

Improved Sampling in Ab Initio-Based Free Energy Calculations of Amino Acids at Solid-Liquid Interfaces: A Tight-Binding Assessment on TiO₂ Anatase (101)

Lorenzo Agosta, Erik G. Brandt, Alexander Lyubartsev

Submitted date: 07/06/2019 • Posted date: 07/06/2019

Licence: CC BY-NC-ND 4.0

Citation information: Agosta, Lorenzo; Brandt, Erik G.; Lyubartsev, Alexander (2019): Improved Sampling in Ab Initio-Based Free Energy Calculations of Amino Acids at Solid-Liquid Interfaces: A Tight-Binding Assessment on TiO₂ Anatase (101). ChemRxiv. Preprint.

Atomistic simulations are powerful for probing molecules at bioinorganic interfaces and excellent complements to scarcely available experimental techniques. The free energy controls the adsorption behavior of molecules on nanosurfaces, and is therefore a quantity of particular importance. Advanced sampling techniques can efficiently explore the adsorption free energy landscape, but molecular simulations with classical (Newtonian) dynamics fail to capture charge transfer and polarization at the solid-liquid interface. First principle simulations do not suffer from this limitation but come with a heavy computational load. Here, we introduce an efficient protocol to explore the free energy of adsorption in the ab initio framework. This approach accurately models the complex phenomena at bio-inorganic surfaces on the nanoscale and properly samples the relevant thermodynamic properties. We present a case study of adsorption of the Lysine and Aspartate amino acids on the anatase (101) TiO₂ surface with the tight binding method. The high values of the calculated adsorption free energies highlight the importance of a proper description of the electronic state for surface binding processes.

File list (1)

draft.pdf (5.15 MiB)

[view on ChemRxiv](#) • [download file](#)

Improved sampling in ab initio-based free energy calculations of amino acids at solid-liquid interfaces: A tight-binding assessment on TiO_2 anatase (101)

Lorenzo Agosta¹, Erik G. Brandt¹ and Alexander P. Lyubartsev¹

¹ *Department of Materials and Environmental Chemistry,
Stockholm University, S-10691 Stockholm, Sweden*

Atomistic simulations are powerful for probing molecules at bioinorganic interfaces and excellent complements to scarcely available experimental techniques. The free energy controls the adsorption behavior of molecules on nanosurfaces, and is therefore a quantity of particular importance. Advanced sampling techniques can efficiently explore the adsorption free energy landscape, but molecular simulations with classical (Newtonian) dynamics fail to capture charge transfer and polarization at the solid-liquid interface. First principle simulations do not suffer from this limitation but come with a heavy computational load. Here, we introduce an efficient protocol to explore the free energy of adsorption in the ab initio framework. This approach accurately models the complex phenomena at bio-inorganic surfaces on the nanoscale and properly samples the relevant thermodynamic properties. We present a case study of adsorption of the Lysine and Aspartate amino acids on the anatase (101) TiO_2 surface with the tight binding method. The high values of the calculated adsorption free energies highlight the importance of a proper description of the electronic state for surface binding processes.

INTRODUCTION

Nano-molecular engineering of inorganic materials is the focus of many scientific applications in the field of nano-biotechnology [1–3] such biomimetics[4], optics[5], biosensor [6, 7] and smart surface materials[8]. The correct characterization of the mechanisms that governs the adsorption of biomolecules with nanostructures represents the ultimate goal in this field. Bio-nanointerfaces are typically defined as restricted area around a nanoparticle, a corona of about 1 nm, where many complex phenomena that regulates the adhesion occur. This region is usually very difficult to access experimentally due to the weak signal generated from such small volume compared with the remained system. The adsorption mechanism on inorganic surfaces like TiO_2 , very suitable for its bio-compatible properties [9, 10], is regulated by different factors. Generally the under-coordinated metal atoms like titanium and the bridging oxygen atoms O_{br} are reactive sites for biomolecules [11], thus the specific interactions between the adsorbate and the oxide surfaces is the principal driven force. Nevertheless oxide surfaces are covered by strongly adsorbed layers of water molecules that compete for the same adsorption sites.[12, 13] Semi-conductor surfaces like TiO_2 can also split naturally water generating a variegate population of hydroxyl groups on the surface that contribute to the adsorption process.[14, 15]. Finally charge transfer phenomena or strong polarization effects can modify significantly the interaction energies. The binding free energy of biomolecules adhering on TiO_2 nanostructures is a fundamental quantity to be evaluated in this context and it requires a careful evaluation of all the energetic contributions generated at the interfaces. Simple models based on coulomb attraction and Lennard-Jones dispersion as implemented in classical MD (CMD) are

