Supporting Information

Constant Potential Electrochemistry with Multiscale Quantum Mechanical / Machine Learning Simulations

Seung-Jae Shin*[1], Kara D. Fong[2], and Aron Walsh[3]

[1] School of Energy and Chemical Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea

[2] Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA

[2] Department of Materials, Imperial College London, Exhibition Road, SW7 2AZ, London, United Kingdom

*e-mail: seungjae.shin@unist.ac.kr

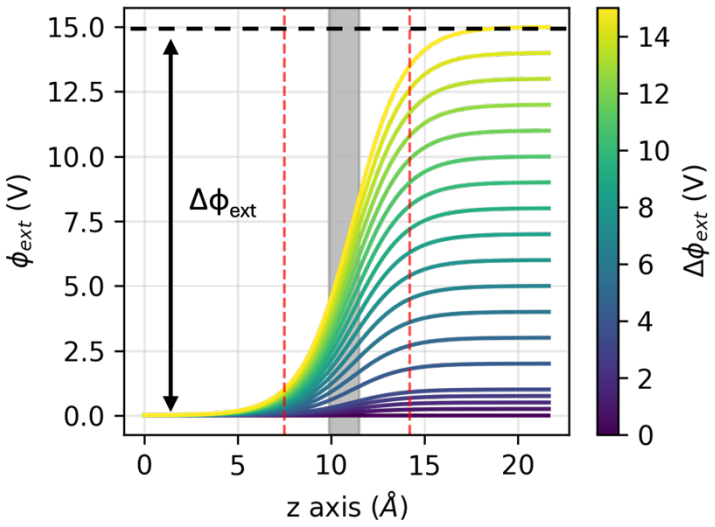
It includes

Supplementary Note 1-3

Supplementary Figure. 1-12

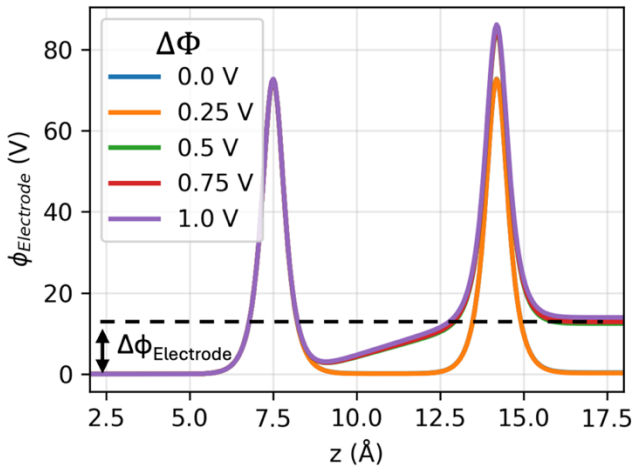
Supplementary Note 1. Validation of QM/ML results with DFT calculations

The QM/ML results were validated against DFT calculations. In these validation tests, the water molecules were described either by DFT or by a machine learning force field plus point charges (MLFF+Q). We optimized the water molecule structures by applying an external sigmoid-shaped potential, ϕ_{ext}. The initial geometry for all optimizations was taken from a snapshot of the QM/ML simulation at a converged total potential, $\Delta\Phi$, of 0.0 V. The external potential difference, $\Delta\phi_{ext}$, was varied, ranging from 0.0 V to 15.0 V (Supplementary Note Figure 1).



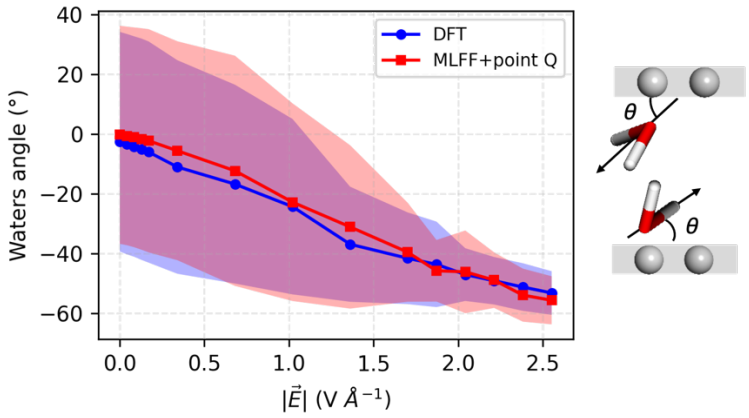
Supplementary Note Figure 1. The external potential profile, ϕ_{ext}, is shown for various potential differences, $\Delta\phi_{ext}$. The red dashed lines indicate the position of the graphene electrode during the QM/ML simulations. They serve as a guideline to show how this sigmoid-shaped potential well reproduces the electrostatic potential from the electrode. The grey shaded region indicates where the water molecules exist for this simulation.

The range of the sigmoid-shaped external potential was determined by referring to the electrostatic potential difference of the electrode $\Delta\phi_{Electrode}$ during the constant potential simulation, which spans 0 V to 15 V (Supplementary Note Figure 2). These potentials were achieved in the QM/ML simulations after $\Delta\Phi$ of the system converged to 0.0 V, 0.25 V, 0.5 V, 0.75 V and 1.0 V, respectively. For the two lowest potentials 0 V and 0.25 V, $\Delta\phi_{Electrode}$ closely matches $\Delta\Phi$, indicating subtle electrolyte response. However, for the higher converged $\Delta\Phi$ as 0.5 V, 0.75 V, and 1.0 V, the $\Delta\phi_{Electrode}$ values are significantly larger as 12.5 V, 12.9 V, and 14.0 V, respectively. This increase is due to the screening response of the water molecules during the QM/ML simulations.

