig. 5.3B) with the negative charge spread around the host silicon.

The standard free energy differences (ΔG^0) between SiOH and the corresponding conjugate base (SiO^- or $\text{Si}(5\text{c})^-$) for the deprotonation processes of all silanols (one by one) on our model surface were obtained from our metadynamics. The pKa values were hence deduced by $pKa = \Delta G^0/2.303RT$. Two groups of pKa values were obtained, with an acidic pKa of 3.9 ± 0.6 (denoted pK1) for the SiOHs that convert to $\text{Si}(5\text{c})^-$ species after deprotonation, and a basic pKa of 7.7 ± 0.9 (denoted pK2) for the SiOHs with SiO^- being the most stable conjugate base.

Among all the silanols on the model surface, 40% of them have an acidic pKa with $\text{Si}(5\text{c})^-$ as the end product after being deprotonated, and the rest of them are basic, undergoing the standard $\text{SiO} \longleftrightarrow \text{SiO}^-$ conversion.

Further DFT-MD simulations were performed to investigate:

1. The effect of $\text{Si}(5\text{c})^-$ species on the deprotonation process of basic SiOHs.
2. The stability of $\text{Si}(5\text{c})^-$ species at $\text{pH} > \text{pK2}$ (with basic SiOH being deprotonated).

These simulations are detailed in Section S2.1 of the SI of paper [“*Unveiling structural evolution and reconstruction of oxide surface in water*”] reported at the end of this chapter.

In a nutshell, we found that:

1. $\text{Si}(5\text{c})^-$ species are stable between $\text{pH} 4 \sim 9$.
2. They convert to SiO^- once pH rises above ~ 9 due to the accumulation of negative surface charge.
3. The basic SiOH groups still convert to SiO^- in the presence of $\text{Si}(5\text{c})^-$ species at the surface.

Combining all the discoveries discussed above, we hence unraveled the change in surface chemistry at the amorphous aqueous silica surface under different pH conditions.

1. In the pH range between $\text{pH} \sim 2$ (Point of Zero Charge for silica) to pK1, the surface is fully hydroxylated.
2. At $\text{pK1} < \text{pH} < \text{pK2}$, the acidic silanols start to be deprotonated, each deprotonated oxygen attacks and covalently bonds to a neighbouring Si-O(H), forming a $\text{Si}(5\text{c})^-$ species.

3. When pH rises above pK_2 , the basic SiOH groups are deprotonated to SiO^- and $\text{Si}(5\text{c})^-$ species become unstable, split up and form SiO^- .

With this, we are hence able to deduce the surface coverages of each surface species (SiOH , $\text{Si}(5\text{c})^-$ and SiO^-) as a function of pH through Henderson-Hasselbalch equations, in order to mimic different pH conditions and hence study the surface chemical composition and spectroscopy at a given pH, as detailed in the next section.

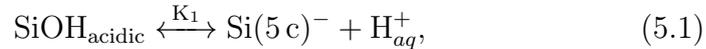
5.3.2 From pKa values to surface coverage vs. pH

Based on the DFT-MD simulation results discussed in the previous section, we hereby determine the variation of the surface coverage of the three surface species as a function of pH : silanols (SiOH , with four-fold Si), silanolates (SiO^-), and five-fold [$\text{Si}(5\text{c})^-$ species], which in most of the cases ($\sim 67\%$ in the model surface of our simulation) carries an OH group in the form of $\text{Si}(5\text{c})\text{OH}^-$. For simplicity, all the five-fold Si species are denoted as $\text{Si}(5\text{c})^-$ in the following.

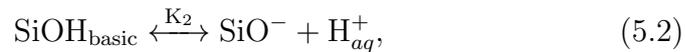
The following parameters/knowledge deduced from the DFT based metadynamics simulations are needed. The initial coverage of the surface is made by silanols with the coverage of $4.5 \text{ OH}/\text{nm}^2$. We found that $\sim 40\%$ (denoted as a) of the initial silanol groups are acidic with a pK_a of 3.9 ± 0.6 (denoted as pK_1) while the rest are basic SiOH groups with a higher pK_a of 7.7 ± 0.9 (denoted as pK_2).

The three surface species transform to each other based on the following mechanism:

- When $pK_1 < \text{pH} < pK_2$, an acidic SiOH is deprotonated and bridged to a nearby Si, which in most of the cases carries an OH group. This thus creates a five-coordinated $\text{Si}(5\text{c})^-$ species at the surface with the following reaction as illustrated in Fig. 5.3:



- When $\text{pH} > pK_2$, basic SiOH is deprotonated to SiO^- , which in the meantime causes the breakage of the Si-O-Si bridge of $\text{Si}(5\text{c})^-$ species due to the accumulation of negative charge at the surface and the formation of SiO^- . We therefore have the 2 following chemical equilibria:



Based on DFT-MD results, we consider that the equilibrium between $\text{Si}(5\text{c})^-$ and SiO^- has the same pH dependence as the one between basic

5.3 Silica surface chemistry as a function of pH revealed by DFT-MD biased metadynamics simulations and theoretical Si-O stretch SFG calculations

SiOH and SiO^- , both with associated pK_2 .

Summing up the two equilibria above, we get:



Using the Henderson-Hasselbalch equation, we now express the surface coverage in SiOH, $\text{Si}(5c)^-$ and SiO^- species (denoted C_{SiOH} , $C_{\text{Si}(5c)^-}$ and C_{SiO^-}) as follows:

$$pH = pK_1 + \log_{10} \frac{C_{\text{Si}(5c)^-}}{C_{\text{SiOH}_{\text{acidic}}}}, \quad (5.5)$$

$$pH = pK_2 + \log_{10} \frac{C_{\text{SiO}^-}}{C_{\text{SiOH}_{\text{basic}}}}, \quad (5.6)$$

$$pH = pK_2 + \log_{10} \frac{C_{\text{SiO}^-}}{C_{\text{Si}(5c)^-}}. \quad (5.7)$$

The coverage of $\text{Si}(5c)^-$ depends on two equilibria:

- At $pK_1 < pH < pK_2$, $\text{SiOH}_{\text{acidic}} \longleftrightarrow \text{Si}(5c)^-$. By doing 10 to the power of both the left and right terms of eq. 5.5, we obtain the coverage in $\text{Si}(5c)^-$:

$$C_{\text{Si}(5c)^-}^{\text{produced}} = C_{\text{SiOH}_{\text{acidic}}}^{\text{initial}} \frac{10^{pH-pK_1}}{1 + 10^{pH-pK_1}} \quad (5.8)$$

where $C_{\text{SiOH}_{\text{acidic}}}^{\text{initial}}$ is the percentage of initial acidic silanol groups.

- At $pH > pK_2$, $\text{Si}(5c)^- \longleftrightarrow \text{SiO}^-$. By doing 10 to the power of both the left and right terms of eq. 5.7, we have:

$$C_{\text{Si}(5c)^-}^{\text{consumed}} = C_{\text{Si}(5c)^-}^{\text{produced}} \frac{10^{pH-pK_2}}{1 + 10^{pH-pK_2}}. \quad (5.9)$$

Combining eq. 5.8 and eq. 5.9, we have:

$$\begin{aligned} C_{\text{Si}(5c)^-}(pH) &= C_{\text{Si}(5c)^-}^{\text{produced}} - C_{\text{Si}(5c)^-}^{\text{consumed}} \\ &= C_{\text{SiOH}_{\text{acidic}}}^{\text{initial}} \frac{10^{pH-pK_1}}{1 + 10^{pH-pK_1}} \left(1 - \frac{10^{pH-pK_2}}{1 + 10^{pH-pK_2}} \right). \end{aligned} \quad (5.10)$$

Now, considering that SiO^- is produced from both basic SiOH groups and $\text{Si}(5c)^-$ groups, by doing 10 to the power of both the left and right terms of eq. 5.6, we have the part from basic SiOH groups, and the part from $\text{Si}(5c)^-$ equals the $C_{\text{Si}(5c)^-}^{\text{consumed}}$.

The coverage in SiO^- species as a function of pH is thus:

$$C_{\text{SiO}^-}(\text{pH}) = \underbrace{\left(1 - C_{\text{SiOH}_{\text{acidic}}}^{\text{initial}}\right)}_{\text{from basic SiOH}} \frac{10^{\text{pH}-\text{p}K_2}}{1 + 10^{\text{pH}-\text{p}K_2}} \quad (5.11)$$

$$+ \underbrace{C_{\text{SiOH}_{\text{acidic}}}^{\text{initial}} \frac{10^{2\text{pH}-\text{p}K_1-\text{p}K_2}}{(1 + 10^{\text{pH}-\text{p}K_1})(1 + 10^{\text{pH}-\text{p}K_2})}}_{\text{from Si(5c)}^-}. \quad (5.12)$$

Finally, the coverage in SiOH species is obtained by:

$$C_{\text{SiOH}}(\text{pH}) = 1 - C_{\text{SiO}^-}(\text{pH}) - C_{\text{Si(5c)}^-}(\text{pH}). \quad (5.13)$$

Equations 5.10, 5.12 and 5.13 hence provide all knowledge on the coverage in the 3 possible species at the silica surface as a function of pH.

According to our metadynamics simulations, the $\text{p}K_a$ values of acidic and basic silanol groups are in two broad ranges instead of being two single values, i.e. $\text{p}K_1 = 3.9 \pm 0.6$ and $\text{p}K_2 = 7.7 \pm 0.9$. This shows the heterogeneity of the silica surface in terms of the deprotonation of silanol groups. It is important to take the heterogeneity into account. We hence adopt a gaussian distributions of $\text{p}K_1$ and $\text{p}K_2$ values with central values being the average values of $\text{p}K_1$ and $\text{p}K_2$ deduced from the metadynamics simulation: 3.9 and 7.7, and take the corresponding standard deviations as the Gaussian widths (0.6 for $\text{p}K_1$ and 0.9 for $\text{p}K_2$) to calculate the surface coverage using equations 5.10, 5.12 and 5.13.

Furthermore the $\text{p}K_a$ values calculated from the DFT-MD metadynamics are with respect to an effective surface pH (H^+ concentration at the interface, denoted as surface pH hereafter for simplicity) as the deprotonation reaction happens at the surface, whereas the experimentally measured $\text{p}K_a$ values are based on effective bulk pH (H^+ concentration in the solution, denoted as bulk pH for simplicity). In order to calculate the surface coverage as a function of bulk pH and directly compare to experimental results, we hence transform the $\text{p}K_a$ values deduced from metadynamics with respect to surface pH to the values with respect to bulk pH, as detailed in the next section.

5.3.3 Correction of the surface coverage with respect to bulk pH

At the silica/water interface, the surface pH differs from the bulk pH due to the attraction between negative surface charges and protons and the repulsion between negative surface charges and OH^- . The difference between the surface pH and bulk pH is thus likely to become considerable with increasing

surface deprotonation. Therefore, in order to quantitatively compare the pH dependent surface chemistry uncovered by our simulations to the experiment, it is necessary to transform the calculated pKa values obtained with respect to surface pH to the ones obtained with respect to bulk pH.

The correspondence between surface pH and bulk pH values depends on the surface potential (see eq. 5.17).

To calculate the surface potential, we adopt the modified Poisson-Boltzmann equation [141, 142]:

$$\nabla^2\phi = \frac{en_0}{\varepsilon\varepsilon_0} \frac{2 \sinh\left(\frac{e\phi}{k_B T}\right)}{1 + 2\nu \sinh^2\left(\frac{e\phi}{2k_B T}\right)}, \quad (5.14)$$

where ϕ is the electric potential as a function of the distance in the z-direction with respect to the surface, n_0 , k_B , T , e , ε , and ε_0 are the ion concentration in the bulk (both 0.5 M, 10 mM of NaCl are considered, where 0.5 M is consistent with the previous SHG measurements from Ref. [46] and 10 mM corresponds to one of the experimental sets of SFG measurement performed by our collaborators Prof. Wei-Tao Liu *et al.* and Prof. Y.Ron Shen, as shown in sec. 5.2), the Boltzmann constant, the temperature in K, the elementary charge, the relative dielectric constant of liquid water and the dielectric constant of vacuum.

While the traditional Poisson-Boltzmann model treats ions as point charges, the modified Poisson-Boltzmann method considers the finite ion size effect by introducing a parameter ν , expressed as $\nu = 2a^3n_0$, where a is the finite ion size including the hydration shell. We adopt $a = 0.7$ nm, which is assumed to be the Bjerrum length in water at room temperature [172, 173].

The electric potential at the surface ϕ_0 is then calculated by the analytical solution of Eq. 5.14, which takes the form of:

$$\phi_0 = \phi(0) = -\frac{2k_B T}{e} \sinh^{-1} \left\{ \sqrt{\frac{1}{2\nu} \left[\exp\left(\frac{\sigma^2\nu}{4n_0\varepsilon\varepsilon_0 k_B T}\right) - 1 \right]} \right\}. \quad (5.15)$$

where σ is the surface charge density.

With our calculated coverages of surface species as a function of pH, the surface charge density is $\sigma(pH) = (C_{Si(5c)^-} + C_{SiO^-}) \rho_{\text{silanols}} e$, where $C_{Si(5c)^-}$ and C_{SiO^-} are coverages of Si(5c)⁻ and silanates respectively (determined in the previous section, see eq. 5.10 and eq. 5.12), and ρ_{silanols} is the density of initial silanol groups (~ 4.5 SiOH /nm²).

The electric potential $\phi(z)$ can be deduced by the following equation which expresses the electric field as a function of the vertical distance (z) from the surface:

$$\frac{d\phi(z)}{dz} = \frac{2en_0\lambda_D}{\varepsilon\varepsilon_0} \sqrt{\frac{2}{\nu} \ln \left[1 + 2\nu \sinh^2\left(\frac{e\phi}{2k_B T}\right) \right]}, \quad (5.16)$$

where $\lambda_D = \sqrt{\frac{\epsilon\epsilon_0 k_B T}{2e^2 n_0}}$ is the Debye length.

We then calculate the concentration of H^+ in the bulk from the surface H^+ concentration (at $z_0 = 2.5 \text{ \AA}$ from the silica surface, corresponding to the topmost water oxygen layer in contact with the surface in the DFT-MD simulations) by the equation given by the modified Poisson-Boltzmann theory [141, 142]:

$$n_{Bulk}(H^+) = n_{Surface}(H^+) \frac{\left(1 + 2v \sinh^2\left(\frac{e\phi(z_0)}{2k_B T}\right)\right)}{\exp\left(-\frac{e\phi(z_0)}{k_B T}\right)}. \quad (5.17)$$

From Eq. 5.17, we establish the relation between bulk pH and surface pH for the two NaCl concentrations used in SHG and SFG experiments, respectively. This is shown in Fig. 5.4.

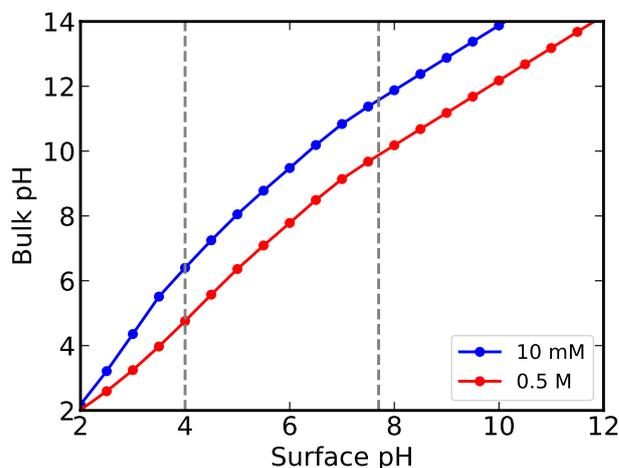


Figure 5.4: Calculated bulk pH versus surface pH values from eq. 5.17. Two different concentrations in electrolytes NaCl are used: 0.5 M to match the SHG results from ref. [46] and 10 mM to match the SFG experiments from Prof. Wei-Tao Liu.

With this, we can thus translate the theoretical pKa values (calculated with respect to surface pH) to bulk pH units as measured in experiments at NaCl concentrations of both 0.5 M for SHG (red) and 10 mM for SFG, as illustrated in Fig. 5.4.

As expected, the discrepancy between surface pH and bulk pH at 0.5 M is smaller than that at 10 mM, as higher ion concentration in the solution screens the surface potential created by the negative surface charges, therefore diminishes the attraction of protons to the surface and diminishes the repulsion of OH^- away from the surface.

While no explicit ions were included in our metadynamics DFT-MD simulations for the calculation of pKa values of SiOH groups, the effect of ion concentrations in the solution can be taken into account when translating our pKa values with respect to surface pH to bulk pH.

We thus end up with three different sets of pKa values, referring to:

- surface pH: $pK1 = 3.9 \pm 0.6$ and $pK2 = 7.7 \pm 0.9$ for respectively acidic SiOH and basic SiOH.
- bulk pH at 10 mM (to compare with experimental SFG spectra from Pr. Liu with 10 mM NaCl): $pK1 = 6.7 \pm 0.6$ and $pK2 = 11.6 \pm 0.9$.
- bulk pH at 0,5 M (to compare with SHG results from ref. [46] at 0.5 M: $pK1 = 4.7 \pm 0.6$ and $pK2 = 9.8 \pm 0.9$.

Taking these bulk pKa values and adopting the method discussed in the previous section, we deduce the coverage in surface species as a function of bulk pH at 10 mM and at 0.5 M NaCl respectively. The representative coverage plot at 10 mM NaCl is presented in Fig. 5.5.

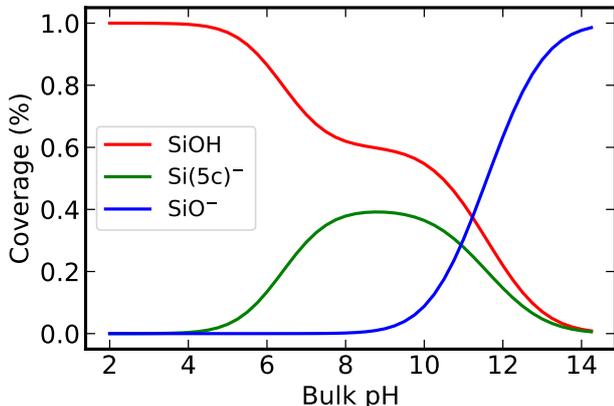


Figure 5.5: Coverage of the silica surface species as a function of bulk pH at 10 mM NaCl. (concentration used in the SFG experiments of Prof. Wei-Tao Liu *et al.* and Prof. Y. Ron Shen).

In the coverage plot, we observe that at pH 2, only SiOH species exists on the surface of silica. As pH increases, acidic SiOH groups are deprotonated and transformed to Si(5c)⁻ species. As pH exceeds 10, basic SiOH groups are deprotonated and converted to SiO⁻, while simultaneously the Si(5c)⁻ species become unstable, split up and transform into SiO⁻.

5.3.4 Theoretical calculation of SFG spectra as a function of bulk pH

Knowing the surface composition at a given bulk pH, we are now able to construct theoretical SFG spectra of the silica surface as a function of pH without performing the actual DFT-MD simulations at various pH values. The theoretical spectra will be compared to the experimental SFG spectra in the Si-O stretch range presented in sec. 5.2, and used to interpret these experimental spectra.

As said before, we do not have simulations as a function of pH. Instead, we have unbiased DFT-MD simulations (~ 20 ps for each) in which each of the three possible species appearing at the silica surface are present. With these simulations, we can calculate the SFG signature of each surface species population. Knowing the concentration (coverage) of each species at a given bulk pH (as deduced in the previous section using equations 5.10, 5.12 and 5.13, see fig. 5.5), we will then be able to obtain the total SFG spectra in the Si-O stretch region by summing up contributions arising from each surface species.

The individual Si-O SFG spectra (in the 700 to 1200 cm^{-1} range) of each surface species are then used as building bricks to construct the total spectra as a function of bulk pH. The theoretical SFG spectrum of the silica/water interface in the Si-O stretching region at a given pH is obtained by the weighted sum of the individual $\chi^{(2)}(\omega)$ signals arising from the Si-OX groups of the three possible species at the silica surface, i.e. Si-OH, Si-O⁻, and Si-OSi:

$$\chi^{(2)}(\omega, pH) = \sum_j^{\text{SiOX species}} C_j(pH) \bar{\chi}_j^{(2)}(\omega), \quad (5.18)$$

where $C_j(pH)$ is the coverage of species j at a given (bulk) pH (see previous section) and $\bar{\chi}_j^{(2)}(\omega)$ is the theoretical SFG signal of the corresponding Si-OX species.

Species-specific SFG spectrum of a given silica surface species

As detailed in Chapter 2, Section 2.6, to calculate the $\bar{\chi}_{j,xxz}^{(2)}(\omega)$ signal (experimentally probed in ssp polarization) of each Si-OX species, we here adopt the same methodology as previously developed for the determination of theoretical $\chi^{(2)}(\omega)$ spectra of both water and SiOH in the O-H stretch range presented in ref. [36] and previously in section 2.3. This methodology focuses on the self-correlation terms only, and we discuss the main principles of this approach in the following.

The species-specific DFT-MD SFG spectrum is obtained by:

$$\bar{\chi}_{j,xxz}^{(2)}(\omega) = \frac{i}{k_B T \omega} \int_0^\infty dt e^{-i\omega t} \langle \dot{\alpha}_{j,xx}(t) \dot{\mu}_{j,z}(0) \rangle, \quad (5.19)$$

where $\mu_{j,z}(0)$ and $\alpha_{j,xx}(t)$ refer to the z and xx components of the dipole and polarizability of a given Si-O group in the laboratory frame and $\dot{\mu}_{j,z}(0)$ and $\dot{\alpha}_{j,xx}(t)$ are their time derivatives. Using basic geometry considerations, we can express the dipole moment of the Si-O bond in the laboratory frame from the calculated one in the molecular frame by:

$$\dot{\mu}_{j,z}(0) \approx \sum_i^{x',y',z'} D_{j,z,i}(0) \dot{\mu}_{j,i}(0) \approx D_{j,z,SiO}(0) \dot{\mu}_{j,SiO}(0), \quad (5.20)$$

where x' , y' , z' are any set of three directions in the molecular frame, and D is a cosine matrix projecting the molecular frame onto the laboratory frame.

5.3 Silica surface chemistry as a function of pH revealed by DFT-MD biased metadynamics simulations and theoretical Si-O stretch SFG calculations

The second equality is written assuming that the dipole component along the Si-O bond direction (i.e. z') is much larger than that in the perpendicular directions (i.e. x' , y').

By introducing a change of variable in the derivative $\dot{\mu}_{j,i}(0)$, we hence obtain:

$$\dot{\mu}_{j,z}(0) \approx D_{j,z, SiO} \frac{d\mu_{j, SiO}}{dr_{j, SiO}} v_{j, SiO}(0), \quad (5.21)$$

where $r_{j, SiO}$ is the Si-O bond coordinate and $v_{j, SiO}(0) = \frac{dr_{j, SiO}}{dt}$. We assume that the orientation of the Si-O bond does not change over time for a given Si-OH or Si-O⁻ group, therefore the dependence of $D_{j,z, SiO}$ on time can be neglected.

With the same methodology for the polarizability, we have:

$$\dot{\alpha}_{j,xx}(t) = D_{j,x, SiO}^2 \frac{d\alpha_{j, SiOSiO}}{dr_{j, SiO}} v_{j, SiO}(t), \quad (5.22)$$

Substituting Eq. 5.21 and Eq. 5.22 into Eq. 5.19, we hence obtain:

$$\begin{aligned} \bar{\chi}_{j,xxz}^{(2)}(\omega) &= D_{j,x, SiO}^2 D_{j,z, SiO} \frac{d\alpha_{j, SiOSiO}}{dr_{j, SiO}} \frac{d\mu_{j, SiO}}{dr_{j, SiO}} \\ &\quad \frac{i}{k_B T \omega} \int_0^\infty dt e^{-i\omega t} \langle v_{j, SiO}(t) v_{j, SiO}(0) \rangle, \end{aligned} \quad (5.23)$$

where $D_{j,x, SiO}$, $D_{j,z, SiO}$, and $v_{j, SiO}(t)$ are obtained from the simulations, while $\frac{d\alpha_{j, SiOSiO}}{dr_{j, SiO}}$ and $\frac{d\mu_{j, SiO}}{dr_{j, SiO}}$ are usually parameterized based on reference ab initio simulations [36, 43].

Here, we set them to unity since we are more interested in the influence on the SFG intensity of a change in species populations (as it happens when changing the pH) than in the change in the SFG activity of the species. For a more accurate description of the SFG activity, $\frac{d\alpha_{j, SiOSiO}}{dr_{j, SiO}}$ and $\frac{d\mu_{j, SiO}}{dr_{j, SiO}}$ must be rigorously evaluated for all x' , y' , z' components. This has not been done in this work.

Construction of the SFG spectrum of the silica surface as a function of pH

We hence calculate the SFG spectrum of the silica surface as a function of pH by summing up the SFG contributions from each species population weighted by the corresponding surface coverage (see eq. 5.18):

$$\begin{aligned} \chi_{\text{total}}^{(2)}(\omega, pH) &= C_{SiOH}(pH) \bar{\chi}_{SiOH}^{(2)}(\omega) + C_{SiO^-}(pH) \bar{\chi}_{SiO^-}^{(2)}(\omega) \\ &\quad + C_{SiOSi}(pH) \bar{\chi}_{SiOSi}^{(2)}(\omega), \end{aligned} \quad (5.24)$$

where Si-O-Si bonds are centrosymmetric thus SFG-inactive for both four-fold (as previously inferred in Ref. [153]) and five-fold species. The centrosymmetry of the five-fold Si-O-Si bonds is further demonstrated by the

fact that $D_{j,z,SiO}$ (the cosine matrix projecting molecular frame onto experimental frame) equals -0.21 ± 0.31 (averaged over all $Si(5c)^-$ population), i.e. approximately 0 within error bar, hence:

$$\chi_{\text{total}}^{(2)}(\omega, pH) = C_{SiOH}(pH)\bar{\chi}_{SiOH}^{(2)}(\omega) + C_{SiO^-}(pH)\bar{\chi}_{SiO^-}^{(2)}(\omega). \quad (5.25)$$

As we will show in the following, the formation of $Si(5c)^-$ species has relevant consequences on the measured SFG signals:

- Formation of a $Si(5c)^-$ converts one SiOH into a Si-O-Si, therefore decreasing the SFG intensity measured by the SiOH band without any other rising contribution.
- The SiOH groups hosted by a $Si(5c)^-$ (in the case of $Si(5c)OH^-$) have a different chemical environment than standard (4-fold coordinated SiOH groups), which could result in a frequency shift.

Indeed, our simulation revealed that when pH is in the range between pK1 and pK2, the Si that carries an OH group can be either 4-coordinated or 5-coordinated. Due to the negative charge of $Si(5c)OH^-$, we observe from the simulations a red-shifted Si-O(H) signal ($\bar{\chi}_{Si(5c)OH^-}^{(2)}(\omega)$) with respect to the normal tetrahedral SiOH (as shown in Fig. 5.6, green for $Si(5c)OH^-$ and red for normal tetrahedral SiOH). Therefore, we deconvolve the $\bar{\chi}_{SiOH}^{(2)}(\omega)$ in eq. 5.25 into a 4-fold $\bar{\chi}_{Si(4c)OH}^{(2)}(\omega)$ and a 5-fold $\bar{\chi}_{Si(5c)OH^-}^{(2)}(\omega)$, and their corresponding coverages are represented by $[C_{SiOH}(pH) - C_{Si(5c)OH^-}(pH)]$ and $C_{Si(5c)OH^-}(pH)$.