unable to catch this complexity. Attentive evaluations of the adsorption free energies were computed to assess the adsorption hierarchy of amino acids interacting with different TiO_2 surfaces [16–19] or amorphous nanoparticles [20]. The general outcome of this studies strongly relies on the possibility for the amino acids to directly adsorb on the TiO_2 surface, penetrating the strong adsorbed layers of water molecules [21]. The force fields used for describing TiO_2 -adsorbates interactions are normally parametrized for neutral small inorganic molecules, discarding polarization effects due to the strong charge localization at the surfaces, such protonation states. Further the water density profiles at the TiO_2 interface calculated in the DFT [21] and CMD frameworks differ significantly, providing either an over-structured water region [19] that completely hinders the direct molecule adhesion, or a too soft barrier that always allows the direct interaction adsorbate-surface [16–18]. Such controversial description of the adsorption mechanism, that lacks a validation against experimental set ups, can be only resolved by first principle calculations. Ab-initio approaches can properly take in account the reactivity generated at the bio-inorganic interfaces but sampling free energy requires extended MD simulations for good convergence, that are normally unfeasible in the context of first principle calculations. Hence the necessity to design a proper protocol to adopt free energy techniques within the ab-initio level of theory. We focused in this work on the computation of adsorption free energies of Lysine (Lys) and Aspartate (Asp) on the hydrated TiO_2 anatase (101) surface whose adsorption energies were calculated in the context of DFT by Agosta et. al. [11], showing a very strong interaction that could not be ascribed to coulombic forces alone. We implemented an efficient and novel protocol to obtain accurate evaluations of the binding free energy of Lys and Asp within a Tight Binding (TB) approach for time

scales that can be also accessed by higher level of theory such density functional theory (DFT). This methodology allows to underpin the real adhesion properties for complex environments like liquid-inorganic interfaces.

METHOD

All the computations were done through the CP2K software[22]. In our knowledge no studies of fully hydrated inorganic surfaces simulated in the TB framework are reported. In order to benchmark the available TB parameters we tested the DFTB libraries as implemented in the Matsci[23] and Mio-Tiorg [24] parametrization. We report that in both cases the amino acids in contact with the TiO_2 anatase (101) surface undergo to nonphysical deprotonation effects of both $C-H$ and $N-H$ groups, suggesting an underestimation of the interacting potential in the overlapping integrals. In order to reproduce the correct protonation state of the amino acids we implemented an harmonic constrain for all the $C-H$ and $N-H$ groups. For the anatase (101) Matsci parameters provide a better agreement with the structure derived from DFT calculation, while with Mio-Tiorg a geometrical distortion arises at the interface. For this reason all the free energy calculation are done using the Matsci parameters. No long range dispersion were used in the calculations.

The simulation boxes were prepared with PACKMOL [25] software. The box sizes were fixed at $10.35 \times 11.4 \times 43$ Å for anatase (101). 4 slabs of TiO_2 were used along the z direction as for previous simulations and the remaining space was filled with a single amino acid and water at its relative density for 1 atm and 310 K. [21] Aspartate and Lysine were implemented in their analogue form, that means considering just their side-chain and substituting the backbone with a CH_3 group. This set up avoids self interactions with first-neighbour periodic images and it resembles the interaction the amino acids would have in a peptide, where the backbone doesn't contribute to the adsorption process. The system total charge were neutralized introducing a counter ion in the form of OH^- or OH_3^+ placed in the opposite surface side with respect the adsorption side of the amino acid.

Metadynamic

Meta-dynamic simulations were done coupling PLUMED software[26] to CP2K. Well tempered adaptive techniques (WTA-MetaD) were adopted [27–29] with a bias factor of 15. Starting with a height of 3.5 kJ/mol Gaussians were added every 25 fs and the widths were updated every 75 ps. These parameters were demonstrated to be a good compromise for ab-initio

Metadynamics simulations, were the accessible sampling is strongly limited in time.