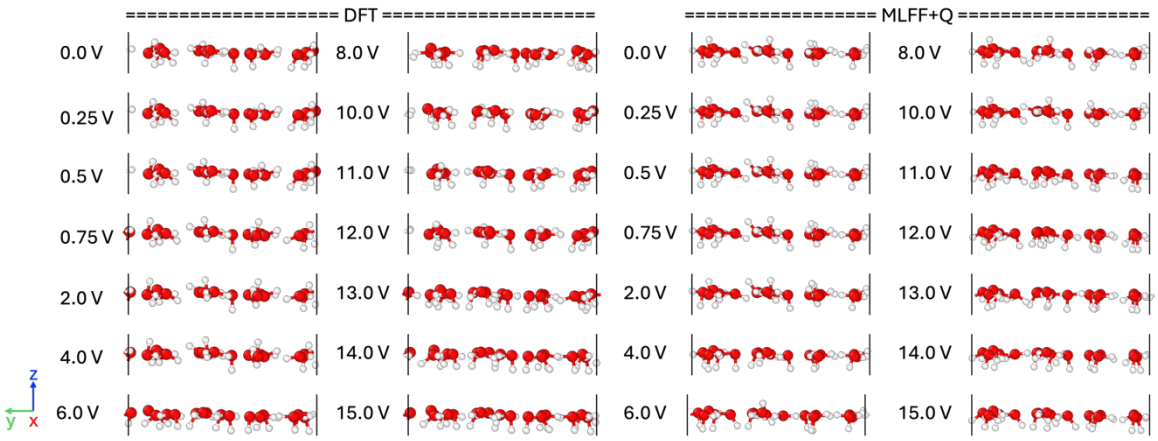


Supplementary Note Figure 2. Electrostatic potential profiles from the electrode, $\phi_{Electrode}$, are plotted as a function of converged total potential, $\Delta\Phi$, calculated during the QM/ML simulations. $\Delta\phi_{Electrode}$ is defined as the difference in $\phi_{Electrode}$ values between the cathode and anode. The legends specify the $\Delta\Phi$, which incorporates the electrostatic potential from the both the electrode and electrolyte in the QM/ML simulations.

The optimized water molecule structures obtained using DFT and MLFF+Q show excellent agreement. This consistency is further reflected in the average dipolar angles of the water molecules across the two methods (Supplementary Note Figures 3 and 4).

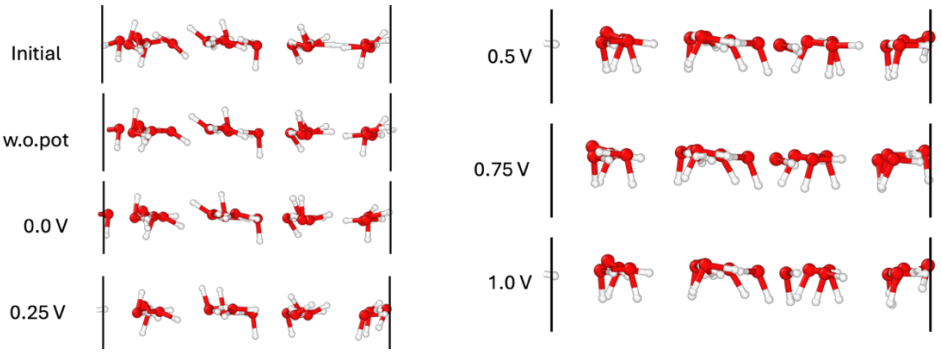


Supplementary Note Figure 3. The average angle of water molecules is plotted as a function of the magnitude of electric field for the grey shaded region in Supplementary Note Figure 1. The corresponding standard deviation is indicated by the shaded region. The angle is defined as that between the vector connecting the oxygen atom to the midpoint of the hydrogen atoms and the electrode surface.



Supplementary Note Figure 4. Optimized structures of water molecules under an external sigmoid-shaped potential at various potential differences, $\Delta\phi_{ext}$. The data are presented in four columns: Columns 1 and 2 show results where the electrolyte was described using DFT, and columns 3 and 4 show results where the electrolyte was described using MLFF plus point charges (MLFF+Q).

A consistent result is achieved using an external potential derived from the polarized graphene electrodes, where the electrolyte components are described with DFT. The external potential for this simulation is shown in Supplementary Note Figure 2. The optimized structures using these external potentials are presented in Supplementary Note Figure 5. Each optimization started from the same initial geometry, specifically the snapshot obtained from the converged QM/ML simulation at $\Delta\Phi$ as 0.0 V. For external potentials corresponding to the $\Delta\Phi$ of QM/ML simulations as 0.0 V and 0.25 V, the resulting optimized structures show small change from the initial state, indicating subtle response from the electrolyte. However, when using the external potentials corresponding to higher $\Delta\Phi$ above 0.5 V, the optimized structures change dramatically to effectively screen the external potential with the electrolyte's dipole moment. This structural response is fully consistent with the findings from the QM/ML simulations (Figure S10).