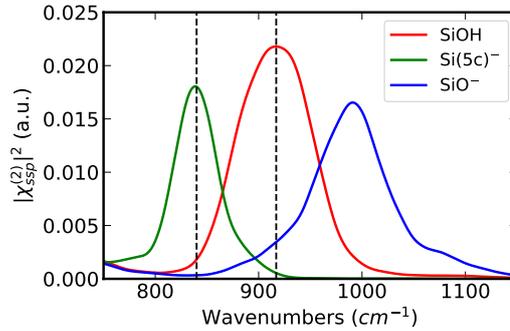


Figure 5.6: Average theoretical $|\chi_{ssp}^{(2)}(\omega)|^2$ spectra of individual 4-fold Si-O(H), Si-O(H) belonging to $Si(5c)^-$ and SiO^- species obtained from DFT-MD simulations in the 750 – 1150 cm^{-1} domain of the Si-O stretching.

We then calculate the theoretical pH-dependent SFG spectra (shown in Fig. 5.7) in ssp polarization with respect to bulk pH at 10 mM by:

$$\begin{aligned} \chi_{\text{total}}^{(2)}(\omega, pH) = & (C_{SiOH}(pH) - C_{Si(5c)OH^-}(pH)) \bar{\chi}_{Si(4c)OH}^{(2)}(\omega) \\ & + C_{Si(5c)OH^-}(pH) \bar{\chi}_{Si(5c)OH^-}^{(2)}(\omega) + C_{SiO^-}(pH) \bar{\chi}_{SiO^-}^{(2)}(\omega). \end{aligned} \quad (5.26)$$

5.3 Silica surface chemistry as a function of pH revealed by DFT-MD biased metadynamics simulations and theoretical Si-O stretch SFG calculations

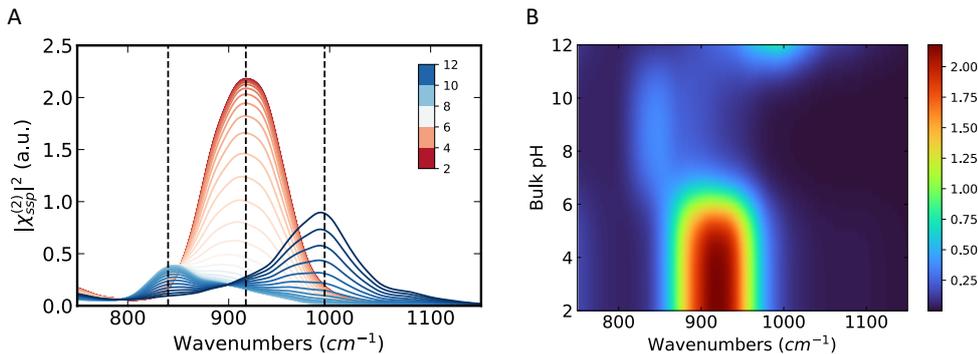


Figure 5.7: Theoretical SFG $|\chi_{ssp}^{(2)}(\omega)|^2$ spectra as a function of bulk pH (10 mM), the dashed black lines represent the position of the average species specific SFG bands of SiOH belonging to $Si(5c)^-$ and SiOH (normal tetrahedrally coordinated Si) calculated from DFT-MD (see also Fig. 5.6), respectively, at 840 cm^{-1} and 920 cm^{-1} .

We now comment the calculated spectra of Fig. 5.7 as a function of increasing pH.

- At pH 2, there is a single band at 920 cm^{-1} corresponding to the signal from SiOH groups only.
- As pH increases above 4, the SiOH band diminishes in intensity while red shifts to 840 cm^{-1} due to the deprotonation of acidic SiOH groups with resultant $Si(5c)OH^-$ species whose Si-O stretch frequency is red-shifted because it is softened by the negative charge.
- When pH exceeds 10, a band at 1000 cm^{-1} emerges which corresponds to the appearance of the SiO^- species, corresponding to the deprotonation of basic SiOH groups and breakage of $Si(5c)^-$ species that both form SiO^- .

Overall, our DFT-MD calculated SFG spectra of aqueous silica surface as a function of pH (10 mM NaCl) are in good agreement with experimental SFG spectra presented in Section 5.2. The discrepancy in frequency between calculated SFG spectra and experimental SFG spectra is due to the DFT functional, as shown ref. [124].

To summarize, in this work, we developed a new method of calculating pH-dependent theoretical SFG spectra in the Si-O stretch region without setting the pH by tuning the proton concentration in the DFT-MD simulations. To achieve our goal, we first calculated pKa values of the surface silanol groups through metadynamics simulations and discovered two populations of silanols with acidic and basic pKa values, respectively. We surprisingly uncovered a new surface species, $Si(5c)^-$, that only exists in the intermediate pH range (pH 4 – 11).

Based on the pKa values, we deduced the surface coverage of each species at a given pH, through Henderson-Hasselbalch equations. After calculating the species-specific SFG signals arising from each surface species (SiOH,

Si(5c)⁻ and SiO⁻), we constructed the theoretical SFG spectrum of the silica/water interface in the Si-O stretch domain at a given pH by summing up SFG signals from each surface species, knowing the coverages of surface species at a given pH.

The combination between metadynamics and both theoretical and experimental SFG calculation revealed a new species, Si(5c), at the silica surface whose formation is triggered by the deprotonation reaction of silanol groups with an acidic pKa value at intermediate pH (pH 4–11). Both the experimental and simulated SFG spectra demonstrated the formation and existence of Si(5c)⁻ species. In particular, we demonstrated that (1) the non-conservation of SiOH and SiO⁻ intensities in the SFG as a function of pH is due to the formation of Si(5c)⁻ by explicitly calculating and comparing theoretical SFG to experimental SFG, and (2) we demonstrated that the red shift of the SiOH band, initially observed in the experimental SFG and confirmed in the theoretical SFG, is due to the formation of the Si(5c)⁻ species.

5.4 Consistency between the discovery of Si(5c)⁻ species and SHG titration experiments

In the work of ref. [46], as well as all the other works dedicated to investigating the bimodal behavior of deprotonation of the surface silanols [46–48, 50–52], the authors all naturally assumed that the deprotonation of the silanols only leads to the formation of SiO⁻ without direct proof. In this joint experimental and theoretical work, we propose a different picture of the silica surface chemistry as a function of pH, with the discovery of the newly discovered Si(5c)⁻ species as the end groups of the acidic SiOH groups after deprotonation.

We now investigate if our new picture can explain the bimodal titration behavior of the silica/water interface initially observed by SHG [46].

The pH dependence in SHG is in response to the variation of SiO₂ surface charge density. Knowing the surface charge density from the coverages in Si(5c)⁻ and SiO⁻ at given pH, we can thus calculate the $\chi_{SHG}^{(2)}$ as follows.

According to Ref. [141], the effective surface nonlinear polarization at 2ω (with two incident laser beams at $\omega = 600$ nm in Ref. [46]’s setup) is expressed by $\chi_{eff}^{(2)}$:

$$\chi_{eff}^{(2)} = \chi_{BIL}^{(2)} + \chi_{DL}^{(2)} \quad (5.27)$$

$$\chi_{DL}^{(2)} = \chi_{Bulk}^{(3)} \Psi, \quad (5.28)$$

$$\Psi = \int E_0(z') e^{i\Delta k_z z'} dz', \quad (5.29)$$

where Δk_z is the phase mismatch of the reflected SHG, $\chi_{BIL}^{(2)}$ is the second-order nonlinear susceptibility of the BIL water (i.e., the topmost binding interfacial layer where centrosymmetry is broken by direct interaction with the surface, see Ref. [33] for its structural definition and characterization), Ψ is the potential difference across the DL region, and $\chi_{Bulk}^{(3)}$ is the third-order nonlinear response of water molecules to optical and static fields, which was shown to be constant for both SFG [33, 174] and SHG [141].

For SHG, $\chi_{Bulk}^{(3)} = (9.56 \pm 1.87) \times 10^{-22} \text{ m}^2/\text{V}^2$, as determined in Ref. [141].

Since in section 5.3.3 we have calculated σ , ϕ_0 , and $\phi(z)$ as a function of (bulk) pH at 0.5 M ion concentration (same ionic strength as in the SHG experiment in ref. [46]), we can derive the electrostatic field $E_0(z')$ from $\phi(z)$, with $E_0(z')\hat{z} = -\nabla\phi(z)$, where \hat{z} is the unit vector pointing along the z-direction to the surface), and then deduce the complex Ψ from Eq. 5.29 with $\Delta k_z^{-1} = 25 \text{ nm}$ [175]. We thus can calculate the $\chi_{DL}^{(2)}$ as a function of pH by Eq. 5.28.

$\chi_{eff}^{(2)}$ is a combination of $\chi_{BIL}^{(2)}$ and $\chi_{DL}^{(2)}$ contributions (Eq. 5.27). Now that we have the theoretical $\chi_{DL}^{(2)}$ by eq. 5.28, we need to deduce the $\chi_{BIL}^{(2)}$ term in order to compare with the SHG results ($\chi_{eff}^{(2)}$) reported in Ref. [46].

It should be noted that the change in $\chi_{BIL}^{(2)}$ as a function of bulk pH was shown in previous studies to be much smaller than the change in the $\chi_{DL}^{(2)}$ term, as expected from Eq. 5.28 since the $\chi_{DL}^{(2)}$ intensity is proportional to the surface field and therefore has much more pronounced pH dependence than $\chi_{BIL}^{(2)}$ (see e.g. Ref. [110]). Following the data analyses performed in Ref. [46], we hence assume in the following that the pH dependence of $\chi_{BIL}^{(2)}$ is negligible with respect to the pH dependence of $\chi_{DL}^{(2)}$. At $\text{pH} \sim 2$ (PZC) the surface charge is zero, hence $\chi_{DL}^{(2)} = 0$ and $\chi_{eff}^{(2)} = \chi_{BIL}^{(2)}$. We thus take the $\chi_{eff}^{(2)}$ value at $\text{pH}=2$ from the SHG results reported in Ref. [46] as a constant $\chi_{BIL}^{(2)}$ from pH 2 to 12.

By adding this $\chi_{BIL}^{(2)}$ (assumed to be independent of pH) term to our calculated $\chi_{DL}^{(2)}$ at each pH, we hence obtain our theoretical $\chi_{eff}^{(2)}$ as a function of pH. This is reported in Fig. 5.8 in blue. This theoretical SHG results are in good agreement with that reported in Ref. [46], in black in Fig. 5.8). Note that the $\chi_{eff}^{(2)}$ extracted from Ref. [46] are in arbitrary units, and therefore we scale them to the theoretical spectra to match the SHG intensity at the highest pH.

The good agreement between the $\chi_{eff}^{(2)}$ results obtained from the DFT-MD surface charge values (blue point line) and the experimental results extracted from Ref. [46] (black line) shown in Fig. 5.8 proves that our model that includes the newly discovered $\text{Si}(5\text{c})^-$ species is consistent with the historical SHG measurements of Ref. [46]. This measurement historically gave birth

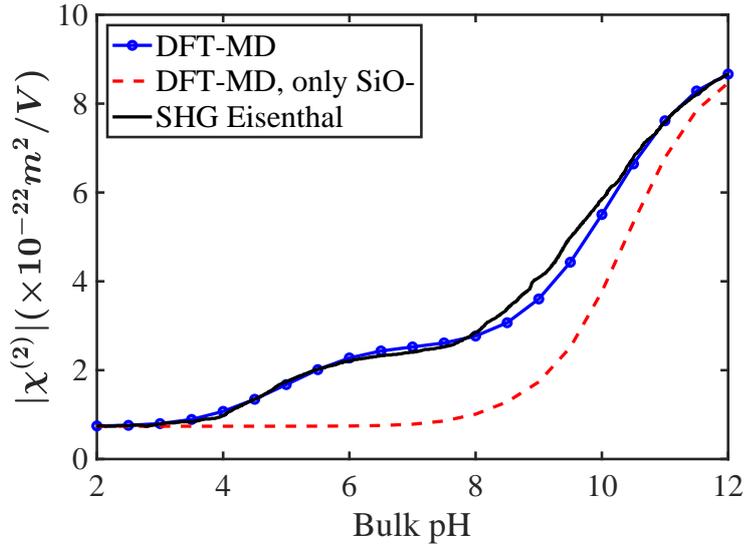


Figure 5.8: Comparison of the theoretical SHG $|\chi_{DL}^{(2)}|$ obtained from the DFT-MD surface charge values (see text) and the experimentally derived $|\chi_{DL}^{(2)}|$ from Ref. [46] as a function of bulk pH. The red dashed line reports the theoretical $|\chi_{DL}^{(2)}|$ spectrum obtained by considering only the surface charge carried by SiO^- species for comparison. The mismatch between this curve and the experimental one in the $4 < \text{pH} < 11$ range is due to the formation of negatively charged $\text{Si}(5\text{c})^-$ species at the surface.

to the commonly assumed two-sites/two-states model, assuming only two surface species, i.e. SiOH and SiO^- .

However, we hereby provide a new interpretation: at pH ranging from 4.7 to 9.8, the negative charge produced at the surface comes from $\text{Si}(5\text{c})^-$ species instead of SiO^- , while the SiO^- solely contributes to the negative charge at pH greater than 9.8.

To further prove the necessity to account for $\text{Si}(5\text{c})^-$ species to interpret the SHG experiment data of Eisenthal *et al.* [46], we plot the $\chi_{eff}^{(2)}$ values calculated by only considering the surface charge coming from SiO^- species, shown in red dashed curve in Fig. 5.8. There is a clear discrepancy between these $\chi_{eff}^{(2)}$ values and the SHG measurements in Ref. [46] (black curve) in the pH range from 4 to 11 because of the negative charged $\text{Si}(5\text{c})^-$ species not being taken into account. Therefore, considering the negative surface charge coming from both $\text{Si}(5\text{c})^-$ and SiO^- allows to quantitatively reproduce and explain the SHG results in Ref. [46].

5.5 Requisites for the $\text{Si}(5\text{c})^-$ species formation at the silica surface

The discovery of the third silica surface species, $\text{Si}(5\text{c})$, through metadynamics simulations explained why the one-to-one conversion between SiOH and

SiO^- was not satisfied as revealed by the pH-dependent experimental SFG spectra recorded by Prof. Wei-Tao Liu's group and Y. Ron Shen. The good agreement between pH-dependent theoretical SFG spectra calculated with DFT-MD parameters and the experimental SFG in the Si-O domain confirmed the novel picture that three surface silica species are inter-converting to each other under different pH conditions on the aqueous silica surface, which also satisfactorily explained the multimodal titration behavior of the silica/water interface observed by SHG experiments [46].

The remaining question is, what leads to the formation of $\text{Si}(5\text{c})^-$ species? To answer this, we investigated the local structure around surface silanols, as presented in the following.

5.5.1 Interaction between acidic silanols and water?

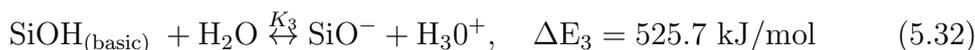
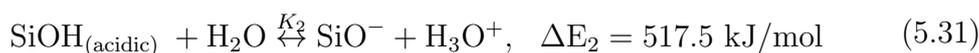
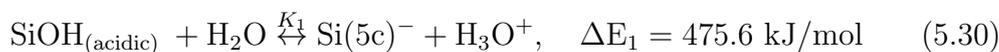
With the conjecture that the formation of $\text{Si}(5\text{c})^-$ species might be due to interaction of silanols with water, we explored the water density exposed to acidic silanols and the type of H-bonds formed between each acidic silanol and water (donor or acceptor of HBs to water). No correlation was found between the $\text{Si}(5\text{c})^-$ formation and the interactions between acidic silanols and water (see details in Section S4 of the SI of our paper [*Unveiling structural evolution and reconstruction of oxide surface in water*]).

To further confirm that water is not to be of importance for the formation of $\text{Si}(5\text{c})^-$ species on the silica surface, we have performed supplementary metadynamics simulations by deprotonating the acidic silanols one by one in absence of water at room temperature. The same $\text{Si}(5\text{c})^-$ species was found on the silica surface in vacuum, proving that the interaction of silanols with water is not essential for the $\text{Si}(5\text{c})^-$ formation.

We then investigated up to which extent the trends observed at room temperature in presence or absence of water could be qualitatively preserved in vacuum at 0 K. To compare the stability of $\text{Si}(5\text{c})^-$ and SiO^- deprotonated states of silanols at 0 K in the absence of water, we have performed a set of geometry optimization calculations (using exactly the same electronic set as for the DFT-MD simulations, i.e. same functional, pseudopotential, basis sets, etc.) for the silica surface in vacuum (the simulation box contains one water or one hydronium ion to simulate the deprotonation reaction and to keep the simulation box neutral).

Four configurations have been optimized: (1) fully hydroxylated silica surface and one water molecule, (2) silica surface with one $\text{Si}(5\text{c})^-$ species formed by deprotonating an acidic silanol group and one hydronium ion, (3) silica surface with one SiO^- species formed by deprotonating an acidic silanol group and one hydronium ion and (4) silica surface with one SiO^- species formed by the deprotonation of a basic silanol group and one hydronium ion. The water molecule or the hydronium ion in the simulation boxes is kept far away from the silica surface and constrained there during the geometry optimization calculations.

At $T = 0$ K, there is no entropic contribution to the free energy of the deprotonation reaction. By making the difference between the total energy of the different deprotonated states and that of the fully hydroxylated state of the silica surface, we obtain the energy difference (ΔE) for the three deprotonation reactions as reported below:



Even at 0K where temperature and entropic effects do not contribute, we still observe that we have two populations of SiOH, a more acidic and a more basic population, with the acidic population having $\text{Si}(5\text{c})^-$ as the most stable conjugate base (as $\Delta E_1 < \Delta E_2$). Importantly, the SiOHs that were acidic at the aqueous surface are still acidic at the vacuum interface, same outcome for the basic silanols. However, the energy cost (see ΔE values above) for the deprotonation reactions at 0 K is one order of magnitude higher than that obtained from metadynamics at room temperature: 22.4 kJ/mol for deprotonation of acidic SiOH into $\text{Si}(5\text{c})^-$ (Fig. 5.2, section 5.3.1), instead of 475.6 kJ/mol reported above in eq. 5.30.

Even if a direct quantitative comparison between the results of geometry optimization and metadynamics should be avoided, this finding still suggests that the SiOH deprotonation is an entropically favored process in liquid water at room temperature. The ΔE for an acidic silanol to be deprotonated and form $\text{Si}(5\text{c})^-$ is lower by 41.9 kJ/mol than that required to form a SiO^- in vacuum. This indicates that for the acidic silanols, $\text{Si}(5\text{c})^-$ is enthalpically a more stable deprotonation product than SiO^- , qualitatively in line with the results obtained from metadynamics performed at room temperature.

Despite being crude, these geometry optimization calculations performed at 0 K in vacuum confirmed that the requisite for $\text{Si}(5\text{c})^-$ species to be the most stable conjugate base of acidic silanols is not the interaction with water.

5.5.2 Specific local morphology in proximity of acidic silanols?

We then questioned whether the acidic silanols had specific surface local morphology which favored the attack of deprotonated acidic silanol to covalently bond to a Si in proximity. The radial distribution functions of each silanol oxygen with respect to all the Si atoms were plotted (Fig. 5.9A). We found that the distances between acidic silanol oxygens and their Si atoms in proximity (excluding the covalent-bonded ones) are systematically smaller than 3.5 Å, whereas that of the basic silanols are larger than 3.5 Å.

We then investigated the relative orientation of the acidic silanol Si-OH bond (that performs the nucleophilic attack) with respect to the SiO_4 tetrahedron being attacked. The angle formed between O of an acidic silanol

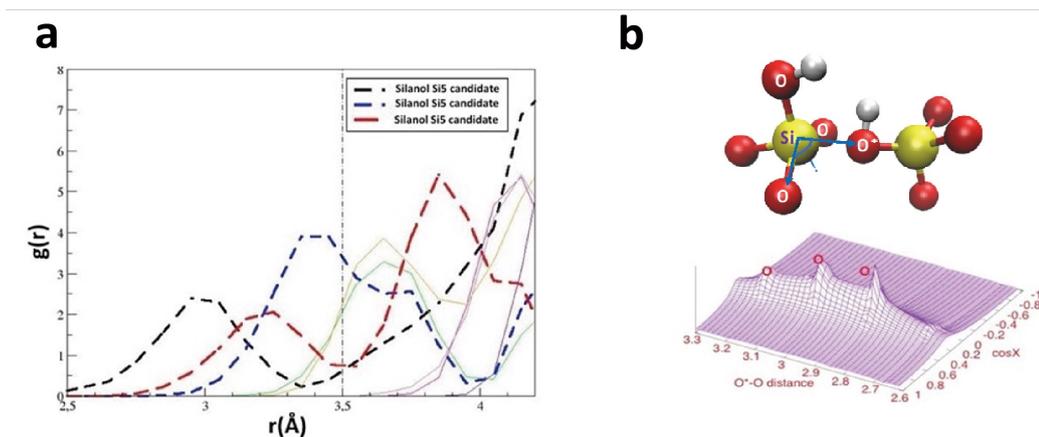


Figure 5.9: (A) Si-O radial distribution curves of non-covalently bonded Si atoms around the oxygen atom of each of the eight SiOH groups located on the surface of the 4.5 OH/nm² silica model. There are 8 plots reported, among which the red, black and blue dashed ones correspond to the SiOH groups whose deprotonation leads to the formation of $\text{Si}(5\text{c})^-$. (b) The top frame is a snapshot from the DFT-MD simulations that describes the SiO_4 tetrahedron in optimum orientation with respect to the reactive SiOH for the formation of the $\text{Si}(5\text{c})^-$ species. The bottom frame shows the distribution of $\text{Si}(5\text{c})^-$ formation with as a function of the O^*-O distance and of the angle x (both are defined in the top frame) between SiOH and SiO_4 . The three peaks marked with red circles identify the three most probable geometries for $\text{Si}(5\text{c})^-$ formation found in our simulations.

and the Si-O of the SiO_4 tetrahedron under attack is systematically $\sim 90^\circ$ (Fig. 5.9B, also detailed in Section S4 of the SI of our paper [*“Unveiling structural evolution and reconstruction of oxide surface in water”*]). The above structural analyses inform that proximity and appropriate relative orientation between SiOH and a neighboring SiO_4 tetrahedron are key factors for the formation of $\text{Si}(5\text{c})^-$.

To conclude, the higher stability of $\text{Si}(5\text{c})^-$ over SiO^- as the conjugate base for the acidic SiOH is due to geometric parameters on the silica surface, which must be favorable for the formation of a new SiOSi bond, while it does not strongly depend on the interactions with water.

5.6 Existence of $\text{Si}(5\text{c})^-$ on other amorphous silica and quartz model surfaces?

The next question to answer is whether the morphology of the silica surface has a relevant role in the formation of $\text{Si}(5\text{c})^-$. By this we mean, is the density of silanols (the results reported in the above sections are based on a 4.5 SiOH/nm² density of surface silanols) a relevant parameter, is amorphous versus crystalline surface a relevant parameter (all results discussed above were obtained for amorphous silica)? We investigate these two points in this section.

5.6.1 Other amorphous silica model surfaces

While the experimental amorphous silica surface typically has an average density of 4.5 SiOH/nm², local patches of higher silanol densities due to the heterogeneity of the amorphous surface also exist. Therefore, we also checked the appearance of Si(5c)⁻ species on amorphous surface models with a higher degree of hydroxylation (5.5 and 7.6 SiOH/nm² [163, 176]) through the same metadynamics method as the ones discussed in section 5.3.1.

Note that while we adopted the Ugliengo models [163] for the 4.5 and 5.5 SiOH/nm² surfaces, the surface model from Bernasconi [176] was adopted for the 7.6 SiOH/nm² surface. The two models differ in the way they are constructed, allowing to further validate the robustness of our results. The models from Ugliengo were derived from a bulk cristobalite simulated at very high temperature and then cooled down to room temperature by classical molecular dynamics. The bulk was then cut along one of the cell axes and the resulting dangling bonds were passivated by silanols. The Bernasconi model has been obtained by cutting a bulk silica model simulated at high temperature by classical MD and annealed by ab-initio MD.

The formation of Si(5c)⁻ was systematically observed on these model surfaces regardless of their degrees of hydroxylation and preparation methods.

5.6.2 Crystalline Quartz model surface

The α - (0001)-quartz/water interface shows similar bimodal behavior in titration [50] as amorphous silica observed by simulations. In ref. [50], the bimodal behavior of surface acidities was observed and ascribed to the different O-H orientations and H-bonded properties to water of surface silanols. The out-of-plane silanols (with silanol O-H pointing towards water making strong H-bonds with water) were indeed found to have an acidic pKa of 5.6, while the in-plane silanols (with silanols O-H parallel to the surface and being weak H-bonds acceptors from water) were found with a basic pKa of 8.6.

To understand whether the different conjugate bases (Si(5c)⁻ and SiO⁻) after deprotonation of silanols with different pKas also exist on the crystalline surface, we performed the same metadynamics simulations on α - (0001)-quartz aqueous surface as the ones discussed in section 5.3.1 by adopting the same simulation box and DFT set-up as in ref. [50]. We calculated the pKa values of one in-plane silanol group and one out-of-plane silanol group at the surface and we found again two values of pKa: acidic (3.9) and basic pKa (8.3), respectively, in line with the values reported in ref. [50] using another pKa calculation approach.

However, according to the free energy landscapes depicted in Fig. 5.10, the SiO⁻ is now the most stable conjugate base of both acidic and basic silanols. There is indeed no minimum associated to the Si(5c)⁻ species after the deprotonation of a silanol group with an acidic pKa. There is only a minimum corresponding to SiO⁻ species. Therefore, as also demonstrated in

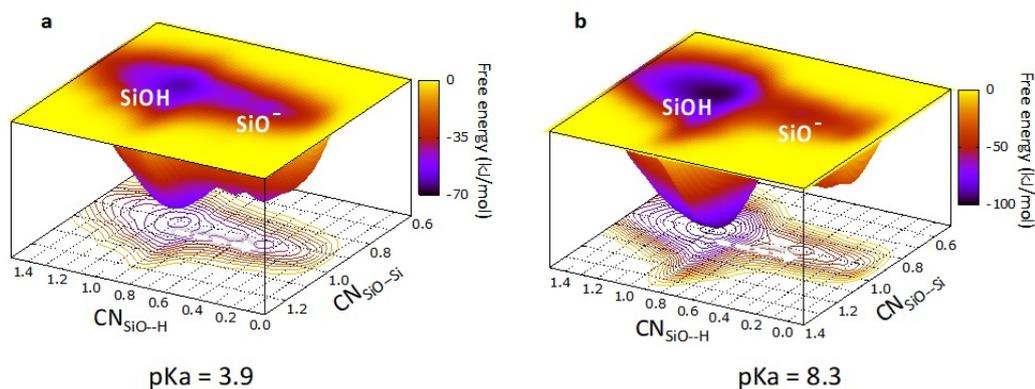


Figure 5.10: Free energy landscapes of α -quartz aqueous surface associated with deprotonation processes of (A) acidic silanols ($\text{pK}_a = 3.9$) and (B) basic silanols ($\text{pK}_a = 8.3$).

ref. [50], the split into acidic and basic SiOH populations on the α -(0001)-quartz surface is indeed due to different O-H orientations of silanols instead of due to different end products of deprotonation obtained in the case of amorphous silica/water interfaces.

To understand why $\text{Si}(5\text{c})^-$ is not a stable conjugate base of acidic silanols on the α -(0001)-quartz aqueous surface, the structural properties of the proximity of silanol groups were investigated. The radial distribution functions of surface Si atoms surrounding O atoms of silanols show that the distances between surface silanol oxygens and their Si atoms in proximity (excluding the covalent-bonded ones) are systematically larger than 3.5 Å. This hence does not comply with the requirements for the $\text{Si}(5\text{c})^-$ formation found in section 5.5.2 ($r_{\text{O-Si}} < 3.5$ Å).