In order to speed up the phase space sampling we made use of 8 multi-walkers [30] starting from different initial configurations. A potential wall at 1 nm away from the outermost line of Ti atoms was introduced in order to restrict the sampled phase space. Each walker was run interchanging information with the remaining walkers every 25 fs. Two independent collective variables were taken in consideration for exploring the free energy landscape. The surface separation distance (SSD) variable defined as the distance between the outermost layer of Ti atoms of the TiO_2 surface and the center of mass of the active group in each amino acid. Using the NH_3^+ and COO^- groups respectively for Lysine (Lys) and Aspartate (Asp) instead of the center of the mass of the whole molecule allows to directly asses the adsorption modes[16]. The second variable was defined as the angle between the normal vector to the TiO_2 surface and an internal vector defined by a couple of atoms in Lys and Asp. This variable was used in order to enhance the sampling of the phase space and it was integrated out in the calculation of the free energy. Up to 240-300 ps could be simulated for each walker.

The standard potential of the mean force (PMF) was calculated in the contest of WTA-MetaD, by recovering the Gaussians history [19, 27, 29]:

$$PMF(z) = \lim_{t \rightarrow \infty} (V(z, t) + \sigma(z, t) + cost.) \quad (1)$$

where $V(z, t)$ is the time recovered bias potential. $\sigma(z, t)$ is the accumulated histogram of the reaction coordinate up to time t , a correction needed when the gaussian width is not constant in time [29].

It has been showed that combining MetaD with the force estimator for calculating the free energy enhances the convergence when the collective variable (CV) is sampled in all its relevant regions [31]. For this reason we implemented an alternative calculation of the PMF based on the thermodynamic integration of the mean forces acting on the NH_3^+ and COO^- atoms along the collective variable SSD sampled during the MetaD:

$$PMF_{TI}(z) = - \int_{r_c}^{r_c + \delta} \langle F(z) \rangle dz \quad (2)$$

where z spans the path of the collective variable from r_c (SSD at minimum value) and $r_c + \delta$ (SSD in bulk solution). δ is the length sampled by the collective variable. We will refer to this combination of MetaD plus force integration as MetaDF in the text.

The final binding free energy is computed as: [19]

$$\Delta G_b = -kT \ln \left(\frac{1}{\delta} \int_{r_c}^{r_c + \delta} e^{\frac{-PMF(z)}{kT}} dz \right) \quad (3)$$

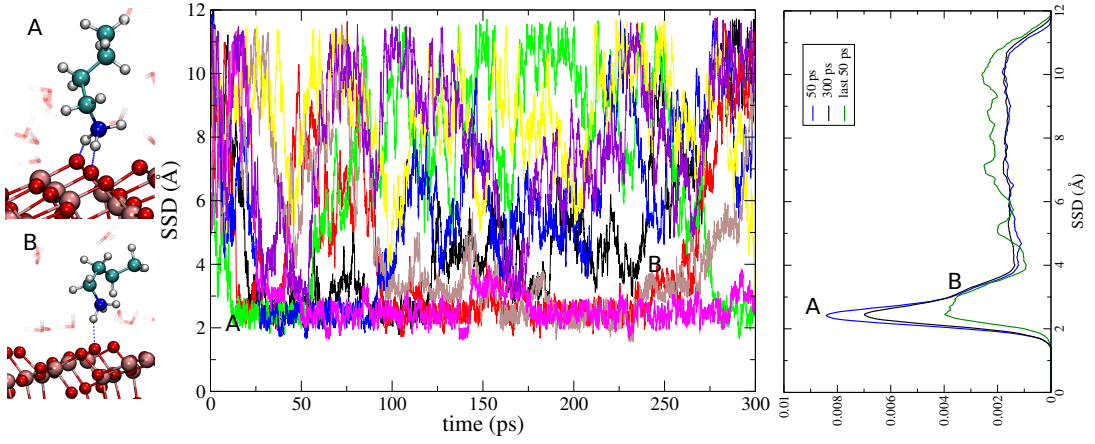


FIG. 1: Trajectories of the 8 walkers projected along the SSD variable for Lysine. The multi-walker approach boosts the collective variable sampling that is well scanned in only 50 ps. On the right panel the distribution of SSD is plotted at different times. No significant difference is shown between 50 and 200 ps; 2 adsorption modes can be distinguished on the anatase (101) surface at $SSD = 2 \text{ \AA}$ and 4 \AA where the flat freely diffusive layer begins.

