Supporting Information for "Free Energy of Proton Transfer at the Water - TiO_2 Interface from *Ab Initio* Deep Potential Molecular Dynamics"

Marcos F. Calegari Andrade,[a] Hsin-Yu Ko,[a] Linfeng Zhang[b], Roberto Car[a] and Annabella Selloni[a]*

[a] *Department of Chemistry, Princeton University, Princeton, NJ 08544, USA*

[b] *Program in Applied and Computational Mathematics,Princeton University, Princeton, NJ 08544, USA*

1 Deep neural network model and training

We trained our *ab initio*-based deep neural network (DNN) potential using the iterative procedure proposed by Zhang *et al.*[1]. As shown in Fig. S1, we started training a coarse DNN potential using computed DFT energies and forces for a small set of 100 equally spaced snapshots taken from *ab initio* simulations of bulk anatase TiO_2, liquid water and the anatase (101)–water interface[2]. This coarse DNN potential is then used to start an iterative training procedure consisting of: 1) exploration of the configurational space for each of the three systems using molecular dynamics simulations at different thermodynamic conditions; 2) evaluation of the maximum standard deviation (MSTD) on atomic forces predicted by DNN potentials trained with random initialization of their parameters; 3) first-principles calculations of energy and atomic forces for the configurations with MSTD larger than a pre-defined threshold; and 4) new training of the DNN after adding the new set of *ab initio* data to the reference data set. Convergence is achieved when the the mean absolute deviation of atomic forces, predicted during a 100 ps DPMD simulation of the anatase (101)–water interface at 330 K, falls bellow 0.05 eV/Å. At the exploration step of our DNN training, we sampled the following thermodynamic conditions for each system

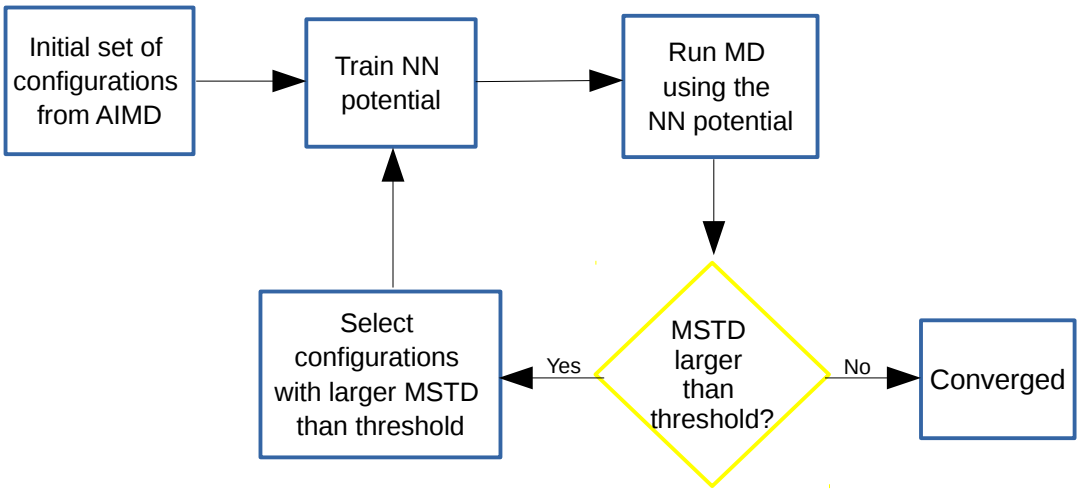

Figure S1: Neural network training procedure

from DPMD:

- anatase TiO_2 (108 atoms): $200-1200$ K at 1 bar;
- Water (192 atoms): $200-400$ K at 1 bar, and $400-800$ K at constant volume (experimental density);
- anatase (101)–water interface (426 atoms): $200-800$ K at constant volume.

Given a converged DNN potential, we further refined it by including configurations collected from enhanced sampling simulations. We adopted an iterative procedure similar to that shown in Fig. S1, but in this case our convergence criterion was based on the free energy barrier for water dissociation.