Therefore, we propose that $\text{Si}(5\text{c})^-$ formation is less probable on crystalline than on amorphous silica surfaces, as amorphous surfaces have a larger “structural flexibility/diversity” leading to a higher probability to meet the local geometrical requirement for $\text{Si}(5\text{c})^-$ formation. This has to be confirmed over a large set of crystalline silica facets.

However, interestingly, we note that in ref. [177], $\text{Si}(5\text{c})^-$ structures have been found on the dry facet (011) of the α -quartz surface due to the reconstruction process following the homolytic fracture of α -quartz crystal at room temperature. Hence, we expect $\text{Si}(5\text{c})^-$ to be possibly formed on certain crystalline quartz aqueous surfaces as long as the required structural conditions are met at a given surface cut. This could be a very nice future project.

5.7 Conclusions

The bimodal behavior of two groups of silanol with respectively acidic and basic pK_a values at the silica surface has long been debated since it was first discovered by Eienthal *et al.* [46] in SHG experiments. Despite a substantial

amount of experimental and theoretical works on it, the origin of this bimodal behavior is still elusive due to the difficulty of probing buried silica surfaces in contact with water. All the existing studies in the literature naturally assume that the deprotonation of SiOH leads to the formation of SiO⁻, without actual direct proof.

We challenged this very fundamental assumption by this joint theory/experiment work. Our experimental collaborator, the group of Prof. Wei-Tao Liu at Fudan University in Shanghai and Prof. Y.R. Shen at UC Berkeley, developed an in situ experimental SFG spectroscopy that probes the Si-O stretch (phonon) range. This allows, for the first time, to monitor the change in the surface structure under different pH conditions. The recorded SFG spectra of the silica surface in water as a function of pH surprisingly revealed (1) the non-conservation of SiOH and SiO⁻ intensities in the SFG as a function of pH and (2) the red shift of SiOH band in the intermediate pH range (4 to 9), which leads to the conclusion that a new surface species other than SiOH and SiO⁻ was formed in that pH range.

In our theory part, by performing metadynamics simulations, we uncovered a new surface species, Si(5c)⁻, a five-coordinated species that only exists in the intermediate pH range (4 to 9). Using the calculated pKa values from metadynamics simulations, we deduced the surface coverage of each silica surface species (SiOH, Si(5c)⁻, SiO⁻) as a function of pH, that allowed us to construct the pH-dependent theoretical SFG signals without performing DFT-MD simulations under different pH conditions. The theoretical SFG spectra matched very well the experimental spectra, and confirmed that the red-shift of the SiOH band was indeed due to the formation of Si(5c)⁻ species.

Additionally, we calculated theoretical SHG based on our newly uncovered 3-species model, which was in perfect agreement with the historical titration SHG plot shown in Ref. [46]. This further confirmed our discovery of the Si(5c)⁻ species.

In quest of the origin of the formation of the Si(5c)⁻ species, a set of additional DFT-MD simulations and geometry analyses confirmed that the formation of Si(5c)⁻ species was actually due to geometric reasons instead of due to interactions between surface and water. We found the existence of Si(5c)⁻ species on other amorphous silica models of various silanol densities where the geometry criteria were met, but no Si(5c)⁻ was formed on the α -(0001)-Quartz surface due to the unsatisfied geometric parameters.

We believe that the discovery of 5-coordinated surface species, the innovative experimental technique, as well as the new strategy of combining metadynamics simulations and pH-dependent theoretical SFG spectra calculation without the need of performing simulations under different pH conditions, will be of high impact on future research in our field.

5.8 Perspectives

As regards the possibility to observe a similar pH-induced surface reconstruction on other oxide surfaces than Silica, such as Alumina, there are some very interesting scenarios that can be foreseen based on our results, which will require and hopefully motivate many future studies. In particular, we have shown that the pH-induced surface reconstruction is dictated by well defined local geometrical parameters of the surface: the proximity and orientation of neighboring SiO_4 surface groups around a candidate site to form $\text{Si}(5c)^-$ after deprotonation has to be favorable to host the formation of a novel Si-O-Si bridge. This requisite is more easily met at amorphous surfaces, where there is a larger diversity of local surface configurations available with respect to crystalline surfaces. However, there is no reason to think such conditions can only be satisfied at amorphous silica surfaces. On the contrary, it is quite reasonable to think that other amorphous surfaces, such as γ -Alumina surfaces, can satisfy them.

Moreover, the pH induced surface reconstruction could be even more likely to be observed at these Alumina surfaces than it is for Silica, as surface Al groups have a larger number of stable coordination environments available as compared to Si: (1) while Si is almost always found with 4-fold coordination, Al surface atoms can be observed in different coordinations ranging from 4-fold to 6-fold depending on surface geometry and degree of crystallinity; (2) while O atoms are only 1-fold or 2-fold coordinated to heavy atoms on silica surfaces, they can be 1-/2-/3-fold coordinated to Al atoms at Alumina surfaces.[178–180] This higher “structural flexibility” can be reasonably expected to result in a lower free energy cost to change the coordination of Al and O atoms on Alumina surfaces as compared to Si and O atoms on Silica surfaces, stabilizing even more surface reconstructions leading to the formation of new Al-O-Al bonds.

Same reasoning can be applied to e.g. titanium oxides or cobalt oxides. These oxides could be investigated theoretically with all tools described in this chapter and the most likely candidates could then be investigated in experiments for final proofs of the surface chemistry.

We now report our paper [“Unveiling structural evolution and reconstruction of oxide surface in water”] where all results presented in this chapter are fully discussed This paper is in its final (third) stage of review for the possible publication on Nature (2023.02).

Unveiling structural evolution and reconstruction of oxide surface in water

Xiaoqun Li^{1,6}, Flavio S. Brigiano^{2,3,6}, Simone Pezzotti^{2,4,6}, Xinyi Liu^{1,6}, Wanlin Chen^{2,6}, Huiling Chen¹, Ying Li¹, Hui Li¹, Y. Ron Shen^{1,5}[✉], Marie-Pierre Gaigeot²[✉], Wei-Tao Liu¹[✉]

¹Physics Department, State Key Laboratory of Surface Physics, Key Laboratory of Micro and Nano Photonic Structures [Ministry of Education (MOE)], Fudan University, Shanghai 200433, China.

²Université Paris-Saclay, Univ Evry, CNRS, LAMBE UMR8587, 91025, Evry-Courcouronnes, France.

³Current address: Laboratoire de Chimie Théorique, Sorbonne Université, UMR 7616 CNRS, 4 Place Jussieu, 75005 Paris, France.

⁴Current address: Physical Chemistry II, Faculty of Chemistry and Biochemistry, Ruhr University Bochum, Universitätsstr. 150, 44780 Bochum, Germany.

⁵Department of Physics, University of California at Berkeley, Berkeley, CA 94702, USA.

⁶These authors contributed equally: Xiaoqun Li, Flavio S. Brigiano, Simone Pezzotti, Xinyi Liu, Wanlin Chen

[✉]Corresponding author. Email: yrshen@berkeley.edu, mgaigeot@univ-evry.fr, wliu@fudan.edu.cn

Aqueous interfaces of oxides are ubiquitous on our planet, playing a key role in the ecosystem: for example, protonation/deprotonation of silicates in water dominates the weathering process that shapes the Earth landscape, and serves as the major sink in global carbon cycle¹. In modern industrial applications, oxide/water interfaces also host many important reactions, such as the oxygen evolution in renewable energy schemes². However, accurate knowledge about these interfaces still lacks at the molecular level, due to difficulties to access buried oxide surfaces^{3,4}. Here we report an experimental scheme enabling *in situ* sum-frequency vibrational spectroscopy of oxide surfaces in liquid water. Application to the silica/water interface revealed surprisingly the emergence of new reaction pathways with

water. With *ab initio* molecular dynamics and metadynamics simulations, we uncovered a surface reconstruction, triggered by deprotonation of surface hydroxylated groups, that led to unconventional five-coordinated silicon species. The result demystifies the multimodal chemistry of aqueous silica discovered decades ago^{5,6}, bringing out new information that radically modifies our current understanding. Most exciting of all, our study provides new opportunities for future in-depth physical and chemical characterizations of other oxide/water interfaces.

Being nearly impenetrable to charged particles/atoms, and incompatible with low temperature and vacuum systems, the solid/liquid interfaces are notoriously difficult to probe. Among all available tools, the surface-specific nonlinear optical methods are arguably the most viable^{7,8}. In 1992, Eienthal and coworkers first employed second harmonic generation (SHG) to monitor the silicon dioxide (silica, SiO₂)/water interface, and discovered the classical bimodal titration behavior⁵. Soon afterwards, Du *et al.* performed the first sum-frequency vibrational spectroscopy (SFVS) of the silica/water interface, initiated the quest for a molecular-level understanding of the interface⁶. Since then, benefiting from the ever-growing theoretical power^{9,10}, enormous advances of our perception on oxide aqueous interfaces have been achieved¹¹. The SiO₂/water interface has served as the most studied model system due to its ubiquity in nature. The first phase-sensitive SFVS of interfacial water was realized with this system¹², as well as the first ultrafast dynamic study of water interfaces¹³, the first investigation of flow effects on surface chemistry¹⁴, and many more¹¹. Each marked a milestone in surface science. However, while the water side has been extensively studied, the oxide side remains barely investigated. Even though SFVS has been used successfully to resolve surface structures of silica/quartz, titania, and alumina in gaseous ambient¹⁵⁻¹⁷, extension to the water environment is challenging: infrared light for excitation of oxide surface vibrations is strongly absorbed by both water and oxide, and has difficulty to access

their interface. Lack of structural information of oxides in liquid water has left many fundamental questions unanswered: what are the lattice sites hosting the multimodal chemistry, how they interact with water molecules and hydrous ions as the water solution varies, and why different titration behaviors are occasionally observed¹⁸⁻²⁶, etc. It also causes difficulties in understanding interfacial water structure and properties, for interpretation of results on water must rely on prior knowledge about the oxide structure.

Here we report an experimental scheme that can change the scene: a scheme that allows *in situ* SFVS of oxide surface lattices in liquid water. Using a high refractive index infrared wafer/oxide/water layer structure – a planar cavity – with appropriate oxide film thickness and excitation geometry, we can specifically optimize the local optical fields at the oxide/water interface, so that the SF signal from the interface is dominant and readily detectable (see Methods). The experimental arrangement of our broadband SFVS measurements is sketched in Fig. 1a. The broadband infrared input, passing through an infrared transparent wafer coated by the oxide thin film, overlapped on the layer structure with the narrowband near-infrared beam incident from the water side, and the SF output was collected in the “reflected” direction (see Methods and Supplementary Section 1.1). We utilized a silicon (Si) /silica/water layer structure, with the near-infrared at 800 nm and broadband infrared centered at 1,000 cm^{-1} to match the Si-O stretch range. Employing the nonlinear optical matrix formalism to calculate the field distributions across the layer structure, we found that despite the strong infrared attenuation in silica, the total SF response from the silica/water interface could be drastically enhanced if the silica film thickness was near 150 nm (Fig. 1b). Meanwhile, the response from the silicon/silica interface remained low, ensuring the SF output to be dominantly from the silica/water interface (see Methods and Supplementary Section 1.2). We prepared the 150 nm amorphous silica film on an *n*-type Si wafer by plasma-enhanced chemical vapor deposition. To be sure of the sample surface quality, we took SF

spectrum of the thin-film-silica/air interface in the Si-O stretch range, and found it closely resembled that of a bulk-silica/air interface (Fig. 1c), both exhibiting a single prominent resonance at $\sim 970\text{ cm}^{-1}$ from the surface silanol (SiOH) groups. The stretch modes of surface siloxane (Si-O-Si) bridges were not seen, for those bonds were oriented randomly over the amorphous silica surface and hence inactive in SFVS¹⁵. Further tests were performed and confirmed the sample surface to be of good quality (see Methods).

***In situ* SFVS of silica surface in water**

We then acquired a series of SF spectra in the Si-O stretch range from the silica/water interface, at varying pH with a sodium chloride (NaCl) buffer to keep the ion activity constant (10 mM, Fig. 1d). The beam polarization combination used was SSP, denoting S-polarized SF output, S-polarized near-infrared input, and P-polarized infrared input, respectively. The protonated SiOH and deprotonated SiO⁻ (Fig. 1e) are generally considered the major surface species on silica reacting with water. At pH 2, the silica surface was fully covered by silanol groups, and the spectrum showed a single peak due to Si-OH stretching vibration at $\sim 980\text{ cm}^{-1}$ with a full-width of $\sim 40\text{ cm}^{-1}$, similar to that of the silica/air interface¹⁵. Increasing pH gradually deprotonated SiOH, with the mode intensity dropping accordingly, accompanied by a redshift to $\sim 940\text{ cm}^{-1}$ before disappearing near pH 10^{27,28}. At that point, a new mode at $\sim 1,020\text{ cm}^{-1}$ emerged. This mode was not detected at the silica/air interface, but could appear on humidified silica particles and was attributed to the Si-O⁻ stretching vibration of silanolate groups^{27,28}. As expected, it grew stronger at higher pH values (Fig. 1d). The above features were robust under pH cycling without noticeable hysteresis (Extended Data Fig. 1). Our result constitutes the first set of SF vibrational spectra from oxide surface lattices ever recorded in liquid water.

Nonetheless, the observed spectra revealed an unexpected relationship between SiOH and SiO⁻. In all existing models for this interface, the two species are believed to have a one-to-one correspondence, converting to each other via $\text{SiOH} + \text{OH}^- \leftrightarrow \text{SiO}^- + \text{H}_2\text{O}$. However, as seen in Fig. 1d, or more clearly from the two-dimensional intensity map in Fig. 2a, the SiO⁻ mode emerged only when the SiOH mode almost disappeared. We further took spectra with the PPP beam polarization combination (all beams being P-polarized), as well as spectra with a different ion concentration (0.5 M) (Figs. 2b-2d). In all cases, we observed SiOH and SiO⁻ bands both being very weak around pH 9~10, indicating that the two groups did not follow the direct one-to-one conversion. In previous studies on silica nanoparticles, both bands were found to coexist over a broad range of pH^{27,28}. However, that was because the pKa values for silanol deprotonation of nanoparticles have significant size dependence, and can be much lower in smaller particles with diameters below ~100 nm²⁹.

Quantitatively, we could deduce the mode amplitudes (labelled A_{SiOH} and A_{SiO^-}) in each spectrum via fitting. From SSP and PPP spectra¹⁵, we concluded that both Si-OH and Si-O⁻ stretches aligned toward the surface normal with little variation versus pH (Figs. 2b-2d). This was further confirmed by the observation that both modes were hardly detectable on SPS spectra (see Supplementary Section 1.1). Meanwhile, neither band exhibited detectable splitting or width change of statistical significance. We were then able to claim that A_{SiOH} and A_{SiO^-} were directly proportional to the surface densities of SiOH and SiO⁻ groups, respectively. Figures 2e and 2f display the normalized amplitudes versus pH. The data could not be explained by the standard Gouy-Chapman-Stern (GCS) model with a single pKa value for the SiOH \leftrightarrow SiO⁻ interconversion (see Supplementary Section 1.1), nor by the model with two pKa values either⁵ (dashed curves in Fig. 2e, f), as the criterion of one-to-one direct conversion between SiOH and SiO⁻ in the model was obviously not satisfied in reality. To search for possible additional surface species, we

extended our SFVS spectral range covering from ~ 800 to $1,200\text{ cm}^{-1}$ (Extended Data Figure 2), and found no new resonance. Overall, the result challenges the conventional two-state model and calls for a new understanding on how the silica surface evolve with pH in water. In the following, we used *ab initio* molecular dynamics simulations to elucidate the structure and chemistry of silica/water interface, in direct comparison to spectroscopic results.

***Ab initio* simulations**

The surface chemistry triggered by deprotonation of silanols was uncovered using metadynamics simulations in the DFT framework. We adopted the amorphous silica surface model with a silanol density of 4.5 OH/nm^2 developed by Ugliengo *et al.*³⁰, which has successfully described the experimental observations in many previous studies^{24,25} (see Methods). One metadynamics simulation was performed for each SiOH being deprotonated (with a total number of 8 on the model surface). Surprisingly, we found only 5 out of 8 of the SiOH groups converted to SiO⁻. The deprotonation of the remaining SiOH triggered a surface reconstruction that led to a five-coordinated silicon species [Si(5c)] as the stable conjugate base. Figures 3a and 3b show, respectively, the representative free energy landscapes for the two reactions: SiOH(I) \leftrightarrow Si(5c), and the conventional SiOH(II) \leftrightarrow SiO⁻. Both landscapes were constructed using two coordination numbers of the reacting oxygen in a chosen silanol: one was specified with respect to all hydrogen atoms at the surface ($\text{CN}_{\text{SiO--H}}$), and the other specified by that with respect to all silicon atoms ($\text{CN}_{\text{SiO--Si}}$) (see Methods). Both coordination numbers were defined by a function of inter-atomic distances introduced in Ref. 31. With such coordinates, a reaction pathway with component parallel to the $\text{CN}_{\text{SiO--Si}}$ axis describes a breakage/formation of a Si-O bond, allowing us to trace possible surface reconstructions that had never been considered before²⁰.

On each free energy landscape, free energy minima identified stable surface species. In Fig. 3a, we found a well-defined minimum for SiOH(I) as the reactant, but none for SiO⁻, meaning that the latter was *not* a stable reaction product. Instead, a surface reconstruction leading to five-coordinated silicon species [Si(5c)] was observed. This reconstruction proceeded through a concerted mechanism (Fig. 3c): upon deprotonating a SiOH(I) into SiO⁻ (a proton transferred to water), the resultant SiO⁻ reacted immediately with a neighboring four-coordinated silicon, forming a Si-O-Si bridge and a five-coordinated Si(5c) group. In most cases (~67%), the silicon under attack was the host of a SiOH(II), and a Si(5c)OH⁻ group is formed after the attack, with negative charges spreading around the Si(5c). In other cases (~33%), the silicon under attack had no hydroxyl attached, and the reaction led to a SiO₅⁻ species, again with negative charges distributed around the central silicon. Supplementary simulations confirmed the barrier-less formation and stability of both Si(5c) species at the room temperature. In contrast to SiOH(I), a different free energy landscape was seen for SiOH(II) (Fig. 3b): a free energy minimum was found for SiO⁻, but not for Si(5c), indicating that the former became the energetically favorable end product (see Supplementary Section 2.1).

The metadynamics further showed that the SiOH(I) groups were acidic with a pKa = 3.9±0.6, while the SiOH(II) groups were basic with a pKa = 7.7±0.9 (see Methods), which were in line with experimental values of ~4 and 9 found by SHG and other methods^{5,11}, successfully reproduced the multimodal deprotonation behavior of silica at varying pH value. We found that: (1) at intermediate pH, the Si(5c) species (including Si(5c)OH⁻ and SiO₅⁻) were the stable deprotonation product; (2) at high pH, the basic SiOH(II) converted to SiO⁻; (3) meanwhile at high pH, due to the increased negative surface charges, Si(5c)OH⁻ after deprotonation became unstable and split into two SiO⁻ groups by bond cleaving; similarly, each SiO₅⁻ split by releasing a SiO⁻ and a recovered four-coordinated silicon (see Supplementary Section 2.1). To check the universality of

our results, we further performed simulations on other silica models with various morphologies, preparation methods and degrees of hydroxylation, all confirming the spontaneous formation of Si(5c) and above-mentioned conclusions (see Supplementary Section 2.2).

Comparison between experiment and theory

To compare to experiments, we first used Henderson-Hasselbalch equations to find the coverages of various surface species at given pH, with parameters deduced from the DFT-MD. Ions were not explicitly included in the simulation, but their concentration was used to relate the effective H^+ concentration at the interface and that in bulk solution via modified Poisson-Boltzmann equation (see Supplementary Section 3.1). With that correction, we obtained $pK_1 = 6.7$ and $pK_2 = 11.6$ for an effective ion concentration of 10 mM. Figure 4a presented the calculated coverages of different species versus the pH. At $pH < pK_2$, deprotonation of SiOH hardly produced SiO^- but converted SiOH to the Si(5c) species (dashed green curve). Above that, the coverage of SiO^- became significant, agreeing with our experimental observation.

We could then calculate $\chi^{(2)}$ spectra in the Si-O stretching vibration region versus pH (see Methods). As shown in Extended Data Figure 3, the species-specific $|\chi_{SSP}^{(2)}|^2$ spectra of the four-coordinated SiOH [Si(4c)OH] and SiO^- species have resonances near 920 and 1,000 cm^{-1} at all pH (top and bottom panels), corresponding to those experimentally detected at $pH < 4$ and $pH > 10$, respectively. The calculated resonance frequencies were slightly lower than the experimental ones because of the DFT functionals used. The Si(5c)OH $^-$ species present in the mid-pH range had a spectrum centered at 840 cm^{-1} (middle panel), redshifted by ~ 80 cm^{-1} from the Si(4c)OH band (middle panel). This redshift resulted from the additional bridging oxygen attached to Si, which elongated the Si(5c)-OH $^-$ bond length by ~ 0.05 Å. Such shift was also observed in silica

nanoparticle studies and attributed to delocalized negative surface charges^{27,28}, our results now reveal the microscopic nature of it. The total $\left| \chi_{SSP}^{(2)} \right|^2$ spectra calculated over a range of pH from 2 to 12 (Figs. 4b & 4c) are in good accordance with the experimental spectra at 10 mM (Figs. 1d & 2a). Overlapping of the Si(4c)OH and Si(5c)OH⁻ modes in the intermediate pH range caused them to merge into one band (the merging would be more seamless if inhomogeneous broadening were included in the calculation). As seen in the experiment, it showed an apparent redshift due to the growth of the Si(5c)OH⁻ mode at increasing pH. At 0.5 M ion concentration, the theoretical pKa values were found to shift to 4.8 and 9.8 (see Supplementary Section 3.1), respectively, also nicely in line with the experimental trends in Fig. 2.

Our simulation could thus satisfactorily explain the multimodal titration behavior observed by SHG⁵. The pH dependence in SHG was in response to the variation of SiO₂ surface charge density. Knowing the charge density from the coverages of Si(5c)⁻, Si(5c)OH⁻ and SiO⁻ at given pH, we could then calculate $\left| \chi_{SHG}^{(2)} \right|$ for SHG (see Supplementary Section 3.2). As seen in Fig. 4d, the SHG variation with pH⁵ could be reproduced only with the Si(5c) species taken into accounts: while the appearance of SiO⁻ is responsible for the sharp rise at pH > 9, the Si(5c) species are responsible for the “plateau” between pH 4 and 9.

With the success of simulation demonstrated, we can now answer the long-standing question: what determined the pKa value of a silanol on silica^{5,12,18,20}? By analyzing the local structure around a silanol, we found that the pKa values did not depend on whether the silanol was vicinal or isolated, neither on the nearby density of water molecules, nor on the local hydrogen bonding environment. Instead, it was the geometric proximity between the silanol and neighboring SiO₄ tetrahedrons being instrumental for the Si(5c) formation, and determining the acidic nature of such silanols (see Supplementary Section 4). This finding further explained why the microscopic

inhomogeneity and surface preparation protocol could significantly alter the silica deprotonation behaviors^{21,26, 29,32}, as it is dominated by the local structure around silanols.

Additional discussion and conclusion

Five-coordinated silicon species are considered rare in nature, appearing under extreme conditions such as in the Earth mantle³³⁻³⁵. Mantle matters, such as molten alkali silicates, can host non-bridging SiO⁻ bonds coordinated by alkaline cations; when an alkaline cation moves away, the remaining SiO⁻ bond will immediately attach to a neighboring four-coordinated silicon and form a Si(5c), closely analogous to the pathway converting deprotonated SiOH to Si(5c)OH⁻ in our system (Fig. 3c). Moreover, theoretical studies have proposed the relevance of five-coordinated silicon as intermediate in SiO₂ nucleation and hydrolysis^{36,37}. As the surface reconstruction is dictated by local geometric parameters, we further expect similar behaviors of other oxides, such as γ -alumina surfaces that have more varieties of coordinations.

To summarize, with *in situ* spectroscopic and theoretical methods we developed, we have reached a new level of understanding of the classical silica/water interface. It also calls for the re-investigation of interfacial water properties, now that we have a more realistic oxide surface structure in hand. Unexpected findings from such an exhaustively studied system assert the relevance and urgency of *in situ* studies on all oxide/water interfaces, especially with joint effects of various techniques (for example, high precision NMR), where new phenomena likely await discovery.

Figures

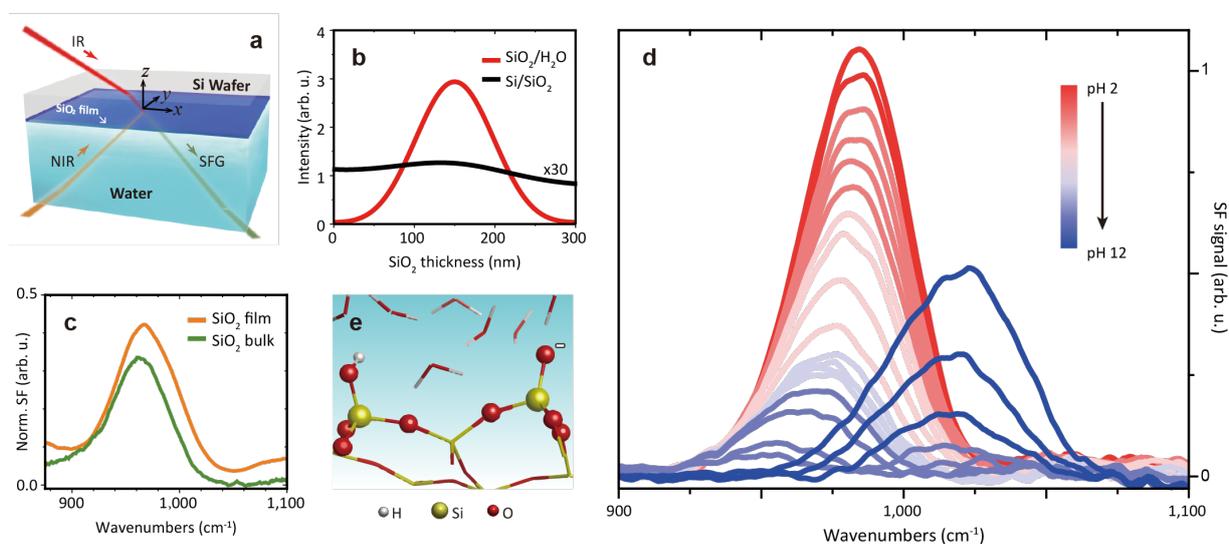


Figure 1 | Experimental arrangement, interfacial field enhancement and sum frequency spectra of silica interfaces. **a**, The sample and beam excitation/detection geometry for SFVS. **b**, Calculated SF intensity from the SiO₂/H₂O (red) and Si/SiO₂ (black) interfaces of a Si/SiO₂/H₂O junction, as functions of the SiO₂ film thickness. The black curve for the Si/SiO₂ interface is magnified by 30 times for clarity. **c**, SF spectra of the Si-OH stretching vibrational mode at the air/bulk-SiO₂ (green) and air/thin-film-SiO₂ (150 nm) (orange) interfaces, both with the Fresnel factors corrected. Peaks at ~970 cm⁻¹ are due to the stretch vibration of surface Si-OH bonds. **d**, SF spectra taken *in situ* from the SiO₂/H₂O interface of a Si/thin-film-SiO₂ (150 nm)/H₂O junction with pH increasing from 2 to 12. The 950~980 cm⁻¹ band within pH < 10 comes from the Si-OH stretching vibrational modes, and the ~1,020 cm⁻¹ band from the Si-O⁻ stretch. All spectra were taken with the SSP beam polarization combination. **e**, Illustration of the SiO₂/H₂O interfacial structure. A protonated SiOH (left) and a deprotonated SiO⁻ (right) are presented in ball-and-stick. Hydrogen, silicon, and oxygen atoms are in white, yellow, and red, respectively.