Although Eq. 3 is the right way to compute the ΔG_b , we noted that related works [16–18] reporting the adsorption free energy of amino acids on TiO_2 surfaces consider only the difference between the minimum and the bulk values in the PMF:

$$\Delta F_{ads} = PMF(r_c + \delta) - PMF(r_c) \quad (4)$$

This absolute difference is always higher than the ΔG_b and it is independent of the chosen path explored by the CV.

RESULTS AND DISCUSSION

Amino-acids possessing charged side chains are considered to be the main responsible for bio-inorganic adhesion processes [11, 32]. From the computational point of view they also are the most challenging cases to be studied, as they can generate strong polarization effects at the interface or induce changes to the protonation state. We selected Lysine and Aspartate, that are respectively positive and negatively charged amino acids in physiological conditions, as representative examples to investigate how a full electron description can affect the calculation of the adsorption free energy.

In Fig. 1 the walkers dynamic along the SSD variable for Lysine is plotted. Two different adsorption modes are identified (A and B) corresponding to bidentate and monodentate adsorption on the O_{br} oxygen atoms. The full phase space described by the collective variable is explored and different walkers are able to penetrate the first layer of adsorbed water molecules linking directly to the TiO_2 surface. The adsorption/desorption dynamic is also sampled different times describing an equilibrium

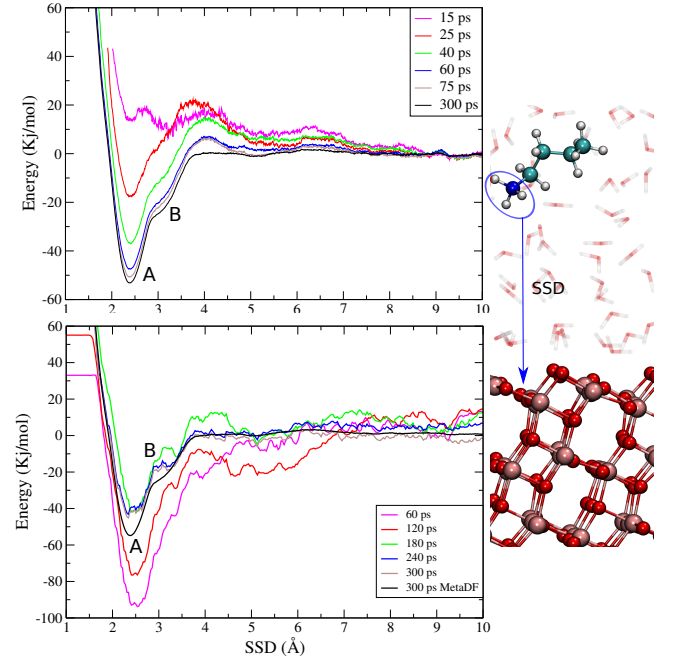


FIG. 2: PMF for Lysine amino acid adsorbing on TiO_2 anatase (101) surface. (Top) PFM obtained by integrating the mean force (MetaDF) acting on NH_3^+ group along the SSD variable (left). Bottom PFM obtained by standard WTA-MetaMD. After 60 ps the PMF profile converges within an error of 3KJ/mol by using MetaDF while standard MetaMD is highly fluctuating after 300 ps.

state and no energy barrier are observed from the bulk to reach the TiO_2 surface. The peak in the distribution of SSD variable (Fig. 1) indicates that the convergence is not reached by the WTA-MetaD.