Our converged DNN potential contains a training set with the following number of configurations: 769 for anatase TiO_2; 5583 for water; 250 for water adsorbed on the anatase (101) in vacuum; and 15248, for the anatase (101)–water interface, for which configurations with some of the interfacial water molecules either dissociated or near the transition state were included.

2 *Ab initio* calculations

First-principles energies and forces were computed with the DFT functional SCAN[3], as implemented in the `PWscf` code of `Quantum ESPRESSO`[4,5] and the `Libxc` package[6]. Norm-conserving pseudopotentials of TM[7] (oxygen and hydrogen) and RRKJ[8] (titanium) types replaced core-valence electron interactions. Kohn-Sham wavefunctions were represented in plane-wave basis including terms up to 110 Ry kinetic energy. Only the Γ point of the Brillouin zone was sampled. A detailed study on the accuracy of the SCAN functional to predict structural properties of TiO_2 can be found in the supporting information of[2].

3 DPMD performance relative to AIMD

In this section we provide a more extensive comparison between results given by DPMD and AIMD with the density functional SCAN. All comparisons are performed using the 1×3 anatase (101) surface supercell described in the main text.

Vibrational properties. We compare the vibrational densities of states (VDOS) of water and TiO_2 as obtained from 40 ps of AIMD[2] and DPMD simulations of the TiO_2–water interface. Interfacial water is discretized into layers, as defined in a previous study by some of the present authors[2], with higher layer indexes referring to water molecules at larger distances from the TiO_2 surface. As shown in Fig. S2, DPMD correctly captures the evolution of the vibrational spectrum of water as a function of the distance from the surface, in good agreement with AIMD. The VDOS of the TiO_2 slab obtained from DPMD also agrees with the *ab initio* spectrum (Fig. S3), showing the ability of DPMD to reproduce the *ab initio* potential energy surface of both water and TiO_2.

Water dissociation on TiO_2. We performed calculations of the minimum energy path to dissociate a water molecule on the anatase (101) surface in vacuum using the Nudged Elastic Band (NEB) method[9]. A single water molecule was adsorbed on one side of a 5 layers 1×3 anatase (101) slab with a 15Å vacuum separating periodic images along the surface normal. As shown in Fig. S4, the NEB-DNN calculations agree, within the deviation of four DNN potentials, with the results from DFT. In particular, our DNN potentials