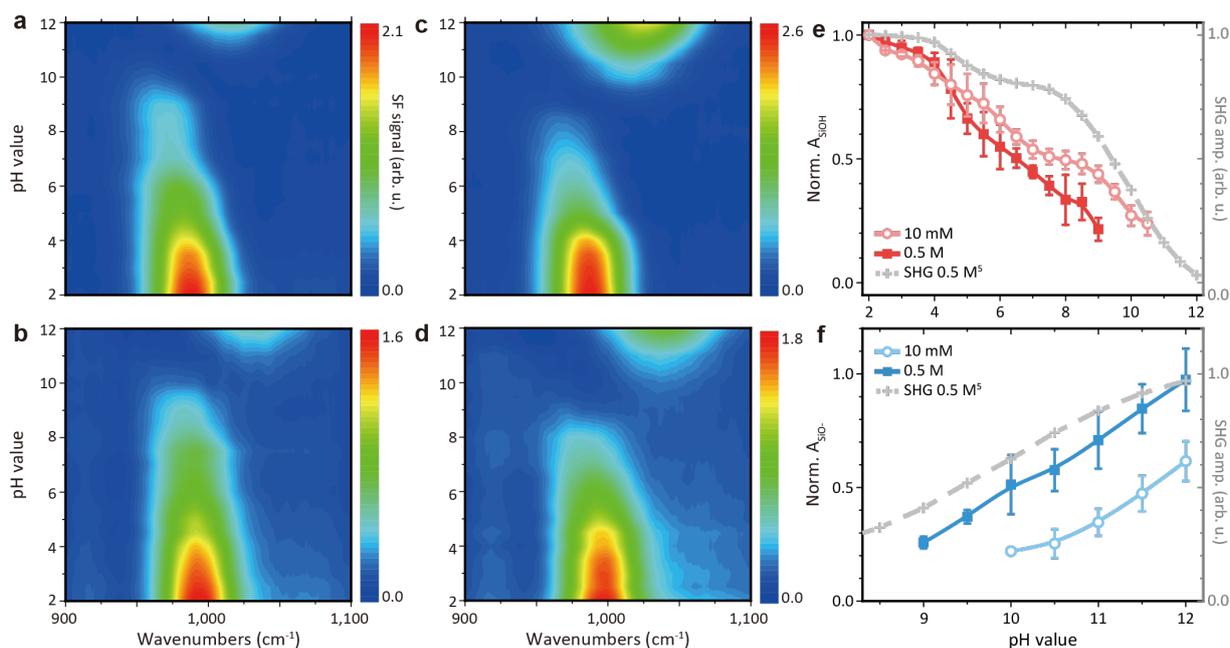


Figure 2 | Evolution of the Si-O stretching vibrational spectra from silica/water interfaces with pH in water. Two-dimensional displays of spectra versus pH taken with different beam polarization combinations and different ion concentrations (NaCl as the buffer solution): **a**, SSP, 10 mM; **b**, PPP, 10 mM; **c**, SSP, 0.5 M; **d**, PPP, 0.5 M. **e**, SSP amplitudes of the SiOH stretching vibrational mode deduced from fitting of the SSP spectra versus pH at 10 mM (empty symbols) and 0.5 M (filled symbols) of NaCl concentration, respectively. **f**, SSP amplitudes of the SiO⁻ stretching vibrational mode at 10 mM (empty symbols) and 0.5 M (filled symbols), respectively. $n = 3$, mean \pm SEM. The dash-cross curves in **e** and **f** were deduced from the model used to explain the SHG results in Ref. 5. Solid curves are guides for eyes. All amplitudes are normalized with respect to the corresponding A_{SiOH} at pH = 2 for comparison.

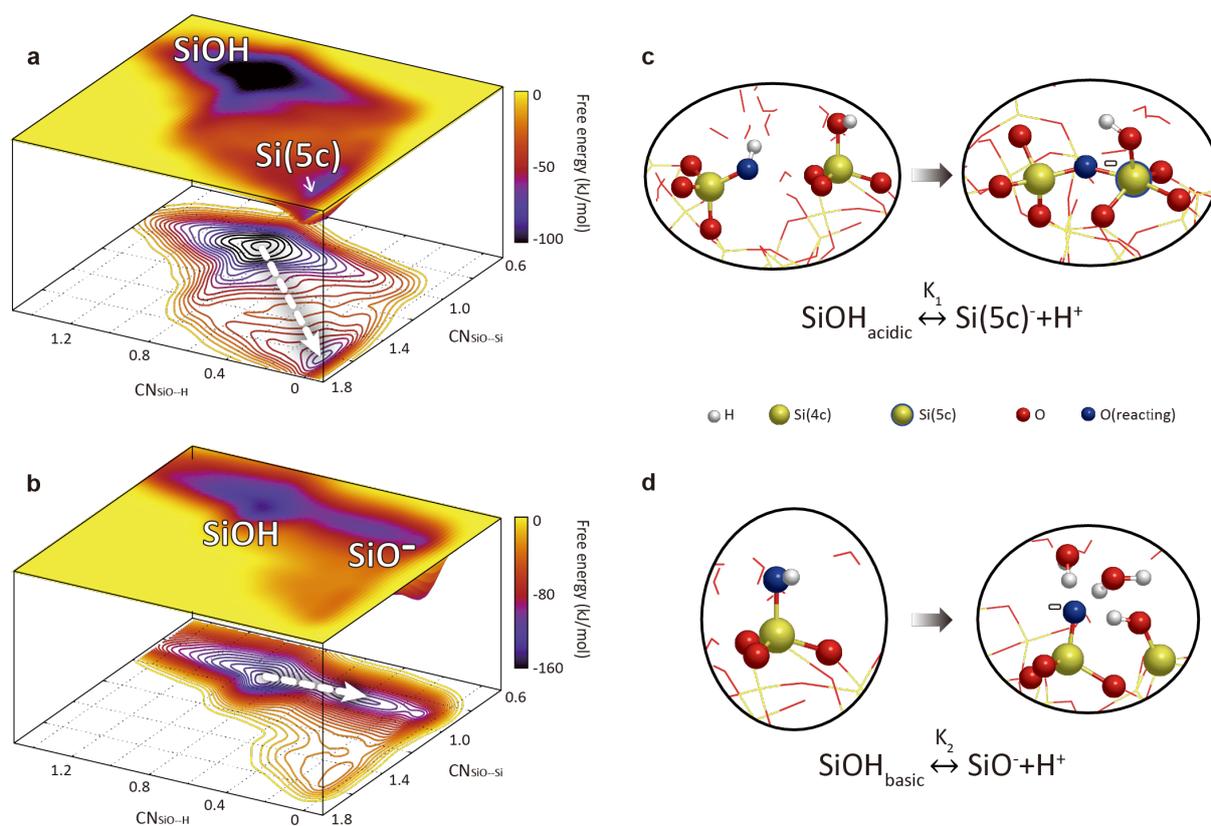


Figure 3 | MD Simulation results of the SiO₂/H₂O interfacial reactions. **a**, Representative free energy landscape associated with deprotonation of an acidic SiOH(I) leading to the formation of Si(5c) as conjugate base. Top and bottom frames describe the 3D and projected 2D free energy maps, respectively. Two free energy minima can be seen, associated with stable surface species of SiOH(I) and Si(5c), respectively. The white dashed arrow marks the transition from the minimum of SiOH(I) to the minimum of Si(5c). **b**, Representative free energy landscape associated with deprotonation of a basic SiOH(II) leading to the conventional formation of SiO⁻. The free energy minimum associated with Si(5c) is no longer present and is replaced by a new minimum for SiO⁻. The white dashed arrow marks the transition from the minimum of SiOH(II) to the minimum of SiO⁻. **c**, Snapshots of the deprotonation process of an acidic SiOH group that leads to the formation of Si(5c), which is as described in **a** and characterized by the reaction equilibrium constant K_1 . **d**, Snapshots of the basic SiOH deprotonation process leading to SiO⁻, as described in **b** and characterized by K_2 . In **c** and **d**, oxygen atoms of the reacting SiOH groups are marked in blue; hydrogen, silicon, and other oxygen atoms are in white, yellow, and red, respectively.

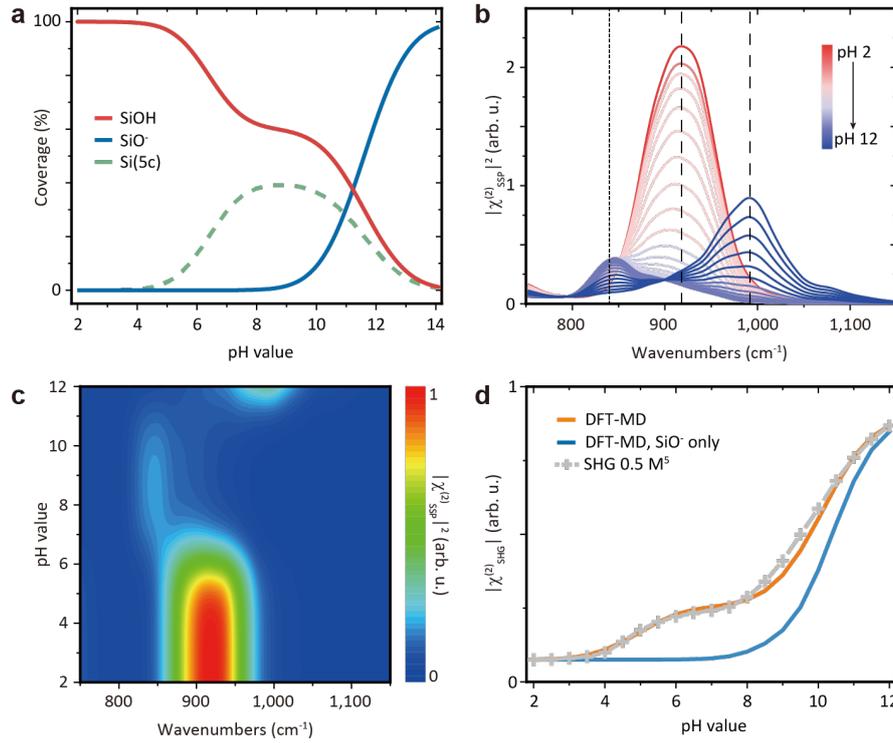


Figure 4 | Calculated surface coverages and nonlinear optical susceptibilities. **a**, Coverages of various surface species versus pH in water: SiOH (red curve), SiO⁻ (blue curve), and Si(5c) (dashed green curve). **b**, Calculated $|\chi_{SSP}^{(2)}|^2$ spectra of the silica/water interface in the Si-O stretch range over a range of pH values. The color scheme is the same as that in Fig. 1d. Vertical lines mark the stretch frequencies of individual Si(5c)-OH⁻, Si(4c)-OH⁻, and Si-O⁻ groups (from low to high frequency). **c**, Two-dimensional display of the $|\chi_{SSP}^{(2)}|^2$ spectra with respect to pH as in **b**, which can be compared to the experimental results in Fig. 2a. **d**, Comparison of the calculated $|\chi_{SSP}^{(2)}|^2$ versus pH with that derived from SHG experiment in Ref. 5 (normalized to match the values at pH 12). The blue curve of theoretical $|\chi_{SHG}^{(2)}|^2$ was calculated by assuming surface charges were carried only by SiO⁻ (i.e. neglecting charges carried by Si(5c) species).

References

1. Taifan, W., Boily, J. & Baltrusaitis, J. Surface chemistry of carbon dioxide revisited. *Surf. Sci. Rep.* **71**, 595-671 (2016).
2. Nong, H. N. et al. Key role of chemistry versus bias in electrocatalytic oxygen evolution. *Nature.* **587**, 408-413 (2020).
3. Zaera, F. Probing Liquid/Solid interfaces at the molecular level. *Chem. Rev.* **112**, 2920-2986 (2012).
4. Gonella, G. et al. Water at charged interfaces. *Nat. Rev. Chem.* **5**, 466-485 (2021).
5. Ong, S., Zhao, X. & Eissenthal, K. B. Polarization of water molecules at a charged interface: Second harmonic studies of the silica/water interface. *Chem. Phys. Lett.* **191**, 327-335 (1992).
6. Du Q, Freysz, E. & Shen, Y. R. Vibrational spectra of water molecules at quartz/water interfaces. *Phys. Rev. Lett.* **72**, 238-241 (1994).
7. Geiger, F. M. Second harmonic generation, sum frequency generation, and $\chi^{(3)}$: Dissecting environmental interfaces with a nonlinear optical swiss army knife. *Annu. Rev. Phys. Chem.* **60**, 61-83 (2009).
8. Covert, P. A. & Hore, D. K. Geochemical insight from nonlinear optical studies of mineral-water interfaces. *Annu. Rev. of Phys. Chem.* **67**, 233-257 (2016).
9. Rimola, A., Costa, D., Sodupe, M., Lambert, J. & Ugliengo, P. Silica surface features and their role in the adsorption of biomolecules: Computational modeling and experiments. *Chem. Rev.* **113**, 4216-4313 (2012).
10. Kubicki, J. D. *Molecular modeling of geochemical reactions: An introduction.* (Wiley, Chichester, West Sussex, United Kingdom, 2016).
11. Backus, E. H. G., Schaefer, J. & Bonn, M. Probing the mineral-water interface with nonlinear optical spectroscopy. *Angew. Chem. Int. Ed.* **60**, 10482-10501 (2020).
12. Ostroverkhov, V., Waychunas, G. A. & Shen, Y. R. New information on water interfacial structure revealed by phase-sensitive surface spectroscopy. *Phys. Rev. Lett.* **94**, 46102 (2005).
13. McGuire, J. A. & Shen, Y. R. Ultrafast vibrational dynamics at water interfaces. *Science.* **313**, 1945-1948 (2006).
14. Lis, D., Backus, E. H. G., Hunger, J., Parekh, S. H. & Bonn, M. Liquid flow along a solid surface reversibly alters interfacial chemistry. *Science.* **344**, 1138 (2014).
15. Liu, W. & Shen, Y. R. Surface vibrational modes of alpha-quartz(0001) probed by sum-frequency spectroscopy. *Phys. Rev. Lett.* **101**, 16101 (2008).
16. Tong, Y. et al. Optically probing Al-O and O-H vibrations to characterize water adsorption and surface reconstruction on alpha-alumina: An experimental and theoretical study. *J. Chem. Phys.* **142**, 54704 (2015).
17. Cao, Y. et al. Evolution of anatase surface active sites probed by in situ sum-frequency phonon spectroscopy. *Sci. Adv.* **2**, e1601162 (2016).
18. Leung, K., Nielsen, I. M. B. & Criscenti, L. J. Elucidating the bimodal acid-base behavior of the water-silica interface from first principles. *J. Am. Chem. Soc.* **131**, 18358-18365 (2009).
19. Flores, S. C., Kherb, J., Konelick, N., Chen, X. & Cremer, P. S. The effects of Hofmeister cations at negatively charged hydrophilic surfaces. *J. Phys. Chem. C.* **116**, 5730-5734 (2012).
20. Pfeiffer-Laplaud, M., Costa, D., Tielens, F., Gageot, M. & Sulpizi, M. Bimodal acidity at the amorphous silica/water interface. *J. Phys. Chem. C.* **119**, 27354-27362 (2015).
21. Darlington, A. M. & Gibbs-Davis, J. M. Bimodal or trimodal? The influence of starting pH on site identity and distribution at the low salt aqueous/silica interface. *J. Phys. Chem. C.* **119**, 16560-16567 (2015).

22. Dalstein, L., Potapova, E. & Tyrode, E. The elusive silica/water interface: Isolated silanols under water as revealed by vibrational sum frequency spectroscopy. *Phys. Chem. Chem. Phys.* **19**, 10343-10349 (2017).
23. Urashima, S., Myalitsin, A., Nihonyanagi, S. & Tahara, T. The topmost water structure at a charged silica/aqueous interface revealed by heterodyne-detected vibrational sum frequency generation spectroscopy. *J. Phys. Chem. Lett.* **9**, 4109-4114 (2018).
24. Cyran, J. D. et al. Molecular hydrophobicity at a macroscopically hydrophilic surface. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 1520-1525 (2019).
25. Tuladhar, A. et al. Ions tune interfacial water structure and modulate hydrophobic interactions at silica surfaces. *J. Am. Chem. Soc.* **142**, 6991-7000 (2020).
26. Wang, H., Hu, X. & Wang, H. Charge-induced $\chi^{(3)}$ susceptibility in interfacial nonlinear optical spectroscopy beyond the bulk aqueous contributions: The case for silica/water interface. *J. Phys. Chem. C* **125**, 26208-26215 (2021).
27. Lagström, T., Gmür, T. A., Quaroni, L., Goel, A. & Brown, M. A. Surface vibrational structure of colloidal silica and its direct correlation with surface charge density. *Langmuir*. **31**, 3621-3626 (2015).
28. Warring, S. L., Beattie, D. A. & McQuillan, A. J. Surficial siloxane-to-silanol interconversion during room-temperature hydration/dehydration of amorphous silica films observed by ATR-IR and TIR-Raman spectroscopy. *Langmuir*. **32**, 1568-1576 (2016).
29. Shi, Y., Ye, M., Du, L. & Weng, Y. Experimental determination of particle size-dependent surface charge density for silica nanospheres. *J. Phys. Chem. C* **122**, 23764-23771 (2018).
30. Ugliengo, P. et al. Realistic models of hydroxylated amorphous silica surfaces and MCM-41 mesoporous material simulated by large-scale periodic B3LYP calculations. *Adv. Mater.* **20**, 4579-4583 (2008).
31. Sprik, M. Computation of the pK of liquid water using coordination constraints. *Chem. Phys.* **258**, 139-150 (2000).
32. Macias-Romero, C., Nahalka, I., Okur, H. I. & Roke, S. Optical imaging of surface chemistry and dynamics in confinement. *Science*. **357**, 784-788 (2017).
33. Stebbins, J. F. NMR evidence for five-coordinated silicon in a silicate glass at atmospheric pressure. *Nature*. **351**, 638-639 (1991).
34. Farnan, I. & Stebbins, J. F. The nature of the glass transition in a Silica-Rich oxide melt. *Science*. **265**, 1206-1209 (1994).
35. Bykova, E. et al. Metastable silica high pressure polymorphs as structural proxies of deep Earth silicate melts. *Nat. Commun.* **9**, 4789 (2018).
36. Xiao, Y. & Lasaga, A. C. Ab initio quantum mechanical studies of the kinetics and mechanisms of quartz dissolution: OH⁻ catalysis. *Geochim. Cosmochim. Ac.* **60**, 2283-2295 (1996).
37. Trinh, T. T., Jansen, A. P. J. & van Santen, R. A. Mechanism of oligomerization reactions of silica. *J. Phys. Chem. B* **110**, 23099-23106 (2006).
38. Wang, H. et al. Gate-controlled sum-frequency vibrational spectroscopy for probing charged oxide/water interfaces. *J. Phys. Chem. Lett.* **10**, 5943-5948 (2019).
39. Shi, H., Zhang, Y., Wang, H. & Liu, W. Matrix formalism for radiating polarization sheets in multilayer structures of arbitrary composition. *Chin. Opt. Lett.* **15**, 81901 (2017).
40. O'Brien, D. B. & Massari, A. M. Modeling multilayer thin film interference effects in interface-specific coherent nonlinear optical spectroscopies. *J. Opt. Soc. Am. B* **30**, 1503-1512 (2013).
41. Hutter, J., Iannuzzi, M., Schiffmann, F. & VandeVondele, J. CP2K: Atomistic simulations of condensed matter systems. *Wires. Comput. Mol. Sci.* **4**, 15-25 (2014).
42. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic

- behavior. *Phys. Rev. A*. **38**, 3098-3100 (1988).
43. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*. **37**, 785-789 (1988).
 44. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).
 45. VandeVondele, J. et al. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* **167**, 103-128 (2005).
 46. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 14101 (2007).
 47. Nair, N. N., Schreiner, E. & Marx, D. Peptide synthesis in aqueous environments: The role of extreme conditions on amino acid activation. *J. Am. Chem. Soc.* **130**, 14148-14160 (2008).
 48. Munoz-Santiburcio, D. & Marx, D. Chemistry in nanoconfined water. *Chem. Sci.* **8**, 3444-3452 (2017).
 49. Ensing, B., Laio, A., Parrinello, M. & Klein, M. L. A recipe for the computation of the free energy barrier and the lowest free energy path of concerted reactions. *J. Phys. Chem. B*. **109**, 6676-6687 (2005).
 50. Khatib, R. et al. Molecular dynamics simulations of SFG librational modes spectra of water at the water-air interface. *J. Phys. Chem. C*. **120**, 18665-18673 (2016).

Methods

SFVS experiment details

A typical broadband multiplex scheme of SFVS experiment was adopted for our measurements as described elsewhere¹⁷. A regenerative amplifier (Spitfire Ace, Spectra Physics) seeded by a Ti:Sapphire oscillator MaiTai SP, Spectra Physics) was used to produce ~7W of 800 nm, 35 fs pulses at 2 kHz repetition rate. Of the 800 nm beam from the amplifier, 40% passed through a beam-splitter to pump an optical parametric amplifier (TOPAS Prime, Spectra Physics) followed by a difference frequency generation stage with a 350 μm -thick GaSe crystal to produce a broadband infrared beam. The rest of the 800 nm beam transmitted through an interference filter (LL01-808-25, Semrock) to generate a narrowband ($\sim 40\text{ cm}^{-1}$) near-infrared beam. The infrared and near-infrared pulses overlapped at the oxide/liquid interface with incident angles of 60° from the substrate side and 45° from the water side, respectively. The SF output from the water side was collected and analyzed by a spectrograph (Acton SP2,300) and a charge-couple (CCD) camera (Princeton Instruments PyLoN 1,340 \times 400). All measurements were conducted in atmosphere and room temperature ($\sim 22^\circ\text{C}$).

Sample preparation and characterization

The geometry of liquid cell is sketched in the main text and described in details elsewhere³⁸. The 0.5 mm-thick Si(100) wafer was n-doped with a resistivity of $\sim 10\ \Omega\cdot\text{cm}$ and double-sided polished. The optimal film thickness was evaluated using the nonlinear optical matrix formalism^{39,40} (see details in Supplementary Section 1.2). The SiO_2 thin film was deposited through chemical vapor deposition procedure on the silicon wafer. The film thickness was then characterized via both ellipsometry (Horiba UVISEL-2) and UV-VIS reflectance (Ocean

Optics, USB 2,000+ UV-VIS) measurements with good agreement. The film surface was found to be laterally uniform with an average rms roughness about 2~4 nm seen by AFM (Bruker Multimode 8). Before the sample was mounted on the cell, the SiO₂ surface was calcinated in UV (Ultraviolet ozone cleaning machine BZS250GF-TC) for 1 hour, then in nonchromix/H₂SO₄ mixture for ~0.5 hour, and finally rinsed thoroughly with deionized water (18.2 MΩ·cm, Thermo Scientific Barnstead MicroPure UV). Prior to all measurements, the Teflon cell was ultrasonicated successively in acetone, ethanol, and deionized water. The acquired SF spectra as in Fig. 1 hardly changed when different surface regions were probed, confirming the sample surface to be uniform. Moreover, we took a series of spectra from the thin-film-silica/water interface in the O-H stretching vibration range, again found good agreement with those obtained from a bulk-silica/water interface (Extended Data Fig. 4)¹².

Preparation of water solutions

Solutions used in this study were prepared with deionized water (18.2 MΩ·cm) and sodium chloride (NaCl, Sinopharm Chemical Reagent Co., Ltd., 99.5%) as buffer to keep the total ionic strength constant at 10 mM or 0.5 M, respectively. The pH value was adjusted by sodium hydroxide (NaOH, Shanghai Dahe Chemical Reagent Co., Ltd., 96%) or hydrochloric acid (HCl, Sinopharm Chemical Reagent Co., Ltd., 36-38%), with pH value measured by pH/Ion meter (Mettler Toledo, SevenCompact pH/Ion meter S220-B). At the beginning of each round of pH variation, the sample was first exposed to a pH 2 solution and allowed to equilibrate with the solution for more than 30 min. Upon changing pH, we removed the old solution and rinsed the cell with new solution 2~3 times. The sample was then let to equilibrate with the new solution for a few minutes before taking the spectra.

DFT-MD and DFT-MD metadynamics set up

All the density functional theory molecular dynamics (DFT-MD) and DFT-MD metadynamics simulations were carried out using the CP2K software package⁴¹ at the Becke-Lee-Yang-Parr functional (BLYP)^{42,43} level of representation that includes the Grimme D2⁴⁴ correction for van der Waals interactions and a combination of Goedecker Teter Hutter (GTH) pseudo-potentials⁴⁴ and Gaussian Plane Waves basis sets⁴⁵. The DVZP-MOLOPT-SR ('SR' denotes shorter range) basis set, augmented with a 400 Ry plane wave basis set was used. The nuclei displacements were predicted from the classical Newton's equations of motions integrated through the velocity Verlet algorithm.

The DFT-MD simulations employed for the calculations of the Vibrational Density of States (VDOS) spectra in the main text were equilibrated for 10 ps in the NVT ensemble (canonical ensemble) with a Canonical Sampling through Velocity Rescaling (CSVR) thermostat⁴⁶ at the target temperature of 300 K and then carried on for 50 ps in the pure NVE ensemble. Calculation of the VDOS was performed over these 50 ps. The rest of the DFT-MD simulations and the DFT-MD metadynamics were done in the NVT ensemble with a CSVR thermostat and a target temperature of 300 K.

We simulated 3 different models of the silica/water interface characterized by different morphologies and degree of hydroxylation. The 4.5 OH/nm² amorphous silica/water interface was modelled in a box of dimension 13.4 Å × 13.3 Å × 35 Å. There were 8 Si-OH groups at the silica surface and the liquid water was modelled with 120 water molecules. For the 5.5 and 7.6

OH/nm² amorphous silica-water interfaces, boxes of dimension 9.117 Å × 16.342 Å × 32 Å were used. The surfaces were respectively composed of 8 and 12 SiOH groups and in both cases the liquid water was modelled with 115 water molecules.

Neutral boxes were employed for all the DFT-MD metadynamics simulations done in this work to calculate the silanol pKa values. The total simulation time amounted to 360 ps for the DFT-MD simulations and to 270 ps for the metadynamics simulations.

Details of the DFT-MD metadynamics simulations

The free energy profiles for the SiOH acid-base equilibrium at the aqueous silica surface were obtained by DFT-MD metadynamics simulations as implemented in the CP2K software package. The acid-base behavior (i.e. protonation/deprotonation of the surface sites) of each silanol over the 4.5 OH/nm² silica surface was simulated and characterized. Eight metadynamics were performed, one per silanol present at the surface.

Our computational approach for the DFT-MD metadynamics was similar to the one successfully adopted in previous works^{47,48} for the study of reactions at aqueous solid interfaces. The type of reaction coordinate $c(A-B)$ adopted was the coordination number³¹ of an atom A with respect to a set of atoms B, defined as:

$$c(A - B) = \sum_{I \in B} \frac{1 - (R_{AI}/R_{AB}^0)^6}{1 - (R_{AI}/R_{AB}^0)^{12}} \quad (1)$$

where R_{AI} is the distance between atoms A and I (belonging to the chosen set of B atoms). R_{AB}^0 is a fixed cut-off parameter based on the equilibrium bond distance between A and B atoms. In particular, we have adopted the following two reaction coordinates:

(1) The silanol oxygen coordination number with respect to all the hydrogen atoms of the system ($CN_{\text{SiO--H}}$), values of 1 and 0.5 identify SiOH and SiO^- species respectively. This coordinate allowed us to model silanol deprotonation events.