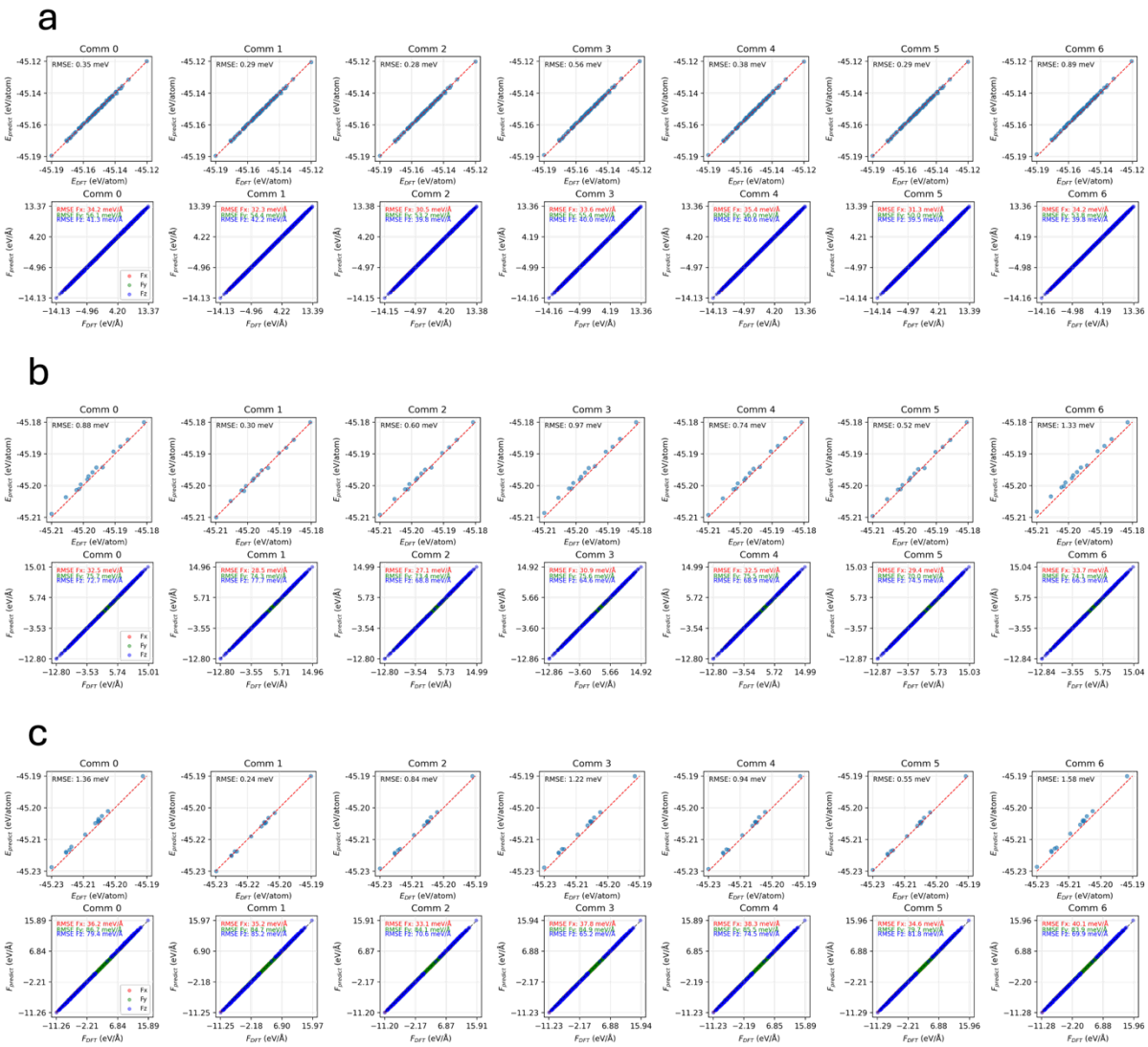


Supplementary Note Figure 5. Optimized electrolyte structures obtained from DFT calculations under the influence of an external potential shown in Supplementary Note Figure 2. The labels on the figure indicate the converged total potential difference $\Delta\Phi$ achieved in the QM/ML simulations.

In our validation simulations, it is notable that water dissociation was not observed despite the application of high external potentials. This outcome is likely a consequence of the simulation methodology: the validation involved only geometry optimization, and the positions of the oxygen atoms were fixed. Furthermore, water dissociation was also absent in the primary QM/ML simulations because the MLFF was not trained for this regime (i.e., rare events like bond breaking are outside its current scope). In real dynamical situations, however, water molecules would be expected to dissociate under these external potential values. Therefore, investigating the relationship between the degree of electrolyte dissociation and the magnitude of the external potential will be an interesting and critical subject for future work.