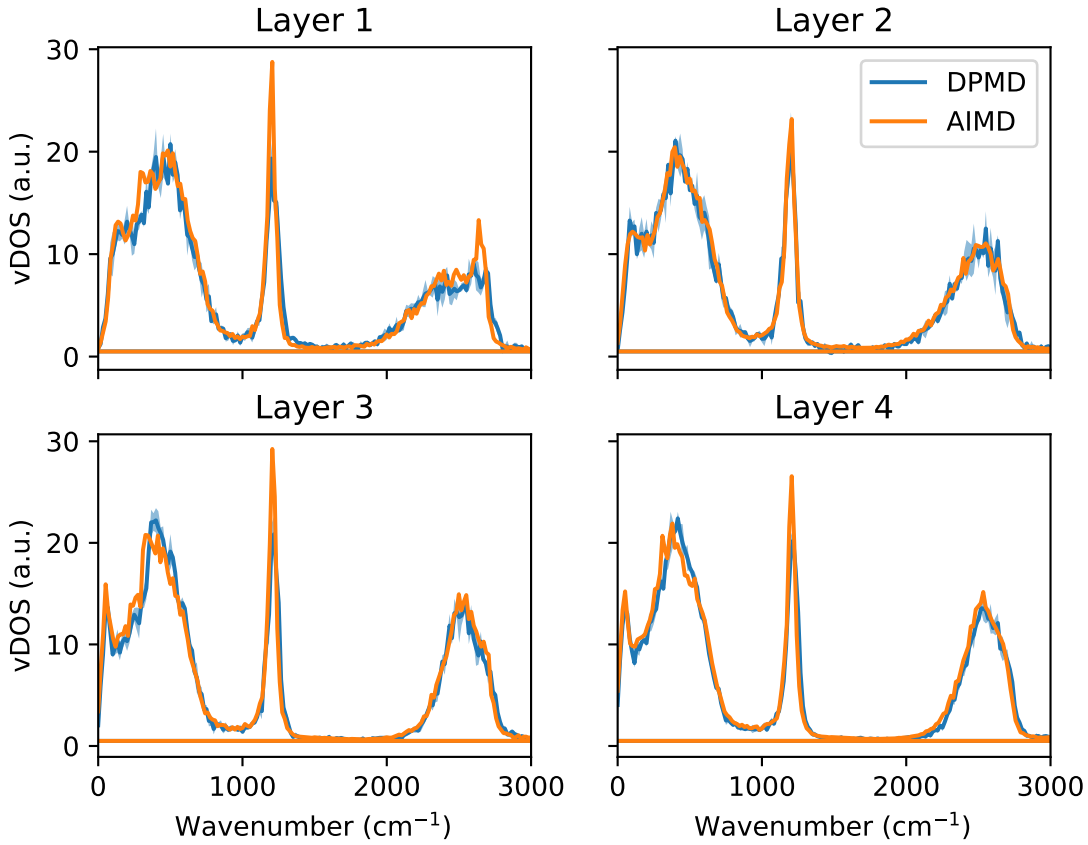

Figure S2: Vibrational density of states of water at the TiO_2–water interface. Water layers are defined as in Ref.[2]. Shaded areas indicate the standard deviation obtained from four independent DPMD simulations.

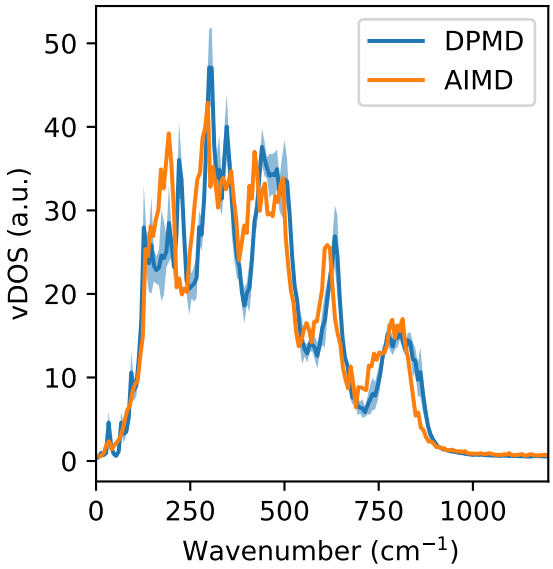

Figure S3: Vibrational density of state of the TiO_2 slab used to simulate the TiO_2–water interface. Water layers are defined as in ref.[2]. Shaded areas indicate the standard deviation obtained from four independent DPMD simulations.

reproduce the energy difference between molecular and dissociative adsorption of water as well as the energy barrier separating these two stable states given by DFT.

We further compared the DNN and *ab initio* results for water dissociation at the anatase (101)–water interface, by evaluating the work to move a H^+ from a TiO_2 undercoordinated oxygen (bridging hydroxyl) to a nearby terminal hydroxyl group. Atomic coordinates along the reaction path were taken from the umbrella sampling described in section 4 (see below). The work was computed as the integral of the force projected on the unit vector connecting a O_{2c} to the nearest H atom. As shown in Fig. S4, the DPMD results are in close agreement with DFT.