In Fig. 2 the adsorption free energy profiles along the SSD variable for Lysine adsorbing on TiO_2 anatase (101)

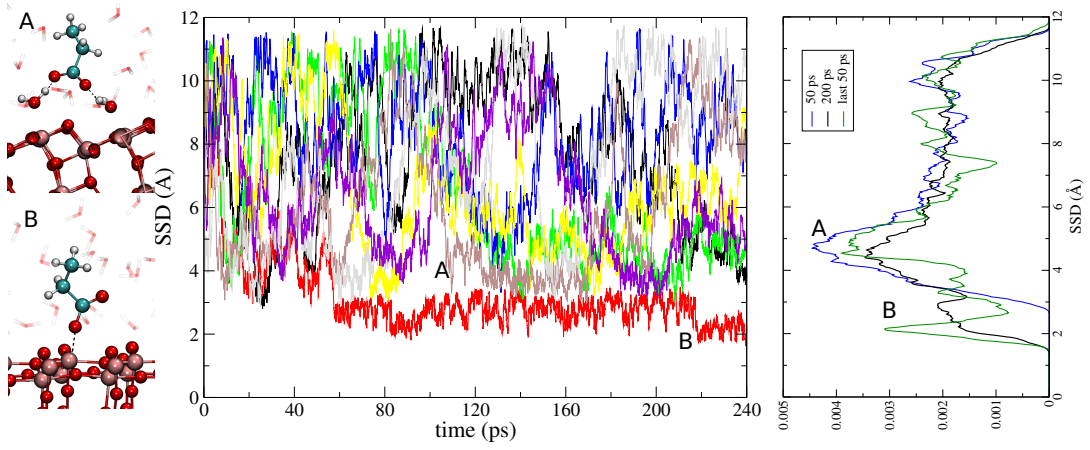


FIG. 3: Trajectories of the 8 walkers projected along the SSD variable for Aspartate. The first adsorption mode A is well sampled after 50 ps. The second adsorption mode has the does not contribute to the adsorption free energy.

surface is reported. The PMF are plotted on intervals of time up to 300 ps per walker in order to study the convergence. After about 60 ps the potential of mean force obtained by integrating the average forces on the NH_3^+ group reaches a smooth profile within an error of 3 KJ/mol. On the contrary standard WTA-MetaD presents an oscillating trend after 300 ps of simulation per walker. The calculated binding free energy is about 43.3 KJ/mol, while the ΔF_{ads} is 54 KJ/mol. These values are about 3 to 10 times higher than any other reported value where only CMD were implemented [16–19]. This high value indicates that important contributions to the total energy derive from the polarization of the electron density at the interface. Further studies aimed to clarify this point are necessary and will be carried on in further studies.

Two distinct adsorption modes are present for the Aspartate molecule as well (Fig. 3 and 4). The first one A arising from the mediation of the first adsorbed water layer and the second one B as a direct interaction with the TiO_2 surface. The PMF in Fig. 4 demonstrates the gain in efficiency by using the MetaDF method with respect to the standard WTA-MetaD. The mode A is well sampled with MetaDF just after 40 ps of simulation with an error 3 KJ/mol, while AWT-MetaD is strongly oscillating after 240 ps. The mode B is appearing after 60 ps and it is explored only by one walker, making its convergence slower. The energy barrier between the 2 modes is about 22 KJ/mol and it derives from the possibility of penetrating the first layer of adsorbed water molecules. The main contribution to the binding free energy is given by the mode A and it is about 5.6 KJ/mol while the ΔF_{ads} is about 17 KJ/mol. Such value for the adsorption of the COO^- group was found for a direct contact to the surface [33] or an adsorption via an hydroxyl group at the surface Sultan et al.. This indicates the fundamental importance of having a good model the reproduces at the same time the water structure at the interface and the

polarization state.