Supplementary Note 2. Comparison of different point charge models

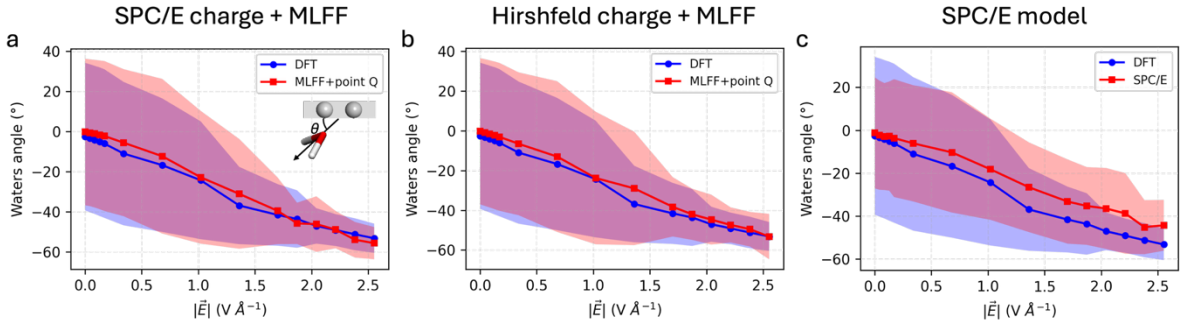
We followed the same procedure as before but replaced the SPC/E model's charge values with Hirshfeld charges. First, we performed DFT simulations using a bulk water model and determined the atomic charges for the water molecules to be -1.1 for Oxygen and +0.55 for Hydrogen. After assigning these values, we decoupled the electrostatic interactions from the DFT datasets and trained new MLFF models based on this updated database. The training set size is identical to that of the models employing SPC/E charges, which includes bulk electrolytes, neutral confined electrolytes, and the confined electrolytes at ΔΦ is 0.5 V and 1.0 V. To evaluate performance, both interpolation and extrapolation tests were also conducted using the structures at ΔΦ is 0.25 V, 0.75 V, and 1.5 V (Supplementary Note Figure 6).



Supplementary Note Figure 6. Test results for 7 committee ML models on the initial dataset. The models were trained using the dataset including bulk and confined electrolyte structures, both with and without applying the potentiostat. The ΔΦ in the training data were 0.5 V and 1.0 V. (**a**) The test dataset consisted of confined electrolyte structures, also without a potentiostat. (**b**) The test dataset consisted of confined electrolyte structures, with a potentiostat where ΔΦ is 0.25 V and 0.75 V. (**c**) The test

dataset consisted of confined electrolyte structures, with a potentiostat where $\Delta\Phi$ is 1.5 V. The energy and force prediction are compared against reference DFT simulation results.

To validate the QM/ML results against DFT calculations, we investigated the average angle of optimized monolayer water molecules under a constant external electric field, as described in Supplementary Note 1. We found that the optimized structures obtained via both methods are in excellent agreement with the DFT results but the standard SPC/E model produces errors of up to ~10° when the external field increases to 2.5 V Å^{-1} (Supplementary Note Figure 7). This indicates that the model is robust to variations in point charges, as the MLFF accurately captures the underlying interactions provided the charges remain within an acceptable range.

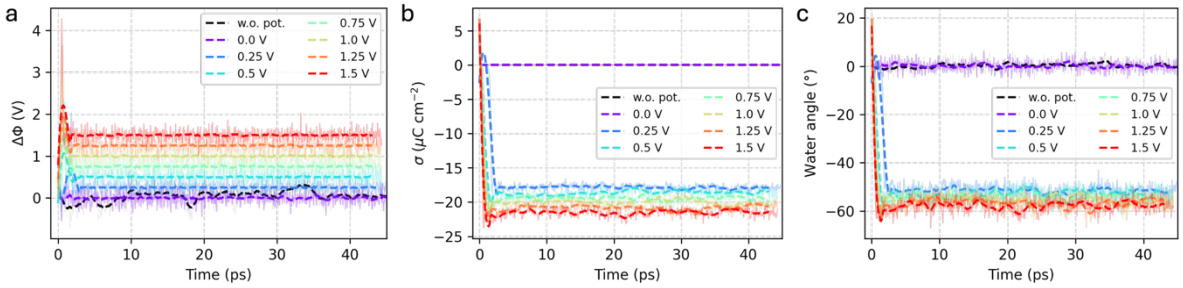


Supplementary Note Figure 7. The average orientation angle of water molecules is plotted as a function of the magnitude of electric field for the grey shaded region in Supplementary Note Figure 1. Each panel shows results employing a different model: (**a**), SPC/E charges with MLFF; (**b**), Hirshfeld charges with MLFF; and (**c**), a standard SPC/E model. The corresponding standard deviation is indicated by the shaded region. The angle is defined as that between the vector connecting the oxygen atom to the midpoint of the hydrogen atoms and the electrode surface.

Furthermore, we employed Hirshfeld charges and corresponding MLFF models to generate a constant-potential ensemble. A stable ensemble was successfully maintained, reproducing the significant potential fluctuations characteristic of systems without a potentiostat (Supplementary Note Figure 8). However, under a constant-potential ensemble, the interface polarization depends on the chosen charge model. This discrepancy stems from the distinct dipole moments of individual molecules. Although the electrolyte exhibits a consistent response to the external field across different charge models, the resultant potential drop varies because each model yields a different screening electric field for the same configuration, leading to a different definition of the total potential drop. For instance, since the Hirshfeld charge model assigns a higher dipole moment to water molecules (3.1 D compared to 2.3 D for the SPC/E model), the interface becomes more strongly polarized. Consequently, we observed that the electrode charging onset potential shifted by 0.25 V, while the surface charge and water reorientation angles were approximately 1.5 times higher due to the increased dipole moment.