4 Umbrella sampling

As mentioned in the main text, we computed the free energy of water dissociation on anatase from a two-step procedure. In the first step we evaluated the free energy ($F(S_O)$) as a function of the minimum distance, S_O, between a particular surface oxygen atom

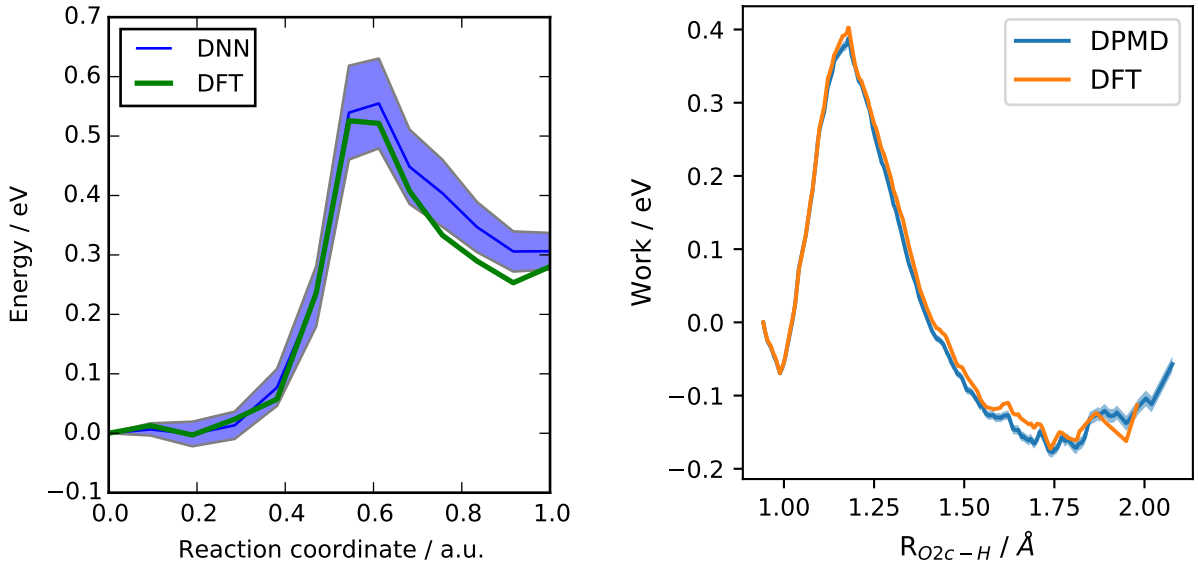

Figure S4: **Left**: Minimum energy path to dissociate an adsorbed water molecule on the anatase (101) surface in vacuum into a terminal and a bridging hydroxyl. **Right**: Work to move a H^+ from a surface O_{2C} to an OH_t adsorbed on an adjacent Ti_{5c} at the anatase (101)-water interface. Shaded areas indicate the standard deviation obtained from four different DNN potentials.

(O) and any hydrogen atom in the system. Here we give further details on the umbrella sampling technique used to evaluate $F(S_O)$.

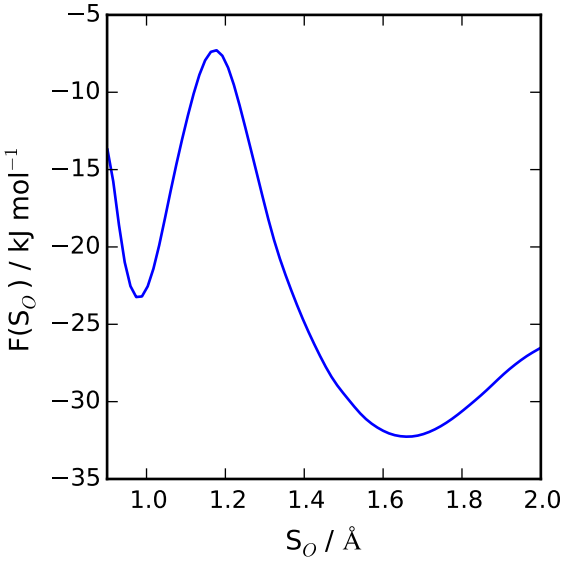

Figure S5: Free energy ($F(S_O)$) as a function of the minimal distance S_O of a surface O_{2c} to any hydrogen in the system.