(2) The silanol oxygen coordination number with respect to all the silicon atoms of silica ($CN_{\text{SiO--Si}}$). This allowed us to sample possible surface reconstruction processes.

The height of the Gaussian hills added during biased metadynamics was 0.8 kJ/mol. The width of the Gaussian hills function (δ_s) was 0.05. To avoid “hill-surfing”, as discussed in Refs. 47-49, a Gaussian hill was added every time the dynamics explored a spot on the reaction coordinate space at a distance $3\delta_s/2$ from the spot where the previous Gaussian hill was deposited:

$$|s(t) - s(t_i)| = 3\delta_s/2 \quad (2)$$

where $s(t)$ is the position along the reaction coordinate at a given time t , and t_i is the time when the last Gaussian hill was deposited.

The computed free energy was considered upon convergence after crossing the energetic barrier forward and backward along the same reaction coordinate. The reaction pathways were found along the minimum energy paths connecting an original silanol and its deprotonation product. The pKa value of a given silanol was calculated from the free energy difference (ΔG) between the energy minima of a given individual silanol species (SiOH) and the conjugate base [SiO^- or $\text{Si}(5c)$] (no integration of the metadynamics values over the well regions was performed):

$$\text{pKa} = \Delta G/2.303RT \quad (3)$$

where T is the temperature of the system in Kelvin (imposed by the NVT ensemble) and R ($8.314 \text{ JK}^{-1}\text{mol}^{-1}$) is the ideal gas constant. To check the robustness of our results, the metadynamics for pKa calculations were repeated for two acidic and two basic silanol groups adopting a different criterion for the rate deposition of the Gaussian hills, i.e. a Gaussian hill was added every 100 steps of the dynamics. Similar results were obtained, with pKa values within ± 0.4 pKa units from the original ones. The pKa values obtained then allowed deduction of surface coverages of SiOH, Si(5c), and SiO⁻ species versus pH.

Theoretical calculation of SFVS as a function of pH

The DFT-MD SFG spectrum was calculated from $|\chi_{total}^{(2)}(\omega, pH)|^2$, with $\chi_{total}^{(2)}(\omega, pH)$ being the total SFG surface nonlinear susceptibility given by

$$\chi_{total}^{(2)}(\omega, pH) = \sum_j C_j(pH) \bar{\chi}_j^{(2)}(\omega), \quad (4)$$

where $\bar{\chi}_j^{(2)}(\omega)$ is the surface nonlinear susceptibility of the j^{th} surface species at full coverage including Si-OH, Si(5c)-OH⁻, and Si-O⁻, and $C_j(pH)$ is the fractional surface coverage of the j^{th} surface species at a given (bulk) pH (see Supplementary Section 3.1). To calculate $\bar{\chi}_{j,xxz}^{(2)}(\omega)$ (experimentally probed by SFG with SSP polarization combination), we adopted the same methodology previously developed for surface $\chi^{(2)}(\omega)$ spectra of both water and SiO-H in the O-H stretch range⁵⁰, which had shown excellent agreement with experiments^{24,25}. We used the following expression in the calculation:

$$\bar{\chi}_{j,xxz}^{(2)}(\omega) = \frac{i}{k_B T \omega} \int_0^\infty dt e^{-i\omega t} \langle \dot{\alpha}_{j,xx}(t) \dot{\mu}_{j,z}(0) \rangle, \quad (5)$$

where $\mu_z(0)$ and $\alpha_{xx}(t)$ refer to the z and xx components of the dipole and polarizability of the surface-group in the laboratory frame, and the angular brackets denote an average of the product.

The calculated $|\bar{\chi}_{j,xxz}^{(2)}(\omega)|$ spectra for the three surface species, Si-OH, Si(5c)-OH⁻, and Si-O⁻, are shown in Extended Data Fig. 3; they peaked at ~840, 920, and 1,000 cm⁻¹, respectively, and their profiles hardly change with pH. The total surface nonlinear susceptibility is given by:

$$\begin{aligned} \chi_{ssp}^{(2)}(\omega, pH) = & \left(C_{SiOH}(pH) - C_{Si(5c)^-}(pH) \right) \bar{\chi}_{SiOH}^{(2)}(\omega) + C_{Si(5c)^-}(pH) \bar{\chi}_{Si(5c)^-}^{(2)}(\omega) \\ & + C_{SiO^-}(pH) \bar{\chi}_{SiO^-}^{(2)}(\omega), \end{aligned} \quad (6)$$

from which the $|\chi_{ssp}^{(2)}(\omega, pH)|^2$ spectra can be found as presented in Figs. 4b and 4c of the main text in comparison with the experimentally observed spectra.

Data availability

All data supporting the findings of this work are available within the paper and its Supplementary Information. Source data can be found at <https://zenodo.org/>. Source data are provided with this paper. All density functional theory calculations were performed using CP2K software package, which is commercially available at <https://www.cp2k.org/download>. The custom codes used to generate the data that support the findings of this study are available from the corresponding authors upon reasonable request.

Method References

38. Wang, H. et al. Gate-controlled sum-frequency vibrational spectroscopy for probing charged oxide/water interfaces. *J. Phys. Chem. Lett.* **10**, 5943-5948 (2019).
39. Shi, H., Zhang, Y., Wang, H. & Liu, W. Matrix formalism for radiating polarization sheets in multilayer structures of arbitrary composition. *Chin. Opt. Lett.* **15**, 81901 (2017).
40. O'Brien, D. B. & Massari, A. M. Modeling multilayer thin film interference effects in interface-specific coherent nonlinear optical spectroscopies. *J. Opt. Soc. Am. B.* **30**, 1503-1512 (2013).
41. Hutter, J., Iannuzzi, M., Schiffmann, F. & VandeVondele, J. CP2K: Atomistic simulations of condensed matter systems. *Wires. Comput. Mol. Sci.* **4**, 15-25 (2014).
42. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A.* **38**, 3098-3100 (1988).

43. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*. **37**, 785-789 (1988).
44. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).
45. VandeVondele, J. et al. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* **167**, 103-128 (2005).
46. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 14101 (2007).
47. Nair, N. N., Schreiner, E. & Marx, D. Peptide synthesis in aqueous environments: The role of extreme conditions on amino acid activation. *J. Am. Chem. Soc.* **130**, 14148-14160 (2008).
48. Munoz-Santiburcio, D. & Marx, D. Chemistry in nanoconfined water. *Chem. Sci.* **8**, 3444-3452 (2017).
49. Ensing, B., Laio, A., Parrinello, M. & Klein, M. L. A recipe for the computation of the free energy barrier and the lowest free energy path of concerted reactions. *J. Phys. Chem. B*. **109**, 6676-6687 (2005).
50. Khatib, R. et al. Molecular dynamics simulations of SFG librational modes spectra of water at the water-air interface. *J. Phys. Chem. C*. **120**, 18665-18673 (2016).

Acknowledgments We thank Daria Ruth Galimberti and Fabio Pietrucci for advice in the metadynamics simulations. This work was supported by National Natural Science Foundation of China and National Key Research and Development Program of China (12250002, 2022YFA1204702, 2019YFA0308404, 11991062); Science and Technology Commission of Shanghai Municipality (22XD1400200, 20JC1415900); National Program for Support of Top-Notch Young Professionals (W03070062); Shu Guang Project by Shanghai Education Development Foundation (18SG02). The sample fabrication was conducted at Fudan Nanofabrication Lab. HPC resources from GENCI-France Grant No. 072484 (CINES/IDRIS/TGCC) are acknowledged. S. P. acknowledges financial support by European Research Council (ERC) Advanced Grant 695437 THz-Calorimetry and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy-EXC2033-390677874-RESOLV for funding this study.

Author contributions W.-T. L. and Y. R. S. conceived the idea and designed the research. X. L. and X. L. performed the sample preparation, SFVS experiments, and analyzed the data. H. C., Y.

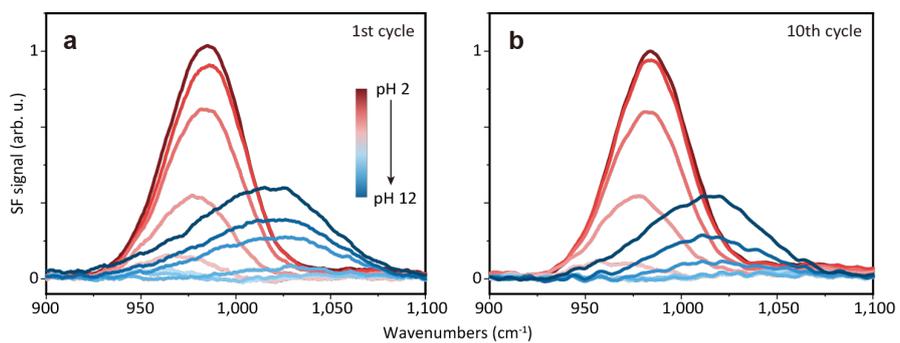
L., and H. L. assisted with ellipsometry, UV-VIS, and AFM characterizations of samples. M.-P. G., F. S. B., S. P., and W. C. developed theoretical approaches, performed simulations, and analyzed the theoretical results. All authors were involved in the discussion and interpretation of the experimental and theoretical results. W.-T. L., M.-P. G, and Y. R. S. wrote the paper with inputs from all authors.

Competing interests The authors declare no competing interests.

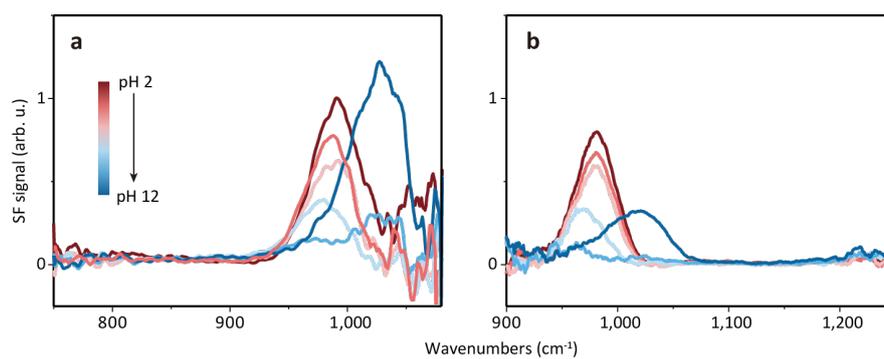
Additional information

Supplementary information is available for this paper.

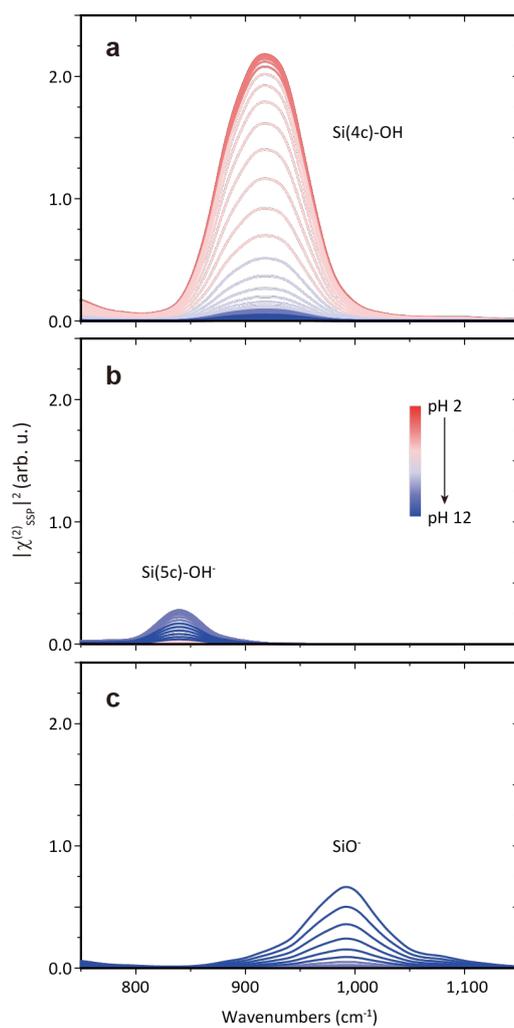
Correspondence and requests for materials should be addressed to Wei-Tao Liu or Marie-Pierre Gaigeot.



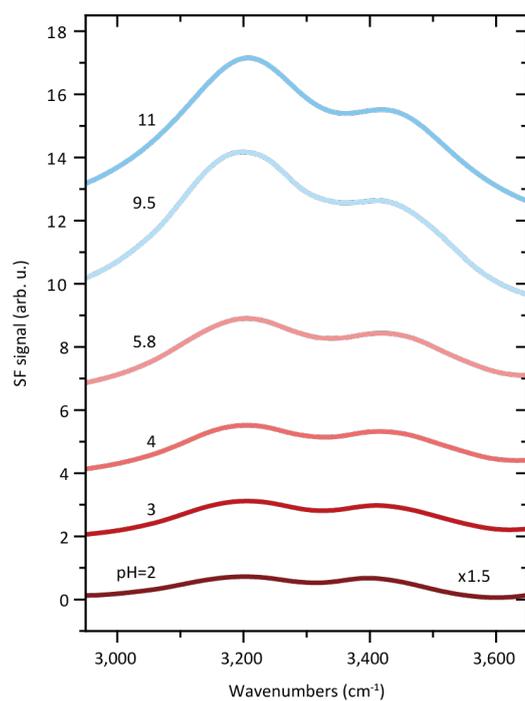
Extended Data Figure 1 |SF spectra of the thin-film-SiO₂/H₂O interface in the Si-O stretch range taken at **(a)** the first round and **(b)** the tenth round of pH variation.



Extended Data Figure 2 | SF spectra from the thin-film-SiO₂/H₂O interface in the spectra range of **(a)** 750-1,080 cm⁻¹ and **(b)** 900-1,250 cm⁻¹.



Extended Data Figure 3 | Calculated $|\chi_{SSP}^{(2)}|^2$ spectra of individual Si(4c)OH (top), Si(5c)OH⁻ (middle), and SiO⁻ (bottom) groups at different pH values and an ion concentration of 10 mM.



Extended Data Figure 4 | SF vibrational spectra in the OH stretching range of the thin-film-SiO₂/H₂O interface at various pH (no buffer ion added).

Supplementary Information of

Unveiling structural evolution and reconstruction of oxide surface in water

Xiaoqun Li^{1,6}, Flavio S. Brigiano^{2,3,6}, Simone Pezzotti^{2,4,6}, Xinyi Liu^{1,6}, Wanlin Chen^{2,6}, Huiling Chen¹, Ying Li¹, Hui Li¹, Yue Ron Shen^{1,5}✉, Marie-Pierre Gaigeot²✉, Wei-Tao Liu¹✉

¹Physics Department, State Key Laboratory of Surface Physics, Key Laboratory of Micro and Nano Photonic Structures [Ministry of Education (MOE)], Fudan University, Shanghai 200433, China.

²Université Paris-Saclay, Univ Evry, CNRS, LAMBE UMR8587, 91025, Evry-Courcouronnes, France,

³Current address: Laboratoire de Chimie Théorique, Sorbonne Université, UMR 7616 CNRS, 4 Place Jussieu, 75005 Paris, France,

⁴Current address: Physical Chemistry II, Faculty of Chemistry and Biochemistry, Ruhr University Bochum, Universitätsstr. 150, 44780 Bochum, Germany,

⁵Department of Physics, University of California at Berkeley, Berkeley, CA 94702, USA.

⁶These authors contributed equally: Xiaoqun Li, Flavio S. Brigiano, Simone Pezzotti, Xinyi Liu, Wanlin Chen

✉Corresponding author. Email: yrshen@berkeley.edu, mgaigeot@univ-evry.fr, wliu@fudan.edu.cn

Table of content

- S1. Basic principle of SFVS and the nonlinear matrix formalism**
- S2. Properties of Si(5c) species at the silica/water interface**
- S3. Theoretical surface coverages and nonlinear optical susceptibilities**
- S4. Other characterizations of the Si(5c) species**

S1. Basic principle of SFVS and the nonlinear matrix formalism

S1.1 Basic principle of SFVS

The basic principle of SFG is described elsewhere¹. Briefly, when the IR frequency (ω_{IR}) is near vibrational resonances, the SF signal (S_{SF}) generated by the incident beams is:

$$S_{SF} \propto \left| \chi_{eff}^{(2)} \right|^2 = \left| \chi_{NR} + \chi_R \right|^2 = \left| \chi_{NR} + \sum_q \frac{A_q}{\omega_{IR} - \omega_q + i\Gamma_q} \right|^2, \quad (S1)$$

where χ_{NR} and χ_R are the non-resonant and resonant contributions, with A_q , ω_q , and Γ_q being the amplitude, frequency, and damping coefficient of the q^{th} resonance mode, respectively. With inhomogeneous broadening taken into account, the SF output of Equation (S1) becomes²:

$$S_{SF} \propto \left| \chi_{NR} + \sum_q \int_{-\infty}^{\infty} \frac{A_q}{\omega_{IR} - \omega' + i\Gamma_q} \frac{1}{\sqrt{2\pi}\sigma_q} \exp\left(-\frac{(\omega' - \omega_q)^2}{2\sigma_q^2}\right) d\omega' \right|^2,$$

where σ_q is the inhomogeneous bandwidth². Figures S1a and S1b show fitting of the experimental spectra taken from the SiO₂-film/water interface at different pH values. The fitting values of Γ_q and σ_q were ~ 43 and 27 cm^{-1} for the Si-OH band, and ~ 45 and 28 cm^{-1} for the Si-O⁻ band. The resonant amplitudes of the two bands deduced from fitting were plotted versus pH in Figs. S1c and S1d. Dashed and dotted curves in Fig. S1d are theoretical curves calculated from the Gouy-Chapman-Stern (GCS) model for surface reaction of $\text{SiOH} \leftrightarrow \text{SiO}^- + \text{H}^+$ with a single pKa value at $\sim 7^3$.

For the SSP (S-SF, S-NIR, P-IR), PPP (P-SF, P-NIR, P-IR), and SPS (S-SF, P-NIR, S-IR) beam polarization combinations, the effective SF amplitudes for an anisotropic interface are⁴:

$$\begin{aligned} A_{SSP} &= \sin \beta_{IR} L_{yy}(\omega_{SF}) L_{yy}(\omega_{NIR}) L_{zz}(\omega_{IR}) A_{yyz}, \\ A_{PPP} &\approx -\cos \beta_{SF} \cos \beta_{NIR} \sin \beta_{IR} L_{xx}(\omega_{SF}) L_{xx}(\omega_{NIR}) L_{zz}(\omega_{IR}) A_{xxz} \\ &\quad + \sin \beta_{SF} \sin \beta_{NIR} \sin \beta_{IR} L_{zz}(\omega_{SF}) L_{zz}(\omega_{NIR}) L_{zz}(\omega_{IR}) A_{zzz}, \\ A_{SPS} &= \sin \beta_{NIR} L_{yy}(\omega_{SF}) L_{zz}(\omega_{NIR}) L_{yy}(\omega_{IR}) A_{yzy}, \end{aligned}$$

with β_n and $L_{ii}(\omega_n)$ being the beam incident angle and Fresnel coefficients, and $\{x, y, z\}$ referring to the lab coordinates with z along the surface normal and x in the incident plane. For linear bonds on an azimuthally isotropic surface, the ratio between A_{PPP} and A_{SSP} are known to be sensitive to the bond polar orientation (as in Ref. 4 of the main text). We deduce from the A_{PPP}/A_{SSP} that both Si-OH and Si-O⁻ bonds remain at $\sim 20 \pm 10^\circ$ away from the surface normal for all pH values. This was further supported by the observation of very weak SPS spectra, which signified the bonds had only small in-plane component (Fig. S1e).

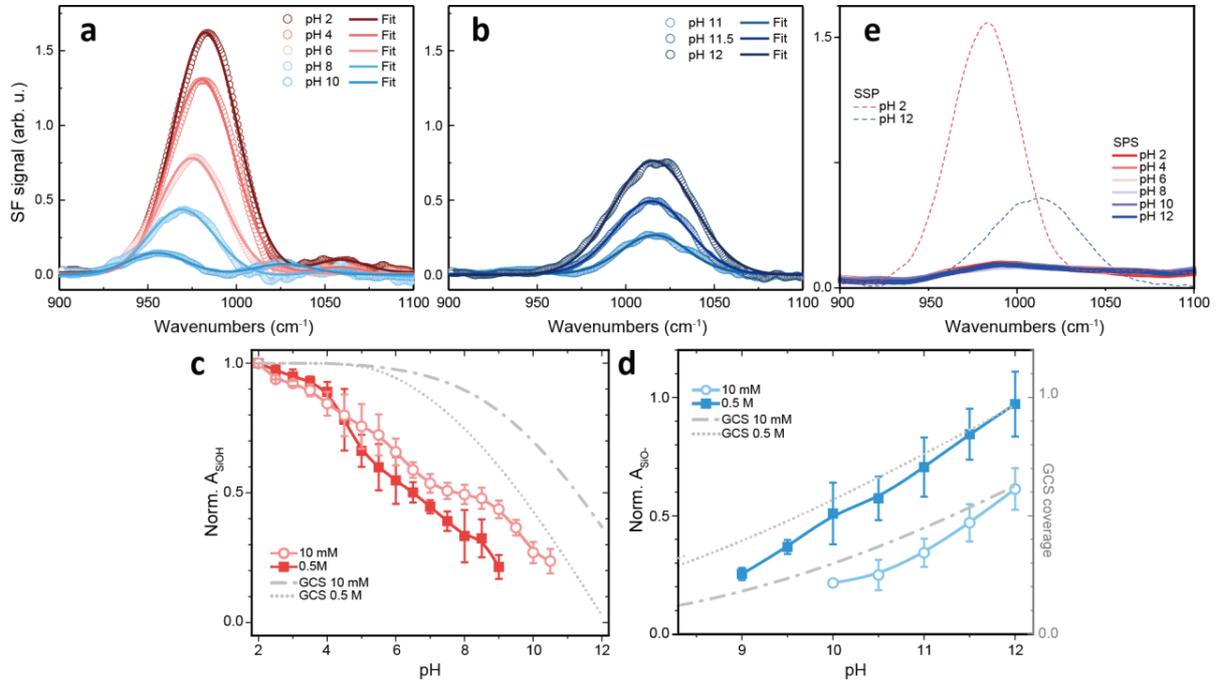


Figure S1 | (a) and (b) Fits of SF spectra vs. pH. (c) and (d) Amplitudes of Si-OH and Si-O⁻ modes deduced for the two ion concentrations, 10 mM (empty symbols) and 0.5 M (filled symbols). The dashed curves are calculated from the single-pKa GCS model of corresponding ion concentrations based on Ref. 3. (e) SPS spectra at different pH values (solid curves) in comparison to SSP spectra at pH 2 and 12 (dotted curves).

S1.2 Nonlinear matrix formalism for local field calculation

The matrix formalism for computing local field strength with nonlinear optical boundary conditions we have developed is described elsewhere⁵. The boundary conditions with the presence

of an interfacial polarization sheet $\mathbf{P}_s(\omega)$ are (x parallel to the incident plane, and z along the surface normal):

$$\Delta E_x = \sigma_z = -\frac{4\pi}{\epsilon'} ik_x P_{sz}, \quad \Delta E_y = 0,$$

$$\Delta H_x = \sigma_y = -\frac{4\pi}{c} i\omega P_{sy}, \quad \Delta H_y = \sigma_x = -\frac{4\pi}{c} i\omega P_{sx}.$$

For S-polarized beam (TE wave), we have $\sigma_E = 0, \sigma_H = \sigma_y$; and for P-polarized beam (TM wave), we have $\sigma_E = \sigma_z, \sigma_H = \sigma_x$. In an n -layer system, the nonlinear characteristic matrix formalism of the J^{th} interface between j^{th} and $(j+1)^{\text{th}}$ media can then be written as (capital I, J, N are indices of interfaces, i, j, n are indices of media):

$$\begin{bmatrix} E_{I,1} \\ H_{I,1} \end{bmatrix} = M_2 M_3 \dots M_j \begin{bmatrix} E_{J,j} \\ H_{J,j} \end{bmatrix},$$

$$\begin{bmatrix} E_{J,j} + \sigma_{J,E} \\ H_{J,j} + \sigma_{J,H} \end{bmatrix} = M_{j+1} M_{j+2} \dots M_n \begin{bmatrix} E_{N,n+1} \\ H_{N,n+1} \end{bmatrix}.$$

M_i is the characteristic matrix of medium with the form of:

$$M_i = \begin{bmatrix} \cos \delta_i & -\frac{i \sin \delta_i}{n_{i,eff}} \\ -in_{i,eff} \sin \delta_i & \cos \delta_i \end{bmatrix},$$

where $n_{i,eff} = n_i \cos \theta_i$ for TE waves, and $n_{i,eff} = n_i / \cos \theta_i$ for TM waves, with n_i being the complex refractive index of, θ_i the beam refraction angle in, and $\delta_i = n_i k_i \cos \theta_i d_i$ the wave propagation phase through the i^{th} section of the medium with thickness d_i . Details about the calculation procedure were provided in Ref. 5. The squared products of all local field factors at the Si/SiO₂ and SiO₂/H₂O interfaces, respectively, of a Si/SiO₂/H₂O junction as functions of the SiO₂ film thickness are presented in Fig. 1b of the main text, with reference to the beam geometry in Fig. 1a. The estimated SF intensity in the Si-O stretch range obtained from the maximum local field intensity at the SiO₂/H₂O interface for the case of ~150 nm-thick SiO₂ film in Fig. 1b was about three times of that from an IR-transparent bulk-SiO₂/H₂O interface in the OH stretch range.

With SF signals in the latter case being routinely observed, the Si-O stretch signals from the thin-film water interface should be readily detectable.

S2. Properties of Si(5c) species at the silica/water interface

S2.1 Stability of Si(5c) species versus pH: theoretical calculations

As mentioned in the main text, two kinds of Si(5c) species were formed upon deprotonation of the more acidic SiOH at the amorphous silica surface, i.e., Si(5c)OH⁻ (66%) and SiO₅⁻ (33%). We first evaluated the stability of the more abundant Si(5c)OH⁻ species. We considered a single Si(5c)OH⁻ on an aqueous 4.5 OH/nm² silica surface and simulated its stability in response to changes of pH using DFT-MD metadynamics simulations. The breakage of the Si(5c)OH⁻ was expected to lead to the formation of two Si(4c), as schematically presented in Fig. S2a. In the metadynamic simulations, pH changed the protonation states of the two oxygen atoms involved in the formation and breakage of Si(5c)OH⁻. Of the two oxygen atoms, the one belonging to the basic (or high-pK_a) silanol group hosting the Si(5c)OH⁻ is labelled O1 (marked in blue in Fig. S2a), and the one belonging to the acidic (or low-pK_a) silanol group that yielded Si(5c)] is labelled O2 (marked in green, Fig. S2a). The first reaction coordinate was chosen as the difference between the coordination number of O2 and O1 with respect to all hydrogen atoms in the system (C_{OH2-OH1}), which described the protonation state of the two silanols and the effective pH in the following way: C_{OH2-OH1} = 0 - 1 = -1 if only the acidic silanol (O2) was deprotonated (between pH 4~9); 0 if both silanols were protonated (C_{OH2-OH1} = 1 - 1 = 0) (corresponding to pH < 4) or deprotonated (C_{OH2-OH1} = 0 - 0 = 0) (pK > 9); C_{OH2-OH1} = 1 - 0 = +1 if only the basic silanol site (O1) was deprotonated. The second reaction coordinate was chosen as the coordination number of O2 with respect to all silicon atoms in the system (C_{Si-O...Si}, green arrow in Fig. S2a), to sample the presence or absence of Si(5c)OH⁻. This coordinate would take the value of 1.6 when the Si(5c)OH⁻ species was stable, and a value of 0.9 when Si(5c)OH⁻ broke into two Si(4c) species. With the above two reaction coordinates, we could hence probe the stability of the molecular species [Si(5c)OH⁻ or Si(4c)] residing at the surface over different pH ranges.