These findings indicate the need for more rigorous future investigations using diverse charge sets. Such efforts should be complemented by experimental measurements or

constant-potential AIMD simulations once these methodologies are further developed and established, helping to identify an optimal model that reproduces the most realistic ensemble. Furthermore, integrating polarizable models, such as the recently developed on-the-fly charge update schemes, is a promising direction for eliminating the sensitivity to fixed charge values in our methodology.



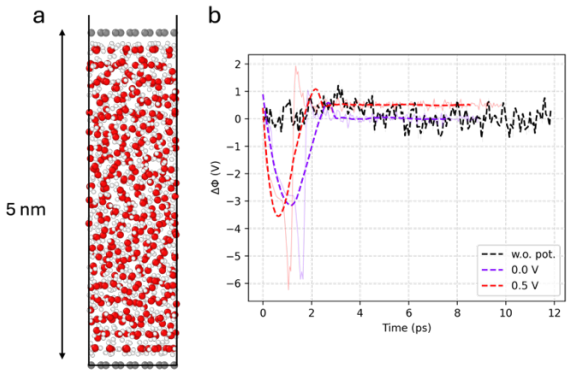
Supplementary Note Figure 8. Temporal evolution of the interface using Hirshfeld charges with monolayer model. (**a**) ΔΦ, (**b**) surface charge (σ), and (**c**) average orientation angles of water molecules as a function of time is plotted for various target potential difference. Averaged lines are displayed as a thicker curve after applying a 2.5 ps moving average.

Supplementary Note 3. Scalability of the constant potential QM/ML method

To demonstrate the length scale scalability, we conducted simulations on a 5 nm system consisting of 442 water molecules confined between bilayer graphene (Supplementary Note Figure 9). In light of Poisson-Boltzmann theory, the diffuse layer width at 0.01 M is approximately 3 nm; thus, this demonstrates that the QM/ML method can utilize a sufficiently large model to capture the intricate details of the electrochemical interface. For both 0.0 V and 0.5 V cases, the potential difference converged to the target values within 3 ps. Without a potentiostat, the potential fluctuated up to 1.0 V. Furthermore, we successfully extended the distance between graphene layers to 10 nm, confirming that our framework can handle even larger simulation cells.

Regarding the timescale and computational efficiency, the advantages of our QM/ML scheme become even more evident when considering the associated costs. While conventional DFT calculations typically scale as N^3, where N is the number of atoms, our MLFF scales almost linearly. For instance, a single DFT run just for the electrode (N=120) costs approximately 1 minute. This means it would require at least 29 hours for a single DFT run if the electrolyte atoms were included together in the 5 nm system (N=1446). In contrast, our QM/ML scheme is significantly more efficient; a single ML run takes only ~0.01 s, and even with the overhead of QM/ML iterations, obtaining information for each snapshot takes only ~2 minutes. Every simulation was conducted on a single CPU node equipped with 128 cores. For our longest trajectory of 150 ps at $\Delta\Phi$ = 0.5 V employing the monolayer model, the total computational time was approximately 2.1 days. Even when scaling to a 5 nm cell, the single ML run time costs only ~0.1 s. Although the DFT component takes slightly longer (~3.5 min) to treat the larger cell, the total cost for a 10 ps run is only about 0.5 days.

While we acknowledge that the current MLFF was trained primarily on small-scale systems, these simulations serve as a proof-of-concept to demonstrate that our framework can conceptually scale to significantly larger systems. To rigorously develop a production-level MLFF specifically for such massive systems, it would be necessary to construct new training datasets derived from larger simulation cells. We anticipate that generating these extensive databases will be the most computationally demanding phase of the development process. Nevertheless, this direction remains a primary focus of our future work to further enhance the efficiency and accuracy of large-scale electrochemical simulations.



Supplementary Note Figure 9. Constant potential ensemble in a large system (**a**) Snapshot of the 5 nm system model. (**b**) Time evolution of $\Delta\Phi$ for the 5 nm model.