Our umbrella sampling simulations consisted of 9 independent DPMD simulations of the 1×3 anatase (101) in contact with a 20 Å slab of water. In each of these simulations, we applied a quadratic (or umbrella) potential $V(S_O) = k(S_O - S_O^{fix})^2$ centered at one of the following values of S_O^{fix}, in Å (with corresponding value of the force constant k given in parenthesis, in kJ mol^{-1} Å^{-2}): 1.0 (200), 1.1 (4000), 1.2 (4000), 1.25 (4000), 1.3 (4000), 1.4 (2000), 1.5 (2000), 1.6 (500), 1.7 (100). The probability distribution of S_O was re-weighted using the WHAM method[10]. The resulting free energy curve is given in Fig. S5.

5 Lattice gas model

To obtain an additional independent estimate of the equilibrium coverage of dissociated water at the anatase (101)–water interface, we used a simple lattice gas model, where each site corresponds to a surface water adsorption site (Ti_{5c}) that is occupied by either a molecular or dissociated water. We assume the adsorption energy of molecular water to

be 8.0 kJ/mol more favorable than dissociated water, as given by our enhanced sampling simulations (Figure 4 in the main text). Our model also includes a constraint present in the real system: no two water molecules can donate H^+ to the same O_{2c} site on the surface. This constraint prevents water dissociation at Ti_{5c} sites where the two adjacent O_{2c}'s are already protonated, decreasing even further the probability of water dissociation at high surface hydroxylation. Our model relies on the negligible correlation between water dissociation at neighboring adsorption sites on anatase (101). On this surface, the distance between neighboring Ti_{5c} sites is indeed quite large, 3.8 Å, thus allowing only weak H-bonds (≈ 2.5 Å) between water or hydroxyl groups adsorbed on such sites. As a result, the water dissociation free energy is largely independent from the degree of surface hydroxylation (see Fig. S5). Similarly, water adsorption on the anatase (101) surface in vacuum depends negligibly on the surface coverage below one monolayer[11].

Our lattice gas model has the following Hamiltonian:

$$H = \begin{cases} \varepsilon \sum_{i=1,j=1}^{N,M} |n_{i,j}|, & \text{if } n_{i,j} - n_{i,j+1} \neq 2 \\ \infty, & \text{otherwise} \end{cases} \quad (1)$$

where ε is the adsorption free energy difference of dissociated vs. molecular water at the anatase (101)–water interface, and $n_{i,j}$ represents the adsorption state of water at site (i,j). We set $\varepsilon = 8.0$ kJ/mol, $n_{i,j} = 0$ for adsorbed molecular water, and $n_{i,j} = \pm 1$ for dissociated water. where $n_{i,j} = -1$ and $n_{i,j} = 1$ represent a dissociated water donating a H^+ to an O_{2c} atom on the left and on the right, respectively (see Fig. S6). The above Hamiltonian prohibits the formation of doubly protonated O_{2c} sites through the energy constraint on the adsorption state $n_{i,j} = 1, n_{i,j+1} = -1$. In this model, rows ($i$) and columns ($j$) represent adsorption sites along the [010] and [$\bar{1}$01] directions of the anatase TiO_2 crystal lattice, respectively. A model with 96 adsorption sites (as in the 4x12 anatase surface supercell used for DPMD simulations) was represented by $N = 12$ columns and $M = 8$ rows. The probability $P(C)$ of observing the surface with C hydroxyl sites was numerically evaluated with Metropolis Monte Carlo sampling, and the total number of hydroxyls is defined by