Figure S2b presents the free energy landscape with different reactive species attached to the respective minima explored by the metadynamics simulation. Minimum 1 was found near $C_{\text{OH2-OH1}} = -1$ and $C_{\text{Si-O...Si}} = 1.6$, corresponding to the case where the low-pKa Si(O2)H was deprotonated, while the high-pKa Si(O1)H group remained protonated and hosted a stable Si(5c); it appeared in the pH range of 4~9. Outside that pH range, when $C_{\text{OH2-OH1}} \neq -1$, there was no minimum related to Si(5c)OH⁻ observed. Instead, minima 2 and 3 were observed for $C_{\text{Si-O...Si}} \sim 0.9$, corresponding to the breakdown of Si(5c)OH⁻ into two Si(4c) when either the high-pKa silanol group is deprotonated (which happens at pH > 9) or both silanols are protonated (pH < 4).

To further confirm the results, we ran two additional *unbiased* DFT-MD simulations (i.e. standard DFT-MD simulations, not combined with enhanced sampling methods) on one single deprotonated Si(5c) species [Si(5c)OH⁻ → Si(5c)O⁻, marked by A in Fig. S2b] at the aqueous 4.5 OH/nm² silica surface. As seen from the absence of a corresponding minimum ($C_{\text{OH2-OH1}} = 0$ and $C_{\text{Si-O...Si}} = 1.6$), the deprotonated Si(5c) systematically breaks into two SiO⁻ groups after few hundreds of fs, which then remain stable for the entire time duration of the DFT-MD simulations (30 ps). This was due to negative charge increase towards Si(5c) after deprotonation of the already negatively charged Si(5c)OH⁻ at high pH (average pKa = 7.7) that strongly reducing its electrophilic character, and inducing a cleavage of the O-Si(5c) bond with disappearance of the Si(5c)OH⁻ species.

Finally, we looked into the stability of the other Si(5c) species, i.e. the SiO₅⁻, at high pH (pH > 9). We simulated by *unbiased* DFT-MD the behavior of SiO₅⁻ at the aqueous 4.5 OH/nm² silica surface with all low-pKa and 60 % of the high-pKa silanol groups already deprotonated. The SiO₅⁻ was found to collapse after few ps of dynamics leading to the formation of one stable SiO⁻ group.

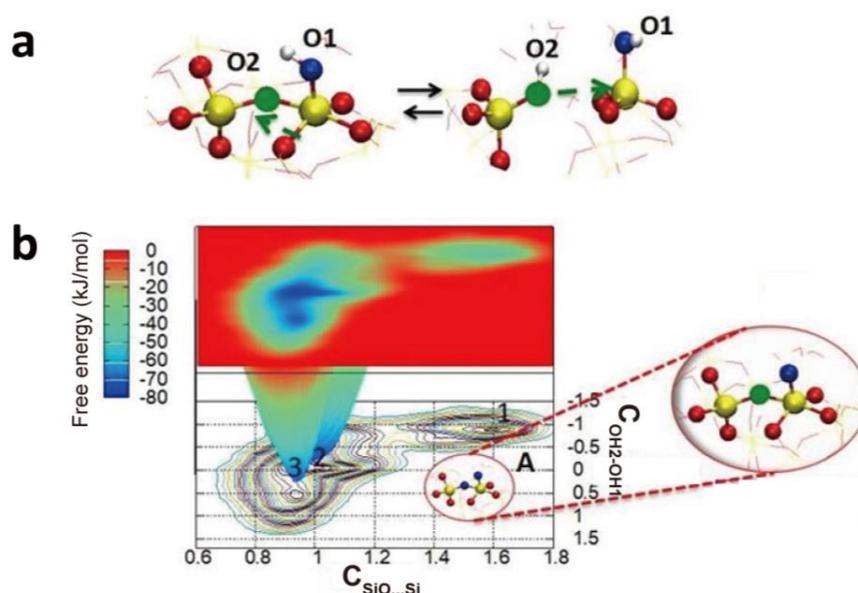


Figure S2 | (a) Scheme of the reaction studied by DFT-MD biased metadynamics simulations. The oxygen atom of the basic (high-pKa) silanol group hosting the Si(5c)OH⁻ is labelled O1 and marked in blue; the oxygen atom of the acidic (low-pKa) silanol group is labelled O2 (green). A new Si-O bond is created during Si(5c) formation (from right to left), while the same bond has to be cleaved to convert Si(5c) back to two Si(4c) species (from left to right). (b) The free energy landscape and associated contour map spanned by C_{OH₂-OH₁} (horizontal axis) and C_{Si-O...Si} (vertical axis). Levels in the contour map correspond to increase of 5 kJ/mol. Minimum 1 indicates Si(5c)OH⁻ to be stable at the surface in the pH range 4-9, whereas minima 2 and 3 represent two stable Si(4c) species for pH > 9 and pH < 4, respectively. For a reaction with the Si(5c)OH⁻ further deprotonated to SiO₅²⁻ (marked by capital A), there was no corresponding free energy minimum, meaning the product would not be stable.

S2.2 Appearance of Si(5c) species in different models of hydroxylated silica surface

To verify if the results were model-independent, two additional silica surface models with varying degrees of hydroxylation were investigated: 5.5 and 7.6 OH/nm² amorphous silica model surfaces^{6,7} in contact with liquid water. Note that the two model surfaces were prepared in different ways. The 5.5 OH/nm² model from Ugliengo et al⁷ was derived from a bulk cristobalite heated to very high temperature and annealed through classical molecular dynamics simulations. The surface was then prepared by cutting this bulk and passivated by silanols. The 7.6 OH/nm² surface adopting the Bernasconi model⁶ was obtained by cutting a simulated bulk silica model at high temperature by classical MD simulations and annealed by ab-initio MD simulations.

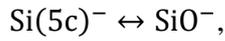
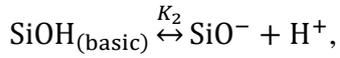
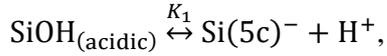
We first identified the silanol groups that had the favorable structural properties (in terms of orientations and proximity) for Si(5c) formation based on the results presented in S4 and main text, and then characterized their chemical behavior after deprotonation by unbiased DFT-MD simulations. It was found that once the silanols were deprotonated, the resultant SiO⁻ indeed approached the neighboring SiO₄ tetrahedron and did the expected nucleophilic attack, leading to the formation of Si(5c). Our results show that when in contact with water in the 4~9 pH range, Si(5c) is systematically and preferentially formed by SiOH groups that satisfy certain orientation and proximity criteria, regardless of the degree of hydroxylation and the way the amorphous surface is prepared in the simulation^{6,7}.

S3. Theoretical surface coverages and nonlinear optical susceptibilities

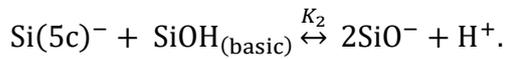
S3.1 Theoretical calculation of the surface coverage as a function of pH

We discuss here how we can find the surface coverages of the three surface species: silanols [Si(4c)OH], silanolates (SiO⁻), and Si(5c)⁻ [including Si(5c)OH⁻ and SiO₅⁻], as functions of bulk pH in water. The DFT-MD simulations yielded that ~ 40% (denoted as *a*) of the initial silanol groups were acidic with pKa of 3.9±0.6 (denoted as pK₁) and the rest basic ones with pKa of 7.7±0.9 (denoted as pK₂). We note that the pKa values here were calculated with respect to the surface pH (pH_S, referring to the H⁺ concentration at an interface). They will be translated into bulk pH in the later section.

At low and high pHs, the three surface species are related to one another through the following reactions



as illustrated in Figs. S3a and S3b, respectively, where H⁺ is released into the solvent. Based on our DFT-MD results presented in the previous section, we found that breaking of Si(5c)⁻ and the deprotonation of the basic SiOH have the same pH dependence, we therefore can combine the two reactions into one:



From the Henderson-Hasselbalch equations, we have the following relations between surface coverage of SiOH, Si(5c)⁻ and SiO⁻ (C_{SiOH} , $C_{\text{Si(5c)}^-}$, and C_{SiO^-} , with 1 referring to a full coverage of SiOH not being deprotonated) and the surface pH (pH_S):

$$pH_S = pK_1 + \log_{10} \left(\frac{C_{\text{Si(5c)}^-}}{C_{\text{SiOH}_{\text{acidic}}}} \right),$$

$$pH_S = pK_2 + \log_{10} \left(\frac{(C_{SiO^-})^2}{C_{SiOH_{basic}} C_{Si(5c)^-}} \right).$$

from which we find:

$$C_{Si(5c)^-}(pH_S) = C_{Si(5c)^-}^{produced} - C_{Si(5c)^-}^{consumed} = a \frac{10^{pH_S - pK_1}}{1 + 10^{pH_S - pK_1}} \left(1 - \frac{10^{pH_S - pK_2}}{1 + 10^{pH_S - pK_2}} \right), \quad (S2)$$

$$C_{SiO^-}(pH_S) = (1 - a) \frac{10^{pH_S - pK_2}}{1 + 10^{pH_S - pK_2}} + a \frac{10^{2pH_S - pK_1 - pK_2}}{(1 + 10^{pH_S - pK_1})(1 + 10^{pH_S - pK_2})}, \quad (S3)$$

$$C_{SiOH}(pH_S) = 1 - C_{SiO^-}(pH_S) - C_{Si(5c)^-}(pH_S). \quad (S4)$$

where a is the percentage of the initial acidic silanol groups (40% of silanols).

Because of inhomogeneity of the surface structure, the values of pK_1 and pK_2 (with respect to the surface pH) in our DFT-MD simulations appear as distributions that can be approximated by Gaussian functions centered at 3.9 and 7.7, respectively, with corresponding standard deviations of 0.6 for pK_1 and 0.9 for pK_2 .

The surface pH (pH_S) and bulk pH, or the surface and bulk H^+ concentrations, $n_S(H^+)$ and $n_B(H^+)$, are related by the surface potential ϕ created by the surface charge density σ given by $\sigma(pH) = (C_{Si(5c)^-} + C_{SiO^-}) \rho_{SiOH} e$, where ρ_{SiOH} is the surface density of SiOH for a neutral interface at low pH and e is the unit electrical charge. From the modified Poisson-Boltzmann equation^{8,9}, we have:

$$n_B(H^+) = \frac{n_{Surface}(H^+) \left(1 + 2v \sinh^2 \left(\frac{e\phi}{2k_B T} \right) \right)}{\exp \left(-\frac{e\phi}{k_B T} \right)},$$

$$\phi_0 = \phi(z = 0) = -\frac{2k_B T}{e} \sinh^{-1} \left\{ \sqrt{\frac{1}{2v} \left[\exp \left(\frac{\sigma^2 v}{4n_0 \epsilon \epsilon_0 k_B T} \right) - 1 \right]} \right\}, \quad (S5)$$

where ϕ is the electric potential in function of the distance in z -direction from the surface, σ , n_0 , k_B , T , e , ϵ , and ϵ_0 are the surface charge density, the ion concentration in the bulk, the Boltzmann constant, the temperature in Kelvin, the elementary charge, the relative dielectric constant of liquid water and the dielectric constant of vacuum. The parameter v is expressed as $v = 2a^3 n_0$, where

we adopt $a = 0.7$ nm as the Bjerrum length to account for the ion size in water at room temperature^{8,9}.

The calculated surface pHs and bulk pH for our case of two different ion concentrations, 10 mM and 0.5 M (see experimental conditions), are plotted in Fig. S3c. Accordingly, the values of pK_1 and pK_2 with respect to bulk pH are changed to 6.7 ± 0.6 and 11.6 ± 0.9 at 10 mM ion concentration, and $pK_1 = 4.7 \pm 0.6$ and $pK_2 = 9.8 \pm 0.9$ at 0.5 M ion concentration, as denoted by the dotted lines in Fig. S3c. From Equation (S5), we can thus translate the theoretical pKa values (calculated with respect to surface pH) to bulk pH units, as illustrated in Fig. S3c. Taking these bulk pKa values and adopting the same method discussed in the previous section, we can then convert pH_s and find the coverages of different species as functions of bulk pH, as presented in Fig. 4a of the main text for the case of 10 mM ion concentration.

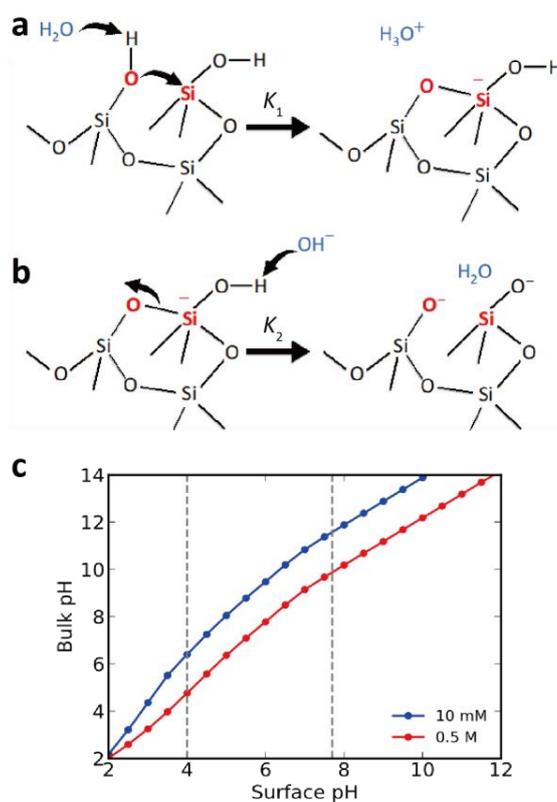


Figure S3 | (a) Formation of Si(5c)⁻ species. (b) Breaking of Si(5c)⁻ species. (c) Calculated bulk pH versus surface pH values.

S3.2 Consistency between the discovery of Si(5c)- species & SHG titration experiments

According to Ref. 10, 11 the effective surface nonlinear susceptibility for the reflected second harmonic generation (SHG) at 2ω has the expression

$$\chi_{eff}^{(2)}(2\omega) = \chi_{BIL}^{(2)} + \chi_{DL}^{(2)}, \quad (S6)$$

$$\chi_{DL}^{(2)} = \chi_{Bulk}^{(3)} \Psi, \quad (S7)$$

$$\Psi = \int_0^\infty E_0(z') e^{i\Delta k_z z'} dz', \quad (S8)$$

where Δk_z is the phase mismatch of the reflected SHG, $\chi_{BIL}^{(2)}$ is the second-order nonlinear susceptibility of the BIL (bonded interfacial layer) water¹², Ψ is the phase-modulated potential difference across the diffuse layer (DL), and $\chi_{Bulk}^{(3)}$ is the third-order nonlinear susceptibility of bulk water. It was found that¹¹ $\chi_{Bulk}^{(3)} = (9.56 \pm 1.87) \times 10^{-22} \text{ m}^2/\text{V}^2$ at $2\omega \sim 20,000 \text{ cm}^{-1}$. Since we knew the surface charge density σ from the fractional surface coverages of Si(5c) and SiO⁻ and the surface potential $\phi(z)$ from the solution of the modified Debye-Boltzmann equation as functions of (bulk) pH (See Sec. S3.1), we could obtain the field $E_0(z')$ from $E_0(z')\hat{z} = -\nabla\phi(z)$, and then deduce the complex Ψ from Equation (S8) with $\Delta k_z^{-1} = 25 \text{ nm}^{12}$. We thus can calculate the $\chi_{DL}^{(2)}$ as a function of pH from Equation (S7).

As the change in $\chi_{BIL}^{(2)}$ with pH was shown in previous studies to be much smaller than the change in the $\chi_{DL}^{(2)}$ term¹³, we hence simply took the $\chi_{eff}^{(2)}$ value at pH 2 (corresponding to $\sigma = 0$) from the SHG results reported in Ref. 14 as $\chi_{BIL}^{(2)}$ for all calculations. We then obtained the total $\chi_{SHG}^{(2)}$ from the sum of $\chi_{BIL}^{(2)}$ and $\chi_{DL}^{(2)}$, as plotted $|\chi_{SHG}^{(2)}|$ versus pH in Fig. 4e (main text) in comparison with the experimental results reported in Ref. 14.

S4. Other characterizations of the Si(5c) species

S4.1 Structure characterization of local environments for Si(5c) formation

Here we focus on specific local structures of the interface that could lead to the formation of Si(5c) species. We first considered the possible role of water in Si(5c) formation using the model surface that had a surface SiOH density of $4.5/\text{nm}^2$. We studied if the silanol acidity and Si(5c) formation could be influenced by the silanols exposure to water. Figure S4a describes the heterogeneous distribution of water density in the bonded interfacial layer (BIL) with silanol groups (circles with index number) at different places. Red circles and indices refer to silanols that would deprotonate to form Si(5c), and the black indices to silanols that would deprotonate to form Si-O⁻. As shown in the Table of Fig. S4b, we found *no* correlation between Si(5c) formation and how SiOH initially interacts with neighboring water.

We next considered the possibility that Si(5c) could result from deprotonation of special types of silanols (e.g. vicinal vs isolated) or SiOH having specific H-bonding with adjacent water molecules (i.e. SiOH bonded to water molecules with a donor bond, an acceptor bond, both donor and acceptor bonds, or no bond at all.) Again, as summarized in the table of Fig. S4b, we found no correlation between Si(5c) formation and H-bonding of SiOH. We can therefore conclude that despite water is an essential reactant in silanol deprotonation, its influence on the acidity of SiOH and Si(5c) formation is not significant.

We then questioned whether there were specific surface morphological factors that can affect the Si(5c) formation. Plotted in Fig. S5a are the Si radial distributions around the oxygen atom of each of the eight SiOH groups on the $4.5 \text{ OH}/\text{nm}^2$ silica surface, taking into accounts all Si atoms in the system. In each radial distribution curve, the first peak (at $r > 2.5 \text{ \AA}$) corresponds to the distance between the oxygen atom and its closest neighboring Si atom that is *not* covalently bonded to it. The three distributions plotted in red, blue, and black dashed curves are the only ones in

which the first peak appears for $r_{O-Si} < 3.5 \text{ \AA}$. Interestingly, they were the only three SiOH groups that formed Si(5c) upon deprotonation. This suggests that the close proximity between the oxygen atom of a given SiOH and a neighboring Si atom is important for the Si(5c) formation.

We also question the effect of SiOH orientation towards a neighboring SiO₄ on Si(5c) formation. We plot in Fig. S5b the probability of Si(5c) formation as a function of the angle x (defined as the angle between the two blue arrows in the sketch at the top of Fig. S5b) that describes the relative orientation of the silanol Si-OH bond (that performs the nucleophilic attack) with respect to the SiO₄ tetrahedron being attacked, and the distance between the oxygen atoms, i.e. O* of SiOH and O of SiO₄ labelled in Fig. S5b. Three peaks appear in this plot, all are obtained for $x \sim 90^\circ$, and each of the peaks is obtained for a different value of the O*-O distance. They reflect ideal geometries for the formation of a Si(5c) species that best avoid steric hindrance between the three oxygens of SiO₄ and the SiOH during the nucleophilic attack. The above structural analysis shows that proximity and appropriate relative orientation between SiOH and a neighboring SiO₄ tetrahedron are key factors for the formation of Si(5c).

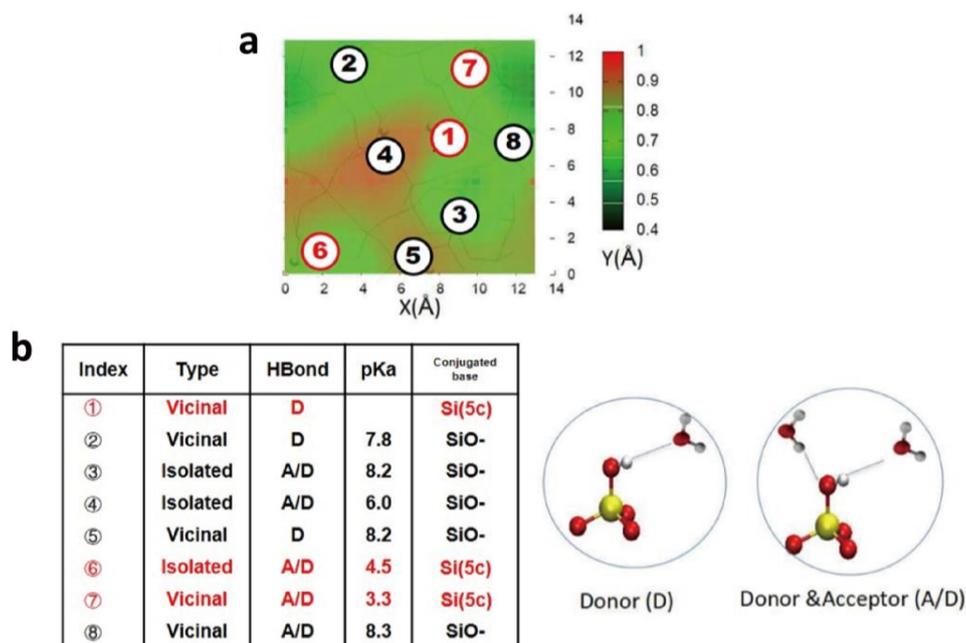


Figure S4 | (a) Water density map in the bonded interfacial layer above the 4.5 SiOH/nm² silica surface. The color coding (vertical scale in the plots) goes from black (low density of water) to red (high density of water). The eight surface Si-OH silanols at the model silica surface are labelled by their index numbers (in circles); the silanols that transform into Si(5c) species, once deprotonated, are labelled in red and the others in black. (b), Table describing some properties of each indexed silanol, i.e., isolated or vicinal, H-bonding pattern with water molecules in the BIL, pKa value and the nature of the conjugate base [SiO⁻ or Si(5c)]. We note that silanols that are not H-bonded to water are not statistically relevant at the investigated interface. On the right of (b) is the scheme of the relevant H-bonding patterns obtained in our DFT-MD simulations.

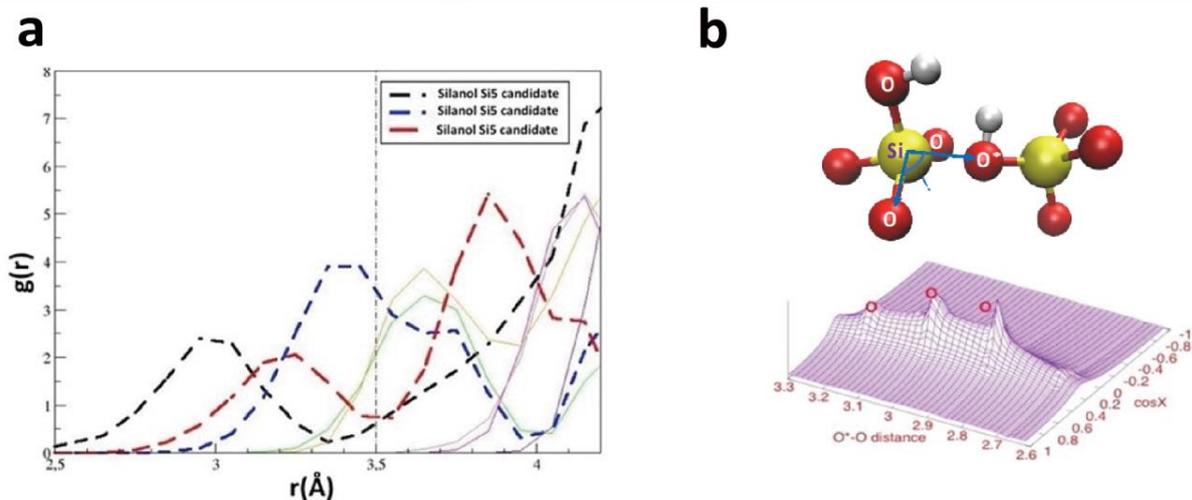


Figure S5 | (a) Si-O radial distribution curves of non-covalently bonded Si atoms around the oxygen atom of each of the eight SiOH groups located on the surface of the 4.5 OH/nm² silica model. There are 8 plots reported, among which the red, black and blue dashed ones correspond to the SiOH groups whose deprotonation leads to the formation of Si(5c). (b) The top frame is a snapshot from the DFT-MD simulations that describes the SiO₄ tetrahedron in optimum orientation with respect to the reactive SiOH for the formation of the Si(5c) species. The bottom frame shows the distribution of Si(5c) formation with as a function of the O*-O distance and of the angle x (both are defined in the top frame) between SiOH and SiO₄. The three peaks marked with red circles identify the three most probable geometries for Si(5c) formation found in our simulations.

S4.2 Comparison to crystalline Quartz surfaces

To investigate the possible existence of Si(5c) species on crystalline silica aqueous surfaces, we performed supplementary metadynamics simulations (same method as presented in the main text, with the same choice for the two coordinates of metadynamics) on the (0001) facet of the α -quartz aqueous interface. Both acidic (pKa = 3.9) and basic silanols (pKa = 8.3) were discovered, in line with the results presented in Ref. 15 for the same quartz surface (pKa values of 5.6 ± 0.6 and 8.5 ± 0.6) despite that our calculation method was different from that of Ref. 15. In Ref. 15 pKa

values were obtained by constrained-dynamics with the “reversible proton insertion method”. Geometrical constraints were applied to locate the H^+ on a specific water molecule with a specific geometry, preventing the proton to diffuse into the water through the Grotthuss mechanism. Consequently, in calculating the free energy, a correction was needed to take into account the entropy contribution of the proton diffusion into water. In our metadynamics simulations, there was no need for this correction, as the proton can freely diffuse into water via the Grotthuss mechanism. It involved interacting with, on average, 3 to 5 water molecules before the proton returned to the surface. The good agreement between the two methods shows that the sampling of our metadynamics was sufficient to capture the entropy contribution to the free energy due to proton diffusion in water.

Though the pKa values for the quartz/water interface are similar to those for the silica/water interface, the underlying chemistry is remarkably different. The free energy landscapes depicted in Fig. S6 for α -(0001)-quartz shows that SiO^- is the stable conjugate base of both acidic and basic silanols, contrary to what we observed for the amorphous surface models. There is *no* minimum on the free energy landscape associated to the Si(5c) species after deprotonation of an acidic silanol. The two pKa values for the α -(0001)-quartz aqueous interface comes from two silanol groups of different orientations, one protruding out of the surface and the other lying close to the surface. As we discussed in Sec. S4.1, the distances between the oxygen of a quartz surface silanol and their neighboring Si atoms were systematically larger than 3.5 Å, thus cannot form Si(5c) after deprotonation.

On the other hand, we note that in Ref. 16, Si(5c) structures were indeed found from calculations on the dry facet (011) of the α -quartz surface due to the reconstruction process following the homolytic fracture of an α -quartz crystal at room temperature. We therefore expect

Si(5c) could be formed also at crystalline quartz aqueous surfaces if the required geometric structural conditions are met with a given surface cut.