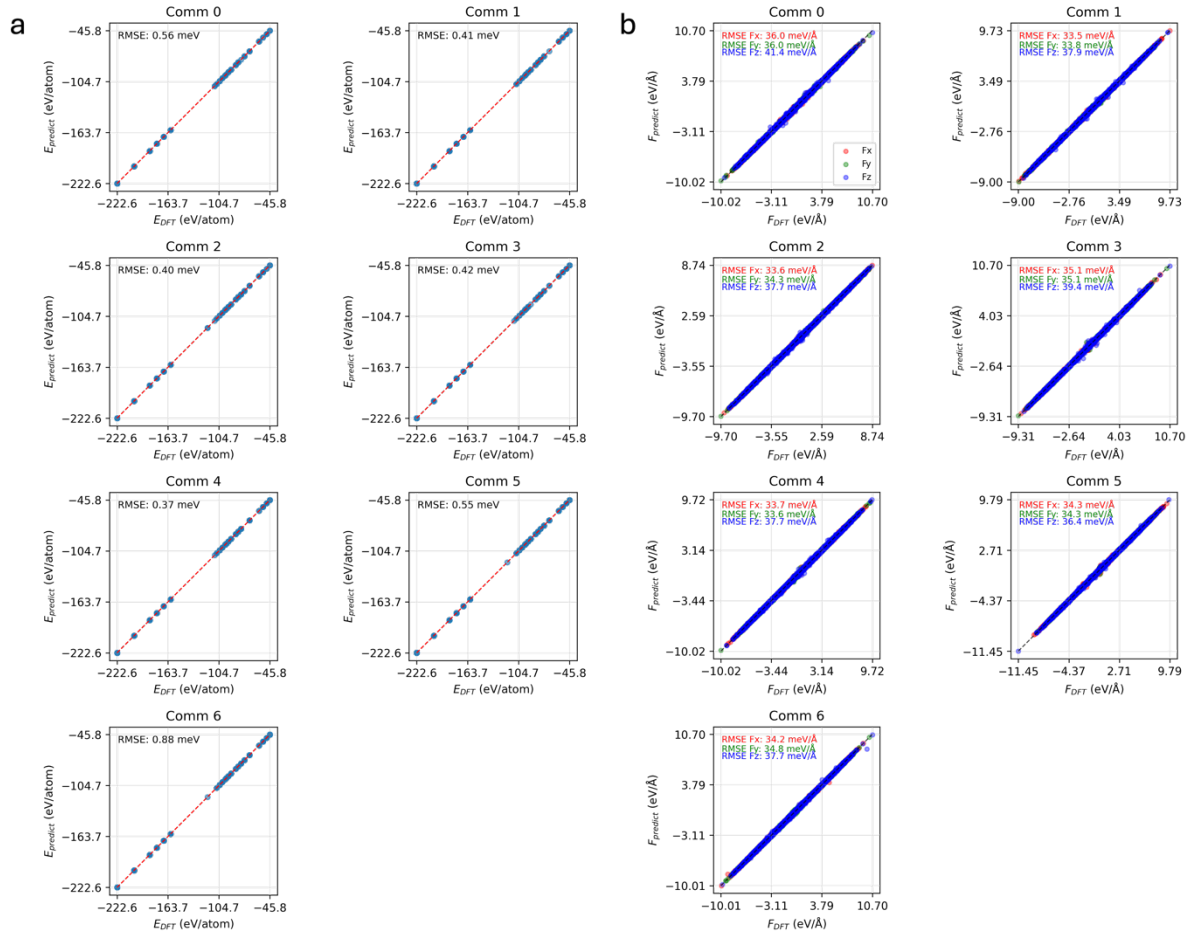


Figure S1. Test results for 7 committee ML models on the initial dataset. The models were trained using a dataset comprising bulk and confined electrolyte structures simulated without a potentiostat. The test dataset consisted of confined electrolyte structures, also without a potentiostat. (**a**) The energy prediction and (**b**) the force prediction are compared against reference DFT simulation results.