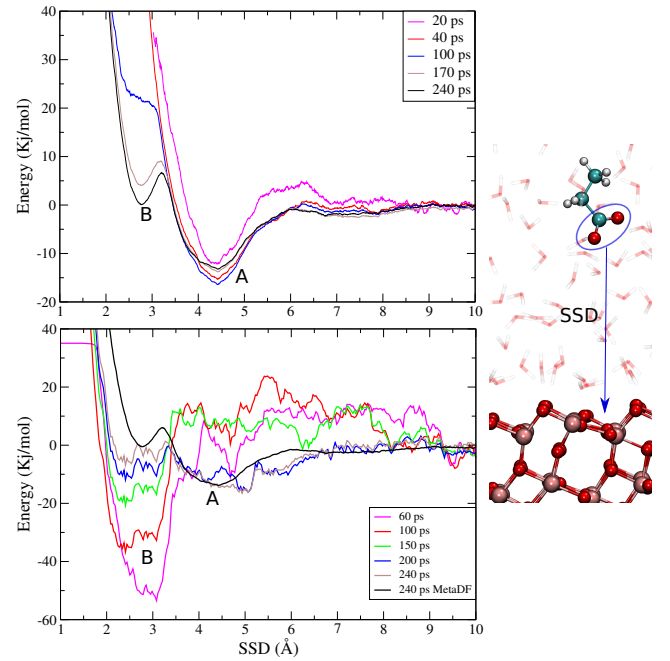


FIG. 4: PMF for Aspartate amino acid adsorbing on TiO_2 anatase (101) surface. (Top) PFM obtained by integrating the mean force (MetaDF) acting on COO^- group along the SSD variable (left). Bottom PFM obtained by standard WTA-MetaD . The 2 minima (A,B) define 2 adsorbing modes of which only B contributes to the free energy. After 60 ps B converges by using MetaDF while standard MetaD is highly fluctuating after 200 ps.

We showed that using the force estimator in the WTA-MetaD framework is the "key" for sampling the free energy landscape with first principle level of theory. While WTA-MetaD is limited by diffusion and converges in the long-time limit due the continuously decreasing of

the bias factor, thermodynamic integration methods, like umbrella sampling, are limited by the choice of an optimal continuous collective variable, unknown a priori. Combining these 2 techniques boosts the exploration of the phase space creating a concatenation of states for continuously integrating the mean force acting on the CV, that is not affected by the bias potential. The bias potential make possible to rapidly explore the phase space on which the forces can be extrapolated and averaged in time. While the forces fluctuations are continuously collected, the gaussian insertion is operated every 25 fs. Once an optimal path for the collective variable is explored the force integration method is at least in the order of 50 times faster in convergence than the standard MetaD. In general this method can be expand to all the cases where the sampling is strongly hindered by diffusion as for strong binding states.

CONCLUSIONS

In conclusion we provided an efficient method for sampling binding free energies when the simulation time, needed for good convergence, is strongly hindered like in the context of ab-initio calculations. For strong binding energies as the ones generated on the TiO_2 adsorption sites it is fundamental to establish a protocol for enhancing the rate of sampling in the bulk and in the adsorption state. Defining the interacting atoms of the amino acids as collective variable and using multi-walkers dynamic assured to span effectively all the important regions of the phase space. To evaluate the PMF we combined the MetaD calculations with the force integration approach that boosted the time required to converge to just 40-60 ps of simulation per walker. This small time window is accessible also for density functional theory simulations and allows to assess the adsorption free energy taking into account all the electronic effects generated at the bio-inorganic interfaces that are normally discarded by using classical MD. We further benchmarked the Tight-Binding parameters to describe the adsorption of amino acids at the liquid-inorganic interface of the TiO_2 anatase (101). We demonstrated that a strong enhancement of the adsorption strength can be ascribed to a proper description of the electronic state at the surface. We expect this methodology to be a mile stone for a thorough and systematic investigation of the bio-nano interfaces.

-
- [1] K. Shiba, Curr. Op. Biotech. **21**, 412 (2010).
 - [2] J. D. Hartgerink, E. Beniash, and S. I. Stupp, Science **294**, 1684 (2001).
 - [3] L. P. B. O. D. J. P. D. P. . T. J. D. Andrew P. Nowak, Victor Breedveld, Nature **417**, 424 (2002).