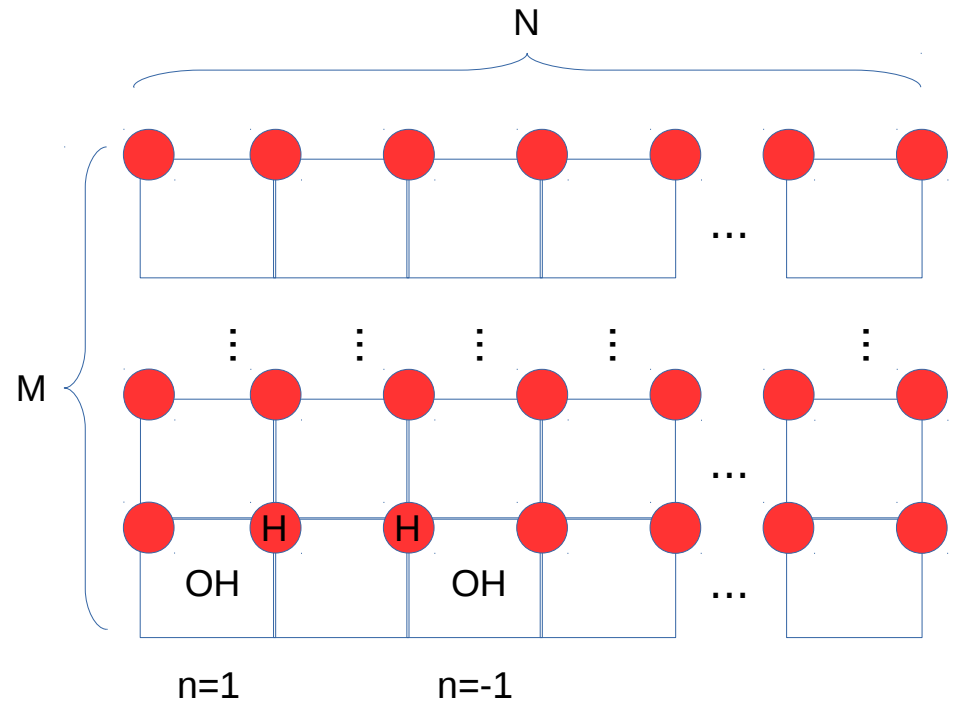

Figure S6: Scheme of the lattice gas model used to compute the free energy of the anatase (101)–water interface as a function of the surface hydroxylation coverage. Empty red circles represent bare O_{2c} sites, while a red circle with a "H" represents a bridging hydroxyl group. Empty squares represent adsorbed molecular water, and squares with a "OH" represent terminal hydroxyl groups on top of Ti_{5c} sites. The model contains $N \times M$ water adsorption sites.

$C = \sum_{i,j} |n_{i,j}|$. The temperature was set at 330 K, the same used in our DPMD simulations. At each Monte Carlo step, a trial change in an adsorption state $n_{i,j}$ randomly selected a molecular ($n_{i,j} = 0$) or dissociative ($n_{i,j} = \pm 1$) state with equal probability. The move was accepted/rejected based on the Metropolis criteria.

From the free energy surface $F(n)$ presented in Fig. S7, we estimated a 5% equilibrium hydroxyl coverage from $\int_0^1 e^{-\beta F(n)} n dn / \int_0^1 e^{-\beta F(n)}$.

Although water dissociation is in itself thermodynamically unfavored, the configurational entropy contribution gives rise to a finite probability of observing dissociated water.

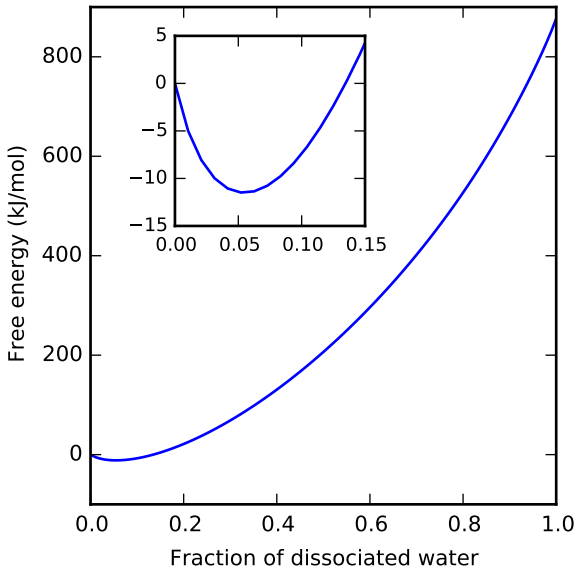

Figure S7: Free energy of the lattice gas model as a function of the fraction of dissociated water. Inset restricts the dissociation fraction to the range from 0 to 0.15 in order to show the minimum in free energy. The lattice gas model was evaluated on a 96 sites model, the same number of water adsorption sites as our DPMD simulations.