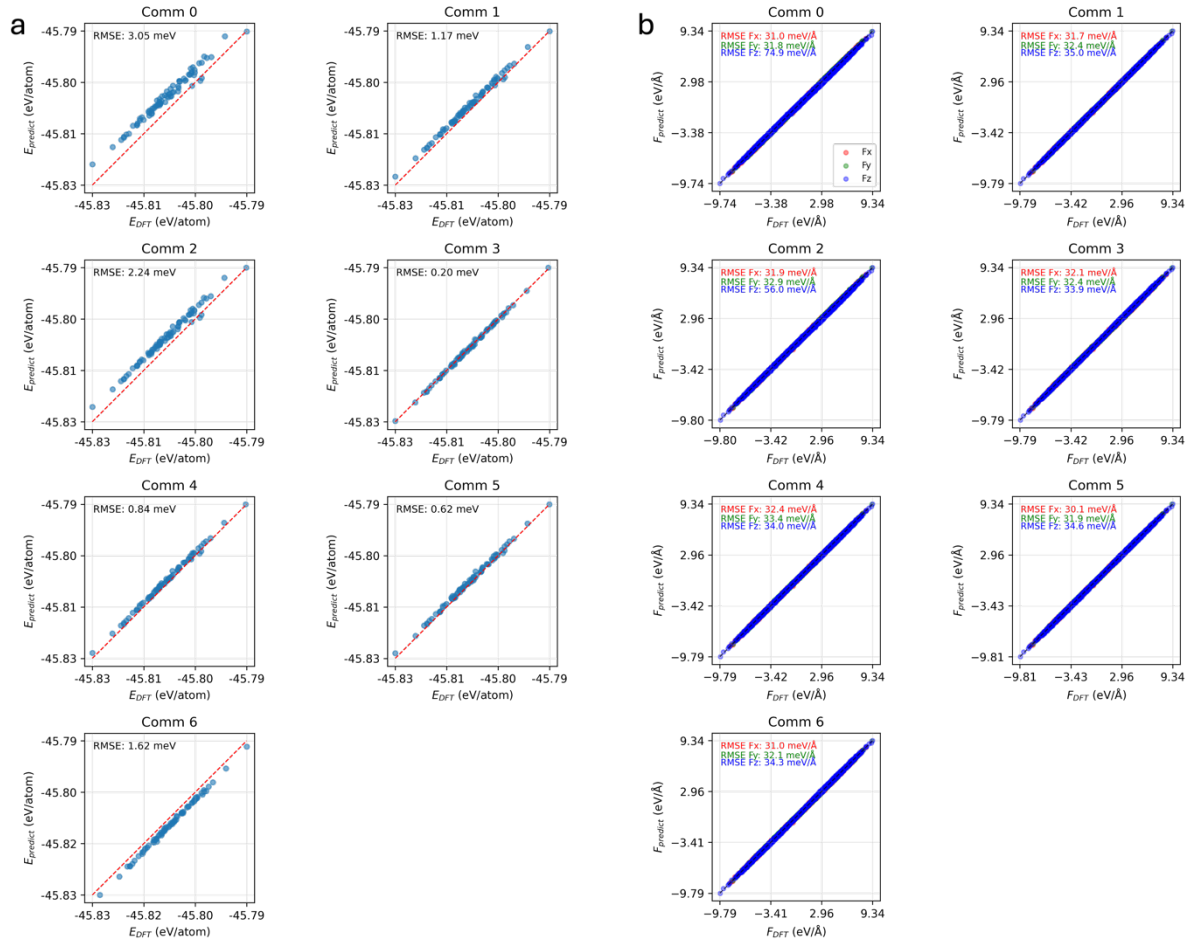


Figure S2. Test results for 7 committee ML models before training with the dataset applying the potentiostat. The models were initially trained using structures of bulk and confined electrolytes simulated without a potentiostat. The test dataset consisted of confined electrolyte structures sampled from a QM/ML simulation achieving 0.5 V as $\Delta\Phi$ via a potentiostat. (**a**) The energy prediction and (**b**) the force prediction is compared against the reference DFT calculations, which were performed on the trajectories sampled from the QM/ML simulations.