- [4] A. K. Y. J. K. S. . F. B. Mehmet Sarikaya, Candan Tamerler, Nature Materials **2**, 577 (2003).
- [5] R. G. Ellis-Behnke, Y.-X. Liang, S.-W. You, D. K. C. Tay, S. Zhang, K.-F. So, and G. E. Schneider, Proceedings of the National Academy of Sciences **103**, 5054 (2006).
- [6] D. Khatayevich, T. Page, C. Gresswell, Y. Hayamizu, W. Grady, and M. Sarikaya, Small **10**, 1505 (2014).
- [7] M. Yemini, M. Reches, J. Rishpon, and E. Gazit, Nano Letters **5**, 183 (2005).
- [8] E. V. and D. V. Andreeva, Advanced Functional Materials **23**, 4483 (2013).
- [9] Progress in Materials Science **54**, 397 (2009), ISSN 0079-6425.
- [10] D.-P. Song, M.-J. Chen, Y.-C. Liang, Q.-S. Bai, J.-X. Chen, and X.-F. Zheng, Acta Biomaterialia **6**, 684 (2010), ISSN 1742-7061, URL <http://www.sciencedirect.com/science/article/pii/S1742706109003171>.
- [11] L. Agosta, G. Zollo, C. Arcangeli, F. Buonocore, F. Gala, and M. Celino, Phys. Chem. Chem. Phys. **17**, 1556 (2015).
- [12] J. Schneider and L. C. Ciacchi, Journal of the American Chemical Society **134**, 2407 (2012).
- [13] A. A. Skelton, T. Liang, and T. R. Walsh, ACS Applied Materials & Interfaces **1**, 1482 (2009).
- [14] A. D. Roddick-Lanzilotta and A. McQuillan, Journal of Colloid and Interface Science **227**, 48 (2000), ISSN 0021-9797, URL <http://www.sciencedirect.com/science/article/pii/S0021979700968644>.
- [15] A. D. Roddick-Lanzilotta, P. A. Connor, and A. J. McQuillan, Langmuir **14**, 6479 (1998).
- [16] A. YazdanYar, U. Aschauer, and P. Bowen, The Journal of Physical Chemistry C **122**, 11355 (2018).
- [17] A. M. Sultan, Z. E. Hughes, and T. R. Walsh, Langmuir **30**, 13321 (2014).
- [18] S. Monti and T. R. Walsh, The Journal of Physical Chemistry C **114**, 22197 (2010).
- [19] E. G. Brandt and A. P. Lyubartsev, The Journal of Physical Chemistry C **119**, 18126 (2015).
- [20] J. M. P.-A. Shengtang Liu, Xuan-Yu Meng and R. Zhou, Scientific Reports **6** (2016).
- [21] L. Agosta, E. G. Brandt, and A. P. Lyubartsev, The Journal of Chemical Physics **147**, 024704 (2017).
- [22] J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, Wiley Interdisciplinary Reviews: Computational Molecular Science **4**, 15 (2014).
- [23] R. Luschtinetz, J. Frenzel, T. Milek, and G. Seifert, The Journal of Physical Chemistry C **113**, 5730 (2009).
- [24] G. Dolgonos, B. Aradi, N. H. Moreira, and T. Frauenheim, Journal of Chemical Theory and Computation **6**, 266 (2010).
- [25] E. G. B. J. M. M. L. Martnez, R. Andrade, Journal of Computational Chemistry **30(13)**, 2157 (2009).
- [26] D. B. C. C. G. B. G.A. Tribello, M. Bonomi, Comp. Phys. Comm. **185** (2014).
- [27] A. Barducci, G. Bussi, and M. Parrinello, Phys. Rev. Lett. **100**, 020603 (2008), URL <https://link.aps.org/doi/10.1103/PhysRevLett.100.020603>.
- [28] O. Valsson, P. Tiwary, and M. Parrinello, Annual Review of Physical Chemistry **67**, 159 (2016).
- [29] D. Branduardi, G. Bussi, and M. Parrinello, Journal of Chemical Theory and Computation **8**, 2247 (2012).
- [30] P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti, and

- M. Parrinello, The Journal of Physical Chemistry B **110**, 3533 (2006).
- [31] M. A. Cuendet and M. E. Tuckerman, Journal of Chemical Theory and Computation **10**, 2975 (2014).
- [32] T. Hayashi, K.-I. Sano, K. Shiba, Y. Kumashiro, K. Iwahori, I. Yamashita, and M. Hara, Nano Letters **6**, 515 (2006).
- [33] C. Li, S. Monti, and V. Carravetta, The Journal of Physical Chemistry C **116**, 18318 (2012).
- [34] A. M. Sultan, Z. C. Westcott, Z. E. Hughes, J. P. Palafox-Hernandez, T. Giesa, V. Puddu, M. J. Buehler, C. C. Perry, and T. R. Walsh, ACS Applied Materials & Interfaces **8**, 18620 (2016).

draft.pdf (5.15 MiB)

[view on ChemRxiv](#) • [download file](#)