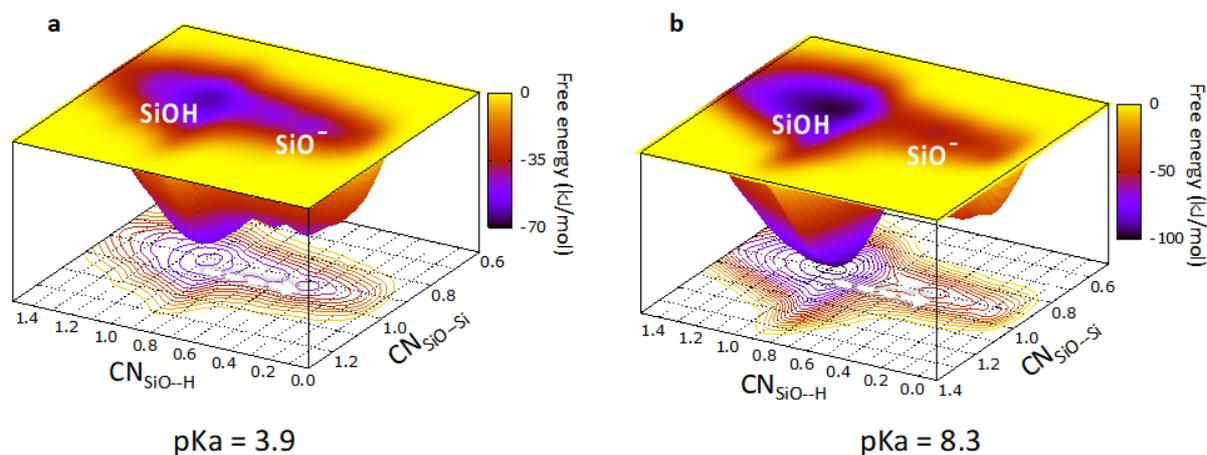


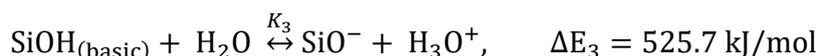
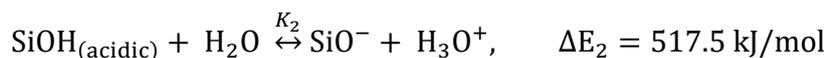
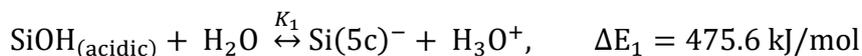
Figure S6 | Free energy landscapes of α -quartz aqueous surface associated with deprotonation processes of (a) acidic silanols (pKa = 3.9) and (b) basic silanols (pKa = 8.3).

S4.3 Stability of Si(5c) vs. SiO⁻ species in vacuum at 0 K

We performed a set of geometry optimization calculations for a silica surface in vacuum (with one water or a hydronium ion to simulate the deprotonation reaction and keep the simulation box neutral). Four configurations were taken: (1) fully hydroxylated silica surface and one water molecule, (2) silica surface with one Si(5c) formed by deprotonating an acidic silanol group and one hydronium ion, (3) silica surface with one SiO⁻ formed by deprotonating an acidic silanol group and one hydronium ion and (4) silica surface with one SiO⁻ formed by deprotonating a basic silanol group and one hydronium ion. The water molecule or the hydronium ion in the simulation boxes was kept far away from the silica surface and constrained during the geometry optimization calculations.

At $T = 0$ K, there is no entropic contribution to the free energy of the deprotonation reaction. We calculate the energies of the different deprotonated states and compared them with the energy

of the fully hydroxylated state of the silica surface, we obtain the energy differences (ΔE) for the three deprotonation reactions:



So even at 0 K where temperature and entropy effects do not contribute, we still observe two populations of SiOH, one more acidic (i.e. with smaller ΔE) and one more basic (i.e. with larger ΔE), with the more acidic population having Si(5c) as the stable conjugate base. The energy difference for formation of Si(5c) was lower by 41.9 kJ/mol than that for formation of SiO⁻. Meanwhile, as expected, the energy costs (ΔE) for the deprotonation reactions at 0 K were one order of magnitude higher than the free energy values at room temperature described in the main text (for example, $\Delta E = 22.4$ kJ/mol for deprotonation into Si(5c) at room temperature). Despite being crude, these geometry optimization calculations performed at 0 K in vacuum confirmed our conclusion that it is the local geometry of acidic silanols, instead of their interaction with water, mainly responsible for the formation of Si(5c).

References

1. Shen, Y. R. *Fundamentals of Sum-Frequency spectroscopy*, 1st edn. (Cambridge University Press, 2016).
2. Ostroverkhov, V., Waychunas, G. A. & Shen, Y. R. Vibrational spectra of water at water/alpha-quartz (0001) interface. *Chem. Phys. Lett.* **386**, 144-148 (2004).
3. Behrens, S. H. & Grier, D. G. The charge of glass and silica surfaces. *J. Chem. Phys.* **115**, 6716-6721 (2001).
4. Liu, W. & Shen, Y. R. Surface vibrational modes of alpha-quartz(0001) probed by sum-frequency spectroscopy. *Phys. Rev. Lett.* **101**, 16101 (2008).
5. Shi, H., Zhang, Y., Wang, H. & Liu, W. Matrix formalism for radiating polarization sheets in multilayer structures of arbitrary composition. *Chin. Opt. Lett.* **15**, 81901 (2017).
6. Masini, P. & Bernasconi, M. Ab initio simulations of hydroxylation and dehydroxylation reactions at surfaces: Amorphous silica and brucite. *J. Phys. Condens. Matter.* **14**, 4133-4144 (2002).
7. Ugliengo, P. et al. Realistic models of hydroxylated amorphous silica surfaces and MCM-41 mesoporous material simulated by large-scale periodic B3LYP calculations. *Adv. Mater.* **20**, 4579-4583 (2008).
8. Andelman, D., Orland, H. & Borukhov, I. Steric effects in electrolytes: A modified Poisson-Boltzmann equation. *Phys. Rev. Lett.* **79**, 435-438 (1997).
9. Bazant, M. Z., Ajdari, A. & Kilic, M. S. Steric effects in the dynamics of electrolytes at large applied voltages. II. Modified Poisson-Nernst-Planck equations. *Phys. Rev. E.* **75**, 21503 (2007).
10. Urashima, S., Myalitsin, A., Nihonyanagi, S. & Tahara, T. The topmost water structure at a charged silica/aqueous interface revealed by heterodyne-detected vibrational sum frequency generation spectroscopy. *J. Phys. Chem. Lett.* **9**, 4109-4114 (2018).
11. Dalstein, L., Chiang, K. & Wen, Y. Direct quantification of water surface charge by Phase-Sensitive second harmonic spectroscopy. *J. Phys. Chem. Lett.* **10**, 5200-5205 (2019).
12. Wang, H., Hu, X. & Wang, H. Charge-induced $\chi^{(3)}$ susceptibility in interfacial nonlinear optical spectroscopy beyond the bulk aqueous contributions: The case for silica/water interface. *J. Phys. Chem. C.* **125**, 26208-26215 (2021).
13. Pezzotti, S., Galimberti, D. R. & Gageot, M. 2D H-bond network as the topmost skin to the air-water interface. *J. Phys. Chem. Lett.* **8**, 3133-3141 (2017).
14. Ong, S., Zhao, X. & Eissenthal, K. B. Polarization of water molecules at a charged interface: Second harmonic studies of the silica/water interface. *Chem. Phys. Lett.* **191**, 327-335 (1992).
15. Sulpizi, M., Gageot, M. & Sprik, M. The Silica-Water interface: How the silanols determine the surface acidity and modulate the water properties. *J. Chem. Theory Comput.* **8**, 1037-1047 (2012).
16. Lopes, P. E. M., Demchuk, E. & Mackerell Jr., A. D. Reconstruction of the (011) surface on α -quartz: A semiclassical Ab initio molecular dynamics study. *Int. J. Quantum Chem.* **109**, 50-64 (2009).

Chapter 6

General conclusions and perspectives

Aqueous interfaces play important roles in many disciplines of science. However, it is challenging to unravel physical and chemical properties at the interface at the microscopic/molecular level, such as molecular hydrophobicity and hydrophilicity, the actual structure of the water in the Binding Interfacial Layer (BIL), wetting, surface chemistry. Experimentally, surface specific techniques, such as Sum Frequency Generation (SFG) and Second Harmonic Generation (SHG) spectroscopies, are often utilized to selectively probe interfaces. The interpretation of these experiments requires however a thorough understanding of the specific synergetic interfacial organization of the water and top surface, which can not be achieved without theoretical modeling.

In this manuscript, DFT-MD (Density Functional Theory based Molecular Dynamics) and FF-MD (Classical Force-Field Molecular Dynamics) simulations, coupled with theoretical and experimental SFG spectroscopies (in collaboration with experimental groups), have been applied to gain a global understanding of interfacial water structures in contact with various aqueous surfaces, ranging from organic to oxide surfaces.

To investigate the molecular hydrophobicity and hydrophilicity and provide a quantitative measurement of it at an interface, we used a platform composed of a set of Self-Assembled Monolayers (SAM) in contact with liquid water, with well-controlled hydrophobicity/hydrophilicity. Specific interfacial water organizations of these SAMs/water interfaces were revealed by coupling MD simulations and a combination of experimental and theoretical SFG spectra, as shown in chapter 3. The molecular interpretation of the SFG spectra that we obtained from the simulations allowed for instance to explain the apparent paradox of the BIL-SFG spectrum of the PEG/water interface, which shows no SFG intensity in the whole O-H stretching frequency range. This is indeed due to a compensation of the spectroscopic contributions from OH groups of BIL water that are H-bonded to the PEGs and the subsequent

water layer. Furthermore, we introduced an approach to include the DL-SFG contribution in the theoretical SFG spectra without the need for the actual DFT-MD simulations at a given pH.

In the same chapter, a novel microscopic metric for hydrophilicity, denoted the H/V descriptor, was also introduced, which for the first time, bridges structural and spectroscopic measures.

Time-resolved and spatially-resolved H/V were further developed that captured dynamics and local hydrophobicity at these heterogeneous interfaces. Among the three studied SAM/water interfaces (OTS, mixed OTS/PEG, and PEG), the PEG/water interface was demonstrated to be highly dynamical due to the driving force of the hydrophilic oxygen atoms of the PEG chains in pursuit of water. This dynamics created local hydrophobic patches at macroscopically hydrophilic PEG/water interface over a nanometer length scale and over a nanosecond time scale. Such dynamical hydrophobicity at the PEG/water interface is relevant for the physical and chemical processes occurring at the ns time-scale, such as ion adsorption and diffusion, chemical reactions, as they are dictated by instantaneous and local properties instead of averaged properties.

Thanks to the coupling between MD simulations and SFG spectroscopy, a SFG spectroscopy database can be accumulated at various aqueous interfaces. Fingerprints of a certain number of surface populations defined by the O-H dipole orientations and surface-water interactions were successfully extracted from DFT-MD calculated theoretical SFG spectra. Based on these, we have proposed a new method for theoretical SFG calculation in chapter 4, denoted the “pop model”. This approach manages to construct and at the same time to interpret SFG spectra from the knowledge of a limited number of pre-computed spectroscopic fingerprints of surface populations, as long as their corresponding abundances are known. The abundance of the interface populations can be obtained either statistically via simulations or analytically from theoretical models. We used both approaches in two different chapters and different contexts, especially one context where DFT-MD and FF-MD simulations are out-of-reach because of the pH conditions required. The model and various applications are described in chapter 4, while its application in the context of silica surface chemistry, is presented in chapter 5.

This approach circumvents the need to calculate the time correlation functions between polarizability and dipole moment or velocity-velocity correlation function (all described in chapter 2). Our new approach highly increases the expediency for theoretical SFG calculation. Most notably, it provides a way to calculate theoretical SFG spectra while DFT-MD or FF-MD simulations are unavailable, as long as the abundance of interface populations can be deduced through existing theoretical models.

A nice example has been shown in chapter 5 by calculating the pH-dependent SFG spectroscopy of the silica surface in the Si-O stretch region, without performing DFT-MD simulations at these different pH conditions.

In this innovative joint experimental/theoretical work, we furthermore revealed the mechanism for the historically debated multimodal behavior of the silica surface deprotonation. On the experimental side, our collaborator Prof. W. Liu’s group developed an in-situ SFG experimental technique in the phonon frequency range (800–1200 cm^{-1}), that enables for the first time to follow the physics and chemistry of the silica surface buried in water, under various pH conditions. On our theory side, we uncovered a new surface species, a five-coordinated Si species $\text{Si}(5\text{c})^-$, by combining DFT-MD simulations and pH-dependent theoretical SFG calculations via the “pop model” method.

The discovery of over-coordinated surface species, the innovative experimental technique, as well as the new strategy of combining metadynamics simulations with pH-dependent theoretical SFG spectra calculations without the need to perform simulations under different pH values will have a significant impact on our future research. In addition, the water properties at this interface need to be re-examined now that we have a more realistic knowledge of the oxide surface structure. In situ studies on various oxide/water interfaces are necessary for discovering new physics and chemistry at the surface in contact with water.

Perspectives

I believe my works open new exciting perspectives in our research domain.

The structural molecular H/V descriptor shed light on our understanding of interfacial hydrophobicity/hydrophilicity at the molecular level. It will be interesting to correlate our H/V descriptor with other microscopic hydrophobicity/hydrophilicity descriptors in the literature, e.g. density fluctuations. This will allow us to combine our understandings with different puzzle pieces to get a complete picture of interfacial hydrophobicity/hydrophilicity at the molecular level. In the meantime, this H/V descriptor can be applied for all investigations of aqueous interfaces to give a quantitative measure of hydrophilicity and hydrophobicity of any aqueous interface. It can certainly be generalized to other solvents that can form H-bonds at the interface.

The time-resolved and spatially-resolved H/V descriptor managed to capture the complexities of hydrophobicity/hydrophilicity at a heterogeneous interface. Combining these descriptors, we uncovered the heterogeneity in hydrophobicity/hydrophilicity at nanosecond and nanometer scales, for a nominally macroscopically hydrophilic interface (PEG/water).

The spatially-resolved H/V map reaches angstrom resolution thanks to the intrinsic locality of the H/V descriptor. The application of this map on other model SAM/water interfaces is ongoing work. The surface heterogeneity and hydrophobicity/hydrophilicity can be tuned through the change in the spatial distribution of OTS and PEG components. This work aims to further investigate the subtle topological effects on local

hydrophobicity/hydrophilicity and how they can affect for instance chemical reactions. This is an exciting perspective to better understand chemistry at the angstrom resolution in space.

The combination of the time-resolved and spatially-resolved H/V descriptor can naturally be extended to metal/water interfaces to investigate how the heterogeneity in surface polarization (i.e. resulting from charge transfer processes) influences the local hydrophobicity/hydrophilicity at the interface. The same approach can be further applied to biological interfaces, such as protein solvation, to quantify the time scale of the dynamics of the biomolecule and that of the induced reorganization of the interface network.

The “pop model” approach for SFG calculation and interpretation can be further developed into more sophisticated approaches. The surface-water populations are, for now, simply defined by the orientation and local H-bond coordination. The definitions of surface-water populations can be further refined.

As a starting point, in an ongoing project in the group, by studying the Quartz/water interfaces, the surface-water populations can be categorized into surface-water motifs composed of water and surface silanols interconnected through H-bonds. These motifs are highly relevant for the physical and chemical properties of the interfaces, such as local hydrophobicity/hydrophilicity, adsorption of ions, etc. Based on these studies, in the future, the motifs at other aqueous interfaces are expected to be automatically identified and defined by graph theory techniques that are developed in the group. This project is in collaboration with the group of Prof. Dominique Barth at the University of Versailles.

With the database of the SFG fingerprints, SFG spectra of aqueous interfaces of interest can be easily constructed by deducing the statistics in the type of motifs from simulations or analytical models.

Furthermore, by correlating the surface properties, e.g. silanol density at silica/water interfaces, morphology, and geometrical parameters, with motifs present at the interfaces, we should be able to establish the relationships between the surface parameters and the motifs through machine learning approaches.

To give a simple example, high silanol densities at silica/water interfaces result in more silanol-silanol H-bonds at the surface, that lead to more collective motifs. Therefore, once the correlations between the surface parameters and motifs are rationalized, most viably through machine learning tools, we should be able to design interfaces with desired motifs components by modulating the surface parameters. This could pave the way for motif-oriented surface engineering.

Ideally, we should be able to anticipate the SFG spectrum of a given interface solely based on its surface parameters, with the well-established connections among surface parameters, motifs, and SFG signals. Reverse engineering should also be possible. The surface properties of a given interface

could be deduced if the SFG spectrum is available. This can be achieved by searching for the motif scenario that builds the best-fitting spectrum.

Last but not least, the discovery of the $\text{Si}(5\text{c})^-$ species revolutionized our understanding of the aqueous silica surface under different pH conditions. It will trigger numerous future studies of the silica/water interface and other oxide interfaces from this perspective.

One of the unanswered questions is for instance how electrolytes affect the pKa values of surface silanols and the stability of the $\text{Si}(5\text{c})^-$ species. This is currently under investigation experimentally by the group of Prof. W. Liu. On the theory side, it would be interesting to use the same approach as presented in chapter 5. This will consist of the same three steps: deducing pKa values by metadynamics simulations, extracting SFG fingerprints of surface populations in presence of electrolytes, and computing SFG spectra of the silica surface in water at different electrolyte types and concentrations by the “pop model”.

Another spontaneous question is, how does the new $\text{Si}(5\text{c})^-$ species shall affect the interfacial water properties? To answer this, the theoretical SFG spectra of the BIL-water at the silica/water interface under different pH conditions should be calculated. This can now be achieved by the “pop model” approach. By extracting the SFG signatures of BIL-water molecules interacting with the three surface species (SiOH , $\text{Si}(5\text{c})^-$ and SiO^-) respectively, we will be able to construct the BIL-water pH-dependent SFG spectra as long as the BIL compositions at different pHs are known.

Furthermore, based on the results obtained for silica surfaces, it is reasonable to expect that pH-induced surface reconstructions can also occur on other oxide surfaces such as alumina. However, the feasibility and nature of such reconstructions would depend on the local surface geometry and the availability of suitable sites for deprotonation and subsequent formation of new surface bridges. In general, amorphous surfaces may be more conducive to such reconstructions due to the greater diversity of local surface configurations available. Nevertheless, it is also possible for crystalline surfaces to satisfy the necessary conditions for surface reconstruction. Further research is needed to investigate the possibility of pH-induced surface reconstructions on different oxide surfaces and to understand the underlying mechanisms and properties of these phenomena.

Bibliography

- (1) Zhang, L.; Yang, Y.; Kao, Y.-T.; Wang, L.; Zhong, D. *jacs* **2009**, *131*, 10677–10691.
- (2) Lo Nostro, P.; Ninham, B. W. *cr* **2012**, *112*, 2286–2322.
- (3) Ohlin, C. A.; Villa, E. M.; Rustad, J. R.; Casey, W. H. *Nat. Mater.* **2010**, *9*, 11–19.
- (4) Michaelides, A.; Morgenstern, K. *Nat. Mater.* **2007**, *6*, 597–601.
- (5) Zhu, J.; Hu, L.; Zhao, P.; Lee, L. Y. S.; Wong, K.-Y. *Chem. Rev.* **2019**, *120*, 851–918.
- (6) Brigiano, F. S.; Gierada, M.; Tielens, F.; Pietrucci, F. *ACS Catal.* **2022**, *12*, 2821–2830.
- (7) Brini, E.; Fennell, C. J.; Fernandez-Serra, M.; Hribar-Lee, B.; Luksic, M.; Dill, K. A. *cr* **2017**, *117*, 12385–12414.
- (8) Acharya, H.; Vembanur, S.; Jamadagni, S. N.; Garde, S. *Faraday Discuss.* **2010**, *146*, 353–365.
- (9) Jamadagni, S. N.; Godawat, R.; Garde, S. *Annu. Rev. Chem. Biomol.* **2011**, *2*, 147–171.
- (10) Lefevre, G. *Adv Colloid Interface Sci.* **2004**, *107*, 109–123.
- (11) Allongue, P. *Phys. Rev. Lett.* **1996**, *77*, 1986–1989.
- (12) Velasco-Velez, J.-J.; Wu, C. H.; Pascal, T. A.; Wan, L. F.; Guo, J.; Prendergast, D.; Salmeron, M. *Science* **2014**, *346*, 831–834.
- (13) Shen, Y. *Annu. rev. Phys. Chem.* **2013**, *64*, 129.
- (14) Hunt, J.; Guyot-Sionnest, P.; Shen, Y. *cpl* **1987**, *133*, 189–192.
- (15) Zhu, X.; Suhr, H.; Shen, Y. *prb* **1987**, *35*, 3047.
- (16) Shen, Y., *Fundamentals of sum-frequency spectroscopy*; Cambridge University Press: 2016.
- (17) Du, Q.; Superfine, R.; Freysz, E.; Shen, Y. *Phys. Rev. Lett.* **1993**, *70*, 2313.
- (18) Eftekhari-Bafrooei, A.; Borguet, E. *J. Am. Chem. Soc.* **2010**, *132*, 3756–3761.
- (19) Zhang, Z.; Piatkowski, L.; Bakker, H. J.; Bonn, M. *Nat. Chem.* **2011**, *3*, 888–893.

- (20) Hsieh, C.-S.; Okuno, M.; Hunger, J.; Backus, E. H.; Nagata, Y.; Bonn, M. *Angew. Chem. Int. Ed.* **2014**, *53*, 8146–8149.
- (21) Tian, C.; Shen, Y. *Surf. Sci. Rep.* **2014**, *69*, 105–131.
- (22) Jubb, A. M.; Hua, W.; Allen, H. C. *Annu. Rev. Phys. Chem.* **2012**, *63*, 107–130.
- (23) Dalstein, L.; Potapova, E.; Tyrode, E. *Phys. Chem. Chem. Phys.* **2017**, *19*, 10343–10349.
- (24) Du, Q.; Freysz, E.; Shen, Y. R. *Science* **1994**, *264*, 826–828.
- (25) Tang, F.; Ohto, T.; Sun, S.; Rouxel, J. R.; Imoto, S.; Backus, E. H.; Mukamel, S.; Bonn, M.; Nagata, Y. *Chem. Rev.* **2020**, *120*, 3633–3667.
- (26) Pezzotti, S.; Serva, A.; Sebastiani, F.; Brigiano, F. S.; Galimberti, D. R.; Potier, L.; Alfarano, S.; Schwaab, G.; Havenith, M.; Gaigneot, M.-P. *J. Phys. Chem. Lett.* **2021**, *12*, 3827–3836.
- (27) Cyran, J. D.; Donovan, M. A.; Vollmer, D.; Brigiano, F. S.; Pezzotti, S.; Galimberti, D. R.; Gaigneot, M.-P.; Bonn, M.; Backus, E. H. *Proc. Natl. Acad. Sci.* **2019**, *116*, 1520–1525.
- (28) DelloStritto, M.; Piontek, S. M.; Klein, M. L.; Borguet, E. *J. Phys. Chem. C* **2018**, *122*, 21284–21294.
- (29) Kim, D.; Kim, E.; Park, S.; Kim, S.; Min, B. K.; Yoon, H. J.; Kwak, K.; Cho, M. *Chem* **2021**, *7*, 1602–1614.
- (30) Godawat, R.; Jamadagni, S. N.; Garde, S. *Proc. Natl. Acad. Sci.* **2009**, *106*, 15119–15124.
- (31) Shin, S.; Willard, A. P. *J. Comput. Chem.* **2018**, *14*, 461–465.
- (32) Pezzotti, S.; Galimberti, D. R.; Gaigneot, M.-P. *J. Phys. Chem. Lett.* **2017**, *8*, 3133–3141.
- (33) Pezzotti, S.; Galimberti, D. R.; Shen, Y. R.; Gaigneot, M.-P. *Phys. Chem. Chem. Phys.* **2018**, *20*, 5190–5199.
- (34) Morita, A.; Hynes, J. T. *J. Phys. Chem. B.* **2002**, *106*, 673–685.
- (35) Morita, A.; Ishiyama, T. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5801–5816.
- (36) Khatib, R.; Backus, E. H.; Bonn, M.; Perez-Haro, M.-J.; Gaigneot, M.-P.; Sulpizi, M. *Sci. Rep.* **2016**, *6*, 1–10.
- (37) Ni, Y.; Skinner, J. L. *J. Chem. Phys.* **2016**, *145*, 031103.
- (38) Ishiyama, T.; Imamura, T.; Morita, A. *Chem. Rev.* **2014**, *114*, 8447–8470.
- (39) Ishiyama, T.; Morita, A. *Ann. Rev. Phys. Chem.* **2017**, *68*, 355–377.
- (40) Gaigneot, M.-P.; Sprik, M. *jpcc* **2003**, *107*, 10344–10358.
- (41) Ohto, T.; Usui, K.; Hasegawa, T.; Bonn, M.; Nagata, Y. *jcp* **2015**, *143*, 124702.

-
- (42) Khatib, R.; Backus, E. H.; Bonn, M.; Perez-Haro, M.-J.; Gaigeot, M.-P.; Sulpizi, M. *Sci. Reports* **2016**, *6*, 1–10.
- (43) Pezzotti, S.; Galimberti, D. R.; Gaigeot, M.-P. *Phys. Chem. Chem. Phys.* **2019**, *21*, 22188–22202.
- (44) Wen, Y.-C.; Zha, S.; Liu, X.; Yang, S.; Guo, P.; Shi, G.; Fang, H.; Shen, Y. R.; Tian, C. *Phys. Rev. Lett.* **2016**, *116*, 016101.
- (45) Diebold, U.; Li, S.-C.; Schmid, M. *Annu. Rev. Phys. Chem.* **2010**, *61*, 129–148.
- (46) Ong, S.; Zhao, X.; Eissenthal, K. B. *cpl* **1992**, *191*, 327–335.
- (47) Ostroverkhov, V.; Waychunas, G. A.; Shen, Y. *prl* **2005**, *94*, 046102.
- (48) Fisk, J. D.; Batten, R.; Jones, G.; O'Reill, J. P.; Shaw, A. M. *jpcc* **2005**, *109*, 14475–14480.
- (49) Darlington, A. M.; Gibbs-Davis, J. M. *jpcc* **2015**, *119*, 16560–16567.
- (50) Sulpizi, M.; Gaigeot, M.-P.; Sprik, M. *J. Chem. Theory Comput.* **2012**, *8*, 1037–1047.
- (51) Pfeiffer-Laplaud, M.; Costa, D.; Tielens, F.; Gaigeot, M.-P.; Sulpizi, M. *jpcc* **2015**, *119*, 27354–27362.
- (52) Tielens, F.; Gierada, M.; Handzlik, J.; Calatayud, M. *Catalysis Today* **2020**, *354*, 3–18.
- (53) Gierada, M.; De Proft, F.; Sulpizi, M.; Tielens, F. *jpcc* **2019**, *123*, 17343–17352.
- (54) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511–519.
- (55) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695–1697.
- (56) Martyna, G. J.; Klein, M. L.; Tuckerman, M. *J. Chem. Phys.* **1992**, *97*, 2635–2643.
- (57) Bougueroua, S.; Chantitch, V.; Chen, W.; Pezzotti, S.; Gaigeot, M.-P. *Vibrational Dynamics Of Molecules* **2022**, 416.
- (58) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101–.
- (59) Nagata, Y.; Ohto, T.; Bonn, M.; Kühne, T. *J. Chem. Phys.* **2016**, *144*, 204705.
- (60) Pfeiffer-Laplaud, M.; Gaigeot, M.-P. *J. Phys. Chem. C* **2016**, *120*, 14034–14047.
- (61) Pfeiffer-Laplaud, M.; Gaigeot, M. *J. Phys. Chem. C* **2016**, 4866–4880.
- (62) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401–146405.
- (63) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
-

- (64) Grimme, S. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 211–228.
- (65) Goerigk, L.; Grimme, S. *jctc* **2010**, *6*, 107–126.
- (66) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (67) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (68) Baer, M. D.; Mundy, C. J.; McGrath, M. J.; Kuo, I.-F. W.; Siepmann, J. I.; Tobias, D. J. *jcp* **2011**, *135*, 124712.
- (69) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *jcp* **2010**, *132*, 154104.
- (70) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. *WIREs Comput Mol Sci* **2014**, *4*, 15–25.
- (71) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comp. Phys. Comm.* **2005**, *167*, 103–128.
- (72) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (73) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- (74) Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B.* **1996**, *54*, 1703–1710.
- (75) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463–1473.
- (76) Sutmann, G. Classical molecular dynamics, 2002.
- (77) Allen, M. P.; Tildesley, D. J., *Computer simulation of liquids*; Oxford university press: 2017.
- (78) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. *jctc* **2013**, *9*, 4046–4063.
- (79) Marti, J.; Padró, J.; Guardia, E. *jml* **1994**, *62*, 17–31.
- (80) Berendsen, H. J.; Postma, J. v.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *The Journal of chemical physics* **1984**, *81*, 3684–3690.
- (81) Kroutil, O.; Pezzotti, S.; Gaigeot, M.-P.; Předota, M. *jpcc* **2020**, *124*, 15253–15263.
- (82) Smirnov, K. S. *pccp* **2017**, *19*, 2950–2960.
- (83) Ferguson, D. M. *jcc* **1995**, *16*, 501–511.
- (84) Allinger, N. L. *jacs* **1977**, *99*, 8127–8134.
- (85) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *jacs* **1989**, *111*, 8551–8566.
- (86) Clementi, E.; Chakravorty, S. J. *jcp* **1990**, *93*, 2591–2602.
- (87) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *jacs* **1996**, *118*, 11225–11236.
- (88) Momany, F. A. *jpc* **1978**, *82*, 592–601.
- (89) Cox, S.; Williams, D. *jcc* **1981**, *2*, 304–323.
- (90) Singh, U. C.; Kollman, P. A. *jcc* **1984**, *5*, 129–145.