6 Water confinement effect

All simulations reported in the main text were performed for a 2 nm water slab confined between TiO_2 surfaces. With this geometry, the density profile of water at the center of the water slab was found to converge to the equilibrium density of bulk water in AIMD simulations with the SCAN functional[2]. To further test the influence of the water slab thickness, we compared DPMD simulations for 2 and 4 nm confined water slabs. As shown in Fig. S8, the density profile of water changes negligibly by increasing the water slab thickness from 2 nm to 4 nm. This indicates that a 2 nm water slab should be sufficient to properly simulate the anatase (101)–water interface with a negligible confinement effect.

Notes and references

[1] L. Zhang, D.-Y. Lin, H. Wang, R. Car and W. E, *Physical Review Materials*, 2019, **3**, 023804.

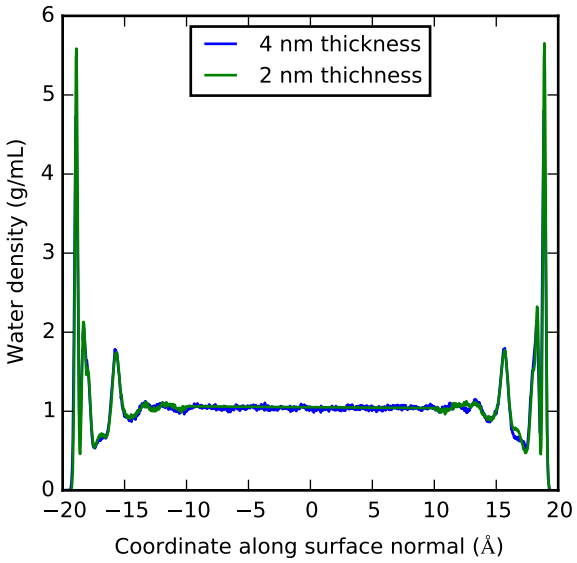

Figure S8: Comparison of the water density profiles obtained from DPMD simulations of anatase (101)–water with 2 nm and 4 nm water slab confined between anatase (101) surfaces. The water density profiles for 2 and 4 nm were matched close to the TiO_2 surfaces.

[2] M. F. Calegari Andrade, H.-Y. Ko, R. Car and A. Selloni, *Journal of Physical Chemistry Letters*, 2018, **9**, 6716–6721.

[3] J. Sun, A. Ruzsinszky and J. P. Perdew, *Physical Review Letters*, 2015, **115**, 036402.

[4] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari and R. M. Wentzcovitch, *Journal of Physics: Condensed matter*, 2009, **21**, 395502.

[5] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. Dal Corso, S. De Gironcoli, P. Delugas, R. A. Distasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia,

M. Kawamura, H. Y. Ko, A. Kokalj, E. Kücükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H. V. Nguyen, A. Otero-De-La-Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu and S. Baroni, *Journal of Physics: Condensed Matter*, 2017, **29**, 465901.

[6] S. Lehtola, C. Steigemann, M. J. Oliveira and M. A. Marques, *SoftwareX*, 2018, **7**, 1–5.

[7] N. Troullier and J. L. Martins, *Physical Review B*, 1991, **43**, 1993–2006.

[8] A. M. Rappe, K. M. Rabe, E. Kaxiras and J. D. Joannopoulos, *Physical Review B*, 1990, **41**, 1227–1230.

[9] G. Henkelman, B. P. Uberuaga and H. Jónsson, *Journal of Chemical Physics*, 2000, **113**, 9901–9904.

[10] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, *Journal of Computational Chemistry*, 1992, **13**, 1011–1021.

[11] A. Vittadini, A. Selloni, F. P. Rotzinger and M. Grätzel, *Physical Review Letters*, 1998, **81**, 2954–2957.