-
- (91) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *jacs* **1984**, *106*, 765–784.
- (92) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *jcc* **1986**, *7*, 230–252.
- (93) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *jpc* **1993**, *97*, 10269–10280.
- (94) Reynolds, C. A.; Essex, J. W.; Richards, W. G. *jacs* **1992**, *114*, 9075–9079.
- (95) MacKerell Jr, A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S., et al. *jpcb* **1998**, *102*, 3586–3616.
- (96) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *Journal of computational chemistry* **2004**, *25*, 1157–1174.
- (97) Schuler, L. D.; Daura, X.; Van Gunsteren, W. F. *Journal of computational chemistry* **2001**, *22*, 1205–1218.
- (98) Whitehead, C. E.; Breneman, C. M.; Sukumar, N.; Ryan, M. D. *jcc* **2003**, *24*, 512–529.
- (99) Williams, D. E. *jcc* **1988**, *9*, 745–763.
- (100) Kramer, C.; Gedeck, P.; Meuwly, M. *jcc* **2012**, *33*, 1673–1688.
- (101) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio Jr, R. A., et al. *jpcb* **2010**, *114*, 2549–2564.
- (102) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; De Vries, A. H. *jpcb* **2007**, *111*, 7812–7824.
- (103) Van Duin, A. C.; Dasgupta, S.; Lorant, F.; Goddard, W. A. *jpca* **2001**, *105*, 9396–9409.
- (104) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. *jcc* **2005**, *26*, 1701–1718.
- (105) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. *jctc* **2008**, *4*, 435–447.
- (106) Roscioni, O. M.; Muccioli, L.; Mityashin, A.; Cornil, J.; Zannoni, C. *J. Phys. Chem. C* **2016**, *120*, 14652–14662.
- (107) Cygan, R. T.; Liang, J.-J.; Kalinichev, A. G. *jpcb* **2004**, *108*, 1255–1266.
- (108) Mityashin, A.; Roscioni, O. M.; Muccioli, L.; Zannoni, C.; Geskin, V.; Cornil, J.; Janssen, D.; Steudel, S.; Genoe, J.; Heremans, P. *ACS Appl. Mater. Interfaces* **2014**, *6*, 15372–15378.
- (109) Willard, A. P.; Chandler, D. *J. Phys. Chem. B* **2010**, *114*, 1954–1958.
- (110) Tuladhar, A.; Dewan, S.; Pezzotti, S.; Brigiano, F. S.; Creazzo, F.; Gageot, M.-P.; Borguet, E. *J. Am. Chem. Soc.* **2020**, *142*, 6991–7000.
-

- (111) Patel, A. J.; Varilly, P.; Chandler, D. *J. Phys. Chem. B* **2010**, *114*, 1632–1637.
- (112) Bonn, M.; Nagata, Y.; Backus, E. H. *Angew. Chem. Int. Ed.* **2015**, *54*, 5560–5576.
- (113) Morita, A.; Ishiyama, T. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5801–5816.
- (114) Morita, A.; Hynes, J. T. *J. Phys. Chem. B* **2002**, *106*, 673–685.
- (115) Du, Q.; Freysz, E.; Shen, Y. R. *Phys. Rev. Lett.* **1993**, *70*, 2313–2316.
- (116) Ostroverkhov, N.; Tian, C.; Shen, Y. *Phys. Rev. Lett.* **2008**, *100*, 096102.
- (117) Nihonyanagi, S.; Kusaka, R.; Inoue, K.; Adhikari, A.; Yamaguchi, S.; Tahara, T. *J. Chem. Phys.* **2015**, *143*, 124707.
- (118) Sun, S.; Liang, R.; Xu, X.; Zhu, H.; Shen, Y. R.; Tian, C. *J. Chem. Phys.* **2016**, *144*, 244711.
- (119) Gaigeot, M.-P.; Sprik, M. *J. Phys. Chem. B* **2003**, *107*, 10344–10358.
- (120) Medders, G. R.; Paesani, F. *J. Am. Chem. Soc.* **2016**, *138*, 3912–3919.
- (121) Sulpizi, M.; Salanne, M.; Sprik, M.; Gaigeot, M.-P. *jpcl* **2013**, *4*, 83–87.
- (122) Bakels, S.; Gaigeot, M.-P.; Rijs, A. M. *cr* **2020**, *120*, 3233–3260.
- (123) Bougueroua, S.; Chantitch, V.; Chen, W.; Pezzotti, S.; Gaigeot, M.-P. *Vibrational Dynamics Of Molecules* **2022**, 416.
- (124) Sulpizi, M.; Salanne, M.; Sprik, M.; Gaigeot, M.-P. *J. Phys. Chem. Lett.* **2013**, *4*, 83–87.
- (125) Ohto, T.; Usui, K.; Hasegawa, T.; Bonn, M.; Nagata, Y. *J. Chem. Phys.* **2015**, *143*, 124702.
- (126) Galimberti, D. R.; Milani, A.; Tommasini, M.; Castiglioni, C.; Gaigeot, M.-P. *jctc* **2017**, *13*, 3802–3813.
- (127) Corcelli, S. A.; Skinner, J. L. *J. Phys. Chem. A.* **2005**, *109*, 6154–6165.
- (128) Galimberti, D. R.; Milani, A.; Tommasini, M.; Castiglioni, C.; Gaigeot, M.-P. *J. Chem. Theo. Comput.* **2017**, *13*, 3802–3813.
- (129) Pezzotti, S.; Gaigeot, M.-P. *Atmosphere* **2018**, *9*, 396.
- (130) Pezzotti, S.; Galimberti, D. R.; Shen, Y. R.; Gaigeot, M.-P. *Minerals* **2018**, *8*, 305.
- (131) Creazzo, F.; Galimberti, D. R.; Pezzotti, S.; Gaigeot, M.-P. *J. Chem. Phys.* **2019**, *150*, 041721.
- (132) Nostro, P. L.; Ninham, B. W. *Chem. Rev.* **2012**, *112*, 2286.
- (133) Et al., G. E. B. *Chem. Rev.* **1999**, *99*, 77.

-
- (134) Gonella, G.; Lutgebaucks, C.; de Beer, A. G. F.; Roke, S. *J. Phys. Chem. C* **2016**, *120*, 9165–9173.
- (135) Ohno, P. E.; Saslow, S. A.; Wang, H.-f.; Geiger, F. M.; Eienthal, K. B. *Nat. Comm.* **2016**, *7*, 13587–13591.
- (136) Joutsuka, T.; Hirano, T.; Sprik, M.; Morita, A. *Phys. Chem. Chem. Phys.* **2018**, *20*, 3040–3053.
- (137) Urashima, S.-h.; Myalitsin, A.; Nihonyanagi, S.; Tahara, T. *J. Phys. Chem. Lett.* **2018**, *9*, 4109–4114.
- (138) Dreier, L. B.; Nagata, Y.; Lutz, H.; Gonella, G.; Hunger, J.; Backus, E. H. G.; Bonn, M. *Science Advances* **2018**, *4*, DOI: [10.1126/sciadv.aap7415](https://doi.org/10.1126/sciadv.aap7415).
- (139) Sanders, S. E.; Petersen, P. B. *J. Chem. Phys.* **2019**, *150*, 204708.
- (140) Tian, C.; Shen, Y. *Proc. Natl. Acad. Sci.* **2009**, *106*, 15148–15153.
- (141) Dalstein, L.; Chiang, K.-Y.; Wen, Y.-C. *jpcl* **2019**, *10*, 5200–5205.
- (142) Urashima, S.-h.; Myalitsin, A.; Nihonyanagi, S.; Tahara, T. *J. Phys. Chem. Lett.* **2018**, *9*, 4109–4114.
- (143) Yan, M.; Kawamata, Y.; Baran, P. S. *Chem. Rev.* **2017**, *117*, 13230–13319.
- (144) Laage, D.; Elsaesser, T.; Hynes, J. T. *Chemical Reviews* **2017**, *117*, 10694–10725.
- (145) Knipping, E. M.; Lakin, M. J.; Foster, K. L.; Jungwirth, P.; Tobias, D. J.; Gerber, R. B.; Dabdub, D.; Finlayson-Pitts, B. J. *Science* **2000**, *288*, 301–306.
- (146) Katrib, Y.; Deiber, G.; Schweitzer, F.; Mirabel, P.; George, C. *J. Aerosol Sci.* **2001**, *32*, 893–911.
- (147) Clifford, D.; Donaldson, D. J. *J. Phys. Chem. A* **2007**, *111*, 9809–9814.
- (148) Sun, S.; Liang, R.; Xu, X.; Zhu, H.; Shen, Y. R.; Tian, C. *jcp* **2016**, *144*, 244711.
- (149) Medders, G. R.; Paesani, F. *J. Am. Chem. Soc.* **2016**, *138*, 3912–3919.
- (150) Tainter, C.; Ni, Y.; Shi, L. a.; Skinner, J. *J. Phys. Chem. Lett.* **2013**, *4*, 12–17.
- (151) Ni, Y.; Skinner, J. *J. Chem. Phys.* **2016**, *145*, 031103.
- (152) Conti Nibali, V.; Pezzotti, S.; Sebastiani, F.; Galimberti, D.; Schwaab, G.; Heyden, M.; Gaigeot, M.-P.; Havenith, M. *The journal of physical chemistry letters* **2020**, *11*, 4809–4816.
- (153) Liu, W.-T.; Shen, Y. *prl* **2008**, *101*, 016101.
- (154) Pezzotti, S. DFT-MD simulations and theoretical SFG spectroscopy to characterize H-Bonded networks at aqueous interfaces : from hydrophobic to hydrophilic environments, 2019SACLE008, Ph.D. Thesis, 2019.
-

- (155) Jeon, S.; Lee, J.; Andrade, J.; De Gennes, P. *J. Colloid Interface Sci.* **1991**, *142*, 149–158.
- (156) Prime, K. L.; Whitesides, G. M. *Science* **1991**, 1164–1167.
- (157) Harder, P.; Grunze, M.; Dahint, R.; Whitesides, G.; Laibinis, P. *J. Phys. Chem. B* **1998**, *102*, 426–436.
- (158) DiBenedetto, S. A.; Facchetti, A.; Ratner, M. A.; Marks, T. J. *Adv. Mater.* **2009**, *21*, 1407–1433.
- (159) Monroe, J. I.; Shell, M. S. *Proc. Natl. Acad. Sci.* **2018**, *115*, 8093–8098.
- (160) Monroe, J.; Barry, M.; DeStefano, A.; Aydogan Gokturk, P.; Jiao, S.; Robinson-Brown, D.; Webber, T.; Crumlin, E. J.; Han, S.; Shell, M. S. *Annu. Rev. Chem. Biomol.* **2020**, *11*, 523–557.
- (161) Dallin, B. C.; Van Lehn, R. C. *J. Phys. Chem. Lett.* **2019**, *10*, 3991–3997.
- (162) Ye, S.; Nihonyanagi, S.; Uosaki, K. *Phys. Chem. Chem. Phys.* **2001**, *3*, 3463–3469.
- (163) Ugliengo, P.; Sodupe, M.; Musso, F.; Bush, I.; Orlando, R.; Dovesi, R. *Adv. Mater.* **2008**, *20*, 4579–4583.
- (164) Chen, W.; Sanders, S. E.; Ozdamar, B.; Louaas, D.; Brigiano, F. S.; Pezzotti, S.; Petersen, P. B.; Gaigeot, M.-P. *npj* **2023**, *14*, 1301–1309.
- (165) Yeon, H.; Wang, C.; Van Lehn, R. C.; Abbott, N. L. *Langmuir* **2017**, *33*, 4628–4637.
- (166) Dallin, B. C.; Yeon, H.; Ostwalt, A. R.; Abbott, N. L.; Van Lehn, R. C. *Langmuir* **2019**, *35*, 2078–2088.
- (167) Taifan, W.; Boily, J.-F.; Baltrusaitis, J. *Surface Science Reports* **2016**, *71*, 595–671.
- (168) Nong, H. N.; Falling, L. J.; Bergmann, A.; Klingenhof, M.; Tran, H. P.; Spöri, C.; Mom, R.; Timoshenko, J.; Zichittella, G.; Knop-Gericke, A., et al. *Nature* **2020**, *587*, 408–413.
- (169) Lagström, T.; Gmür, T. A.; Quaroni, L.; Goel, A.; Brown, M. A. *Langmuir* **2015**, *31*, 3621–3626.
- (170) Warring, S. L.; Beattie, D. A.; McQuillan, A. J. *Langmuir* **2016**, *32*, 1568–1576.
- (171) Sprik, M. *cp* **2000**, *258*, 139–150.
- (172) Borukhov, I.; Andelman, D.; Orland, H. *prl* **1997**, *79*, 435.
- (173) Kilic, M. S.; Bazant, M. Z.; Ajdari, A. *pre* **2007**, *75*, 021502.
- (174) Joutsuka, T.; Hirano, T.; Sprik, M.; Morita, A. *Phys. Chem. Chem. Phys.* **2018**, *20*, 3040–3053.
- (175) Wang, H.; Hu, X.-H.; Wang, H.-F. *jpcc* **2021**, *125*, 26208–26215.

-
- (176) Masini, P.; Bernasconi, M. *J. Condens. Matter Phys.* **2002**, *14*, 4133.
- (177) Lopes, P. E.; Demchuk, E.; Mackerell Jr, A. D. *Int. J. Quantum Chem.* **2009**, *109*, 50–64.
- (178) Gleizes, A. N.; Samélor, D.; Vahlas, C.; Sarou-Kanian, V.; Florian, P.; Massiot, D. In *Advances in Science and Technology*, 2014; Vol. 91, pp 123–133.
- (179) Sung, J.; Zhang, L.; Tian, C.; Shen, Y. R.; Waychunas, G. A. *J. Phys. Chem. C* **2011**, *115*, 13887–13893.
- (180) Ngouana-Wakou, B.; Cornette, P.; Corral Valero, M.; Costa, D.; Raybaud, P. *J. Phys. Chem. C* **2017**, *121*, 10351–10363.
- (181) Lis, D.; Backus, E. H. G.; Hunger, J.; Parekh, S. H.; Bonn, M. *Science* **2014**, *344*, 1138–1142.
- (182) Gragson, D. E.; McCarty, B. M.; Richmond, G. L. *J. Am. Chem. Soc.* **1997**, *119*, 6144–6152.
- (183) Nihonyanagi, S.; Yamaguchi, S.; Tahara, T. *J. Am. Chem. Soc.* **2014**, *136*, 6155–6158.
- (184) Mondal, J. A.; Nihonyanagi, S.; Yamaguchi, S.; Tahara, T. *J. Am. Chem. Soc.* **2012**, *134*, 7842–7850.
- (185) Sung, J.; Zhang, L.; Tian, C.; Shen, Y. R.; Waychunas, G. A. *J. Phys. Chem. C* **2011**, *115*, 13887–13893.
- (186) DelloStritto, M. J.; Piontek, S. M.; Klein, M. L.; Borguet, E. *J. Phys. Chem. Lett.* **2019**, *10*, 2031–2036.
- (187) Pezzotti, S.; Serva, A.; Gaigeot, M.-P. *J. Chem. Phys.* **2018**, *148*, 174701–174710.
- (188) Todorova, T.; Seitsonen, A. P.; Hutter, J.; Kuo, I.-F. W.; Mundy, C. J. *J. Phys. Chem. B* **2006**, *110*, 3685–3691.
- (189) VandeVondele, J.; Mohamed, F.; Krack, M.; Hutter, J.; Sprik, M.; Parrinello, M. *jcp* **2005**, *122*, 014515.
- (190) Gaiduk, A. P.; Gygi, F.; Galli, G. *J. Phys. Chem. Lett.* **2015**, *6*, 2902–2908.
- (191) Behler, J. *Chem. Rev.* **2021**, *121*, 10037–10072.
- (192) Gkeka, P.; Stoltz, G.; Barati Farimani, A.; Belkacemi, Z.; Ceriotti, M.; Chodera, J. D.; Dinner, A. R.; Ferguson, A. L.; Maillet, J.-B.; Minoux, H., et al. *jctc* **2020**, *16*, 4757–4775.
- (193) Toukan, K.; Rahman, A. *prb* **1985**, *31*, 2643.
- (194) Laskin, A.; Wang, H.; Robertson, W. H.; Cowin, J. P.; Ezell, M. J.; Finlayson-Pitts, B. J. *J. Phys. Chem. A* **2006**, *110*, 10619–10627.
- (195) Graham, J. D.; Roberts, J. T.; Anderson, L. D.; Grassian, V. H. *J. Phys. Chem.* **1996**, *100*, 19551–19558.
-

- (196) Hsieh, C.-S.; Okuno, M.; Hunger, J.; Backus, E. H. G.; Nagata, Y.; Bonn, M. *Angew. Chem.* **2014**, *53*, 8146–8149.
- (197) Alfarano, S. R. et al. *Proc. Natl. Acad. Sci.* **2021**, *118*.
- (198) Montenegro, A.; Dutta, C.; Mammetkuliev, M.; Shi, H.; Hou, B.; Bhattacharyya, D.; Zhao, B.; Cronin, S. B.; Benderskii, A. V. *Nature* **2021**, *594*, 62–65.
- (199) Qi, C.; Zhu, Z.; Wang, C.; Zheng, Y. *J. Phys. Chem. Lett.* **2021**, *12*, 931–937.
- (200) Giberti, F.; Hassanali, A. A. *J. Chem. Phys.* **2017**, *146*, 244703.
- (201) Grosjean, B.; Bocquet, M.-L.; Vuilleumier, R. *Nat. Commun.* **2019**, *10*, 1–8.
- (202) Piontek, S. M.; Borguet, E. *J. Phys. Chem. C* **2022**, *126*, 2307–2324.
- (203) Limmer, D. T.; Willard, A. P.; Madden, P.; Chandler, D. *Proc. Natl. Acad. Sci.* **2013**, *110*, 4200–4205.
- (204) Serva, A.; Salanne, M.; Havenith, M.; Pezzotti, S. *Proc. Natl. Acad. Sci.* **2021**, *118*.
- (205) Galli, G. *Proc. Natl. Acad. Sci.* **2007**, *104*, 2557–2558.
- (206) Rana, K.; Boyd, S. A.; Teppen, B. J.; Li, H.; Liu, C.; Johnston, C. T. *Phys. Chem. Chem. Phys.* **2009**, *11*, 2976–2985.
- (207) Li, J.-L.; Car, R.; Tang, C.; Wingreen, N. S. *Proc. Natl. Acad. Sci.* **2007**, *104*, 2626–2630.
- (208) Wang, H.; Xu, Q.; Liu, Z.; Tang, Y.; Wei, G.; Shen, Y. R.; Liu, W.-T. *J. Phys. Chem. Lett.* **2019**, *10*, 5943–5948.
- (209) Oleinikova, A.; Smolin, N.; Brovchenko, I.; Geiger, A.; Winter, R. *J. Phys. Chem. B* **2005**, *109*, 1988–1998.
- (210) Kelkar, A. S.; Dallin, B. C.; Van Lehn, R. C. *J. Phys. Chem. B* **2020**, *124*, 9103–9114.
- (211) Mityashin, A.; Roscioni, O. M.; Muccioli, L.; Zannoni, C.; Geskin, V.; Cornil, J.; Janssen, D.; Steudel, S.; Genoe, J.; Heremans, P. *ACS Applied Materials & Interfaces* **2014**, *6*, 15372–15378.
- (212) Sulpizi, M.; Gaigeot, M.-P.; Sprik, M. *J. Chem. Theory Comput.* **2012**, *8*, 1037–1047.
- (213) Brown Jr, G. E.; Henrich, V.; Casey, W.; Clark, D.; Eggleston, C.; Andrew Felmy, A. F.; Goodman, D. W.; Gratzel, M.; Maciel, G.; McCarthy, M. I., et al. **1999**.
- (214) Uchida, T.; Kurita, Y.; Koizumi, N.; Kubo, M. *J. Polym. Sci.* **1956**, *21*, 313–322.
- (215) Barry, M. E.; Davidson, E. C.; Zhang, C.; Patterson, A. L.; Yu, B.; Leonardi, A. K.; Duzen, N.; Malaviya, K.; Clarke, J. L.; Finlay, J. A., et al. *Macromolecules* **2019**, *52*, 1287–1295.

-
- (216) Giovambattista, N.; Debenedetti, P. G.; Rossky, P. J. *J. Phys. Chem. C* **2007**, *111*, 1323–1332.
- (217) Lee, C.-Y.; McCammon, J. A.; Rossky, P. *J. Chem. Phys.* **1984**, *80*, 4448–4455.
- (218) Rego, N. B.; Xi, E.; Patel, A. J. *Proc. Natl. Acad. Sci.* **2021**, *118*.
- (219) Xi, E.; Venkateshwaran, V.; Li, L.; Rego, N.; Patel, A. J.; Garde, S. *Proc. Natl. Acad. Sci.* **2017**, *114*, 13345–13350.
- (220) Dove, P. M. *Am. J. Sci.* **1994**, *294*, 665–712.
- (221) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 15–25.
- (222) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Commun.* **2005**, *167*, 103–128.
- (223) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
- (224) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (225) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463–1473.
- (226) White, J. A.; Schwegler, E.; Galli, G.; Gygi, F. *J. Chem. Phys.* **2000**, *113*, 4668–4673.
- (227) Backus, E. H.; Schaefer, J.; Bonn, M. *Angew. Chem. Int. Ed.* **2021**, *60*, 10482–10501.
- (228) Zhang, Y.; de Aguiar, H. B.; Hynes, J. T.; Laage, D. *J. Phys. Chem. Lett.* **2020**, *11*, 624–631.
- (229) Adams, E. M.; Pezzotti, S.; Ahlers, J.; Rüttermann, M.; Levin, M.; Goldenzweig, A.; Peleg, Y.; Fleishman, S. J.; Sagi, I.; Havenith, M. *JACS Au* **2021**, *1*, PMID: 34337607, 1076–1085.
- (230) Ohto, T.; Dodia, M.; Xu, J.; Imoto, S.; Tang, F.; Zysk, F.; Kühne, T. D.; Shigeta, Y.; Bonn, M.; Wu, X., et al. *npj* **2019**, *10*, 4914–4919.
- (231) Crespo, Y.; Marinelli, F.; Pietrucci, F.; Laio, A. *Phys. Rev. E* **2010**, *81*, 055701.
- (232) Horinek, D.; Netz, R. R. *prl* **2007**, *99*, 226104.
- (233) Guardia, E.; Rey, R.; Padró, J. *cp* **1991**, *155*, 187–195.

Titre: De la solvation à la chimie des surfaces aqueuses: une approche combinée simulations DFT-MD et spectroscopie SFG

Mots clés: interfaces aqueuses, MD, spectroscopies SFG, hydrophobicité/hydrophilicité, chimie de surface, Silice

Résumé: L'hydrophobicité/hydrophilicité des interfaces aqueuses joue un rôle essentiel sur les propriétés physico-chimiques des interfaces, avec des applications allant de l'adsorption d'ions, l'électrocatalyse, au repliement des protéines. Nous avons développé une métrique pour quantifier l'hydrophobicité/hydrophilicité microscopique en combinant des simulations DFT-MD et la spectroscopie SFG expérimentale et théorique, qui pour la première fois établit les connexions entre les descripteurs structuraux et les quantités spectroscopiques. Cette nouvelle métrique a été appliquée aux interfaces SAMs/eau, révélant également le rôle des échelles nanométriques et nanosecondes des patches hydrophobes à la surface des PEGs. En outre, une nouvelle méthode, appelée "pop modèle" pour le calcul théorique rapide du sig-

nal SFG, est présentée. Elle permet à la fois de calculer et d'interpréter les spectres SFG des interfaces aqueuses pour la surface et l'eau. En utilisant cette méthode, les spectres SFG dépendant du pH de la surface de la silice en contact avec l'eau ont été construits pour la première fois sans effectuer de simulations MD dans différentes conditions de pH qui ne sont pas abordables en simulations atomistiques. Combinés aux simulations DFT-MD métadynamiques, les expériences et la spectroscopie SFG théorique ont mis en évidence une reconstruction de la surface supérieure de l'oxyde de silice déclenchée par la déprotonation des silanols de surface. Une espèce de silicium non conventionnelle à cinq coordinations $\text{Si}(5c)^-$ révolutionne ainsi la compréhension du comportement bimodal des acidités de surface de la silice.

Title: From solvation to chemistry of aqueous surfaces: a combined DFT-MD simulations and SFG spectroscopy approach

Keywords: aqueous interfaces, MD, SFG spectroscopy, hydrophobicity/hydrophilicity, surface chemistry, Silica

Abstract: Hydrophobicity/hydrophilicity of aqueous interfaces plays an essential role on interfacial physico-chemical properties, with applications ranging from ion adsorptions, electrocatalysis to protein folding. We have developed a metric for quantifying the microscopic hydrophobicity/hydrophilicity by combining DFT-MD simulations and both experimental and theoretical SFG spectroscopy, that for the first time built the connections between structural descriptors and spectroscopic quantities. This new metric has been applied to SAMs/water interfaces, also revealing the role of nano-scale and nano-seconds scales of hydrophobic patches at the surface of PEGs. Furthermore, a novel expedient method, named the "pop model" for theoretical SFG calculation is pre-

sented which allows at the same time calculating and interpreting SFG spectra of aqueous interfaces for both surface and water. Using this method, the pH-dependent SFG spectra of the silica surface in contact with water were for the first time constructed without performing MD simulations under different pH conditions that are not affordable in atomistic modeling. Combined with metadynamics DFT-MD simulations, experiments and theoretical SFG spectroscopy have uncovered a reconstruction of the silica oxide top-surface triggered by the deprotonation of the surface silanols. An unconventional five-coordinated silicon species $\text{Si}(5c)^-$ hence revolutionizes the understanding of the bimodal behavior of the silica surface acidities.