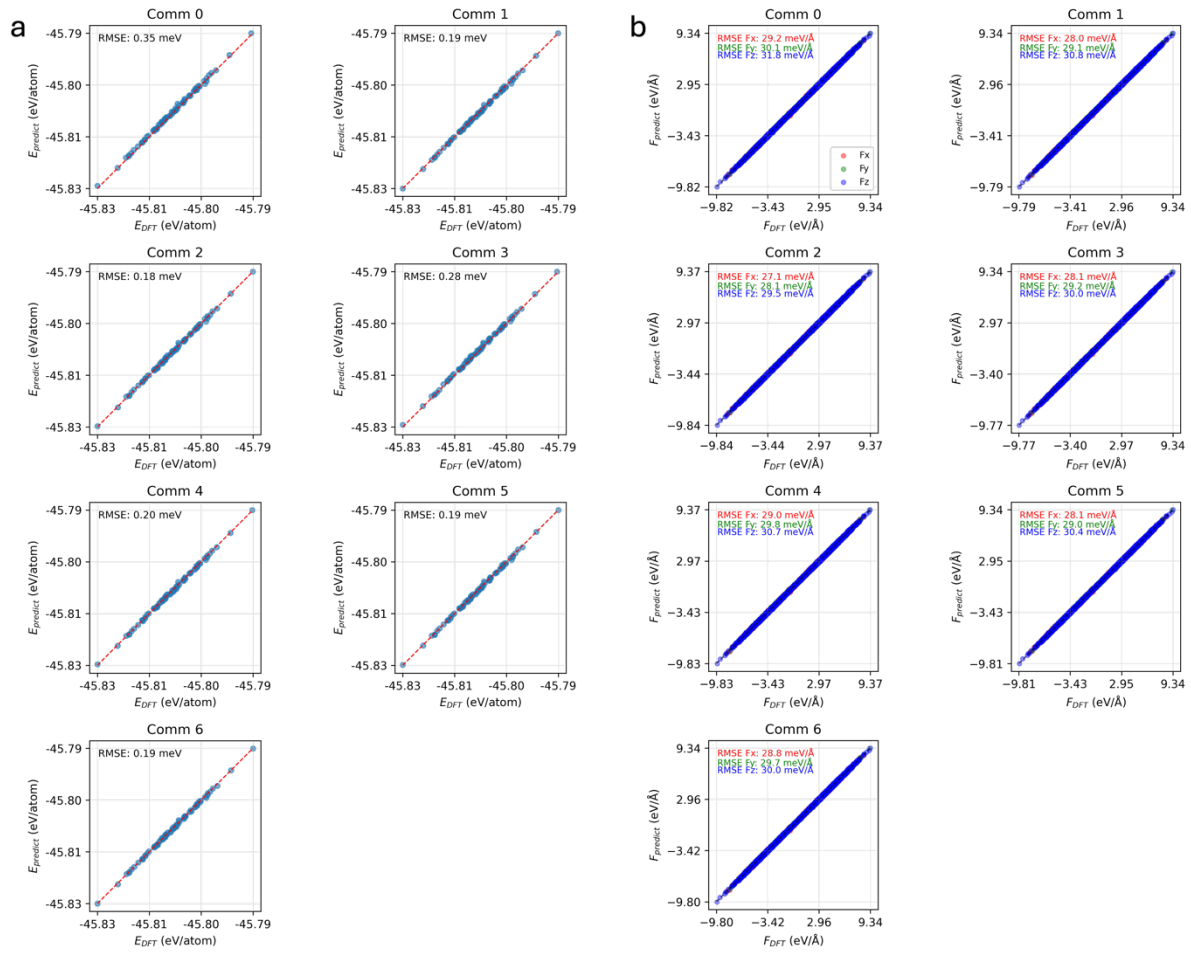


Figure S3. Test results for 7 committee ML models after training with the dataset applying the potentiostat. The training set included bulk and confined electrolyte structures, both with and without applying the potentiostat. The $\Delta\Phi$ in the training data were 0.5 V and 1.0 V. The test dataset consisted of confined electrolyte structures sampled from QM/ML simulations having $\Delta\Phi$ as 0.5 V and 1.0 V. (**a**) The energy prediction and (**b**) the force prediction is compared against the reference DFT calculations, which were performed on the trajectories sampled from the QM/ML simulations.

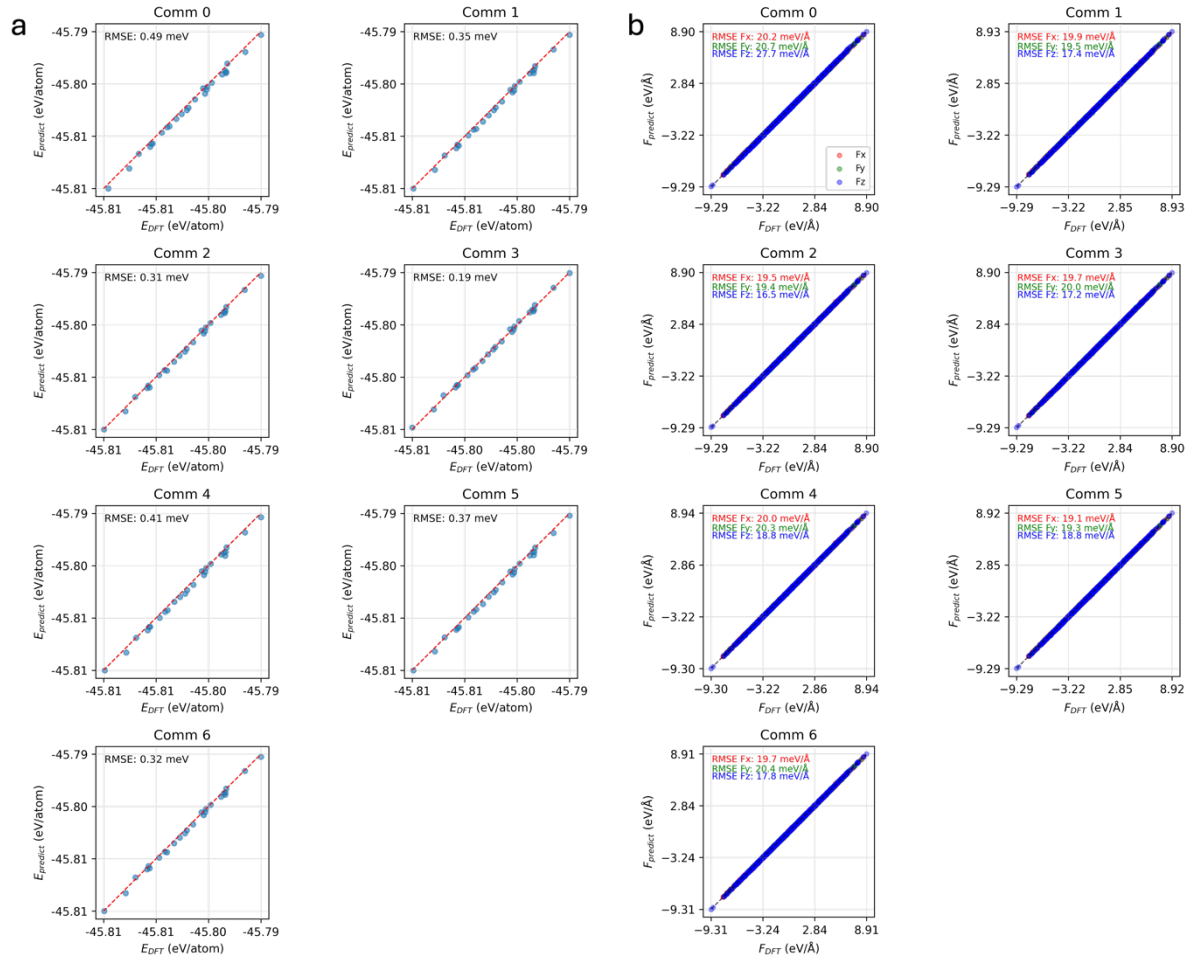


Figure S4. Interpolation performance of the 7-committee ML models. The training was conducted with the dataset with bulk electrolyte, confined electrolyte, with/without using a potentiostat with ΔΦ as 0.5 V and 1.0 V. The test dataset is comprised of confined electrolyte using a ΔΦ as 0.25 V. (**a**) The energy prediction and (**b**) the force prediction is compared against the reference DFT calculations, which were performed on the trajectories sampled from the QM/ML simulations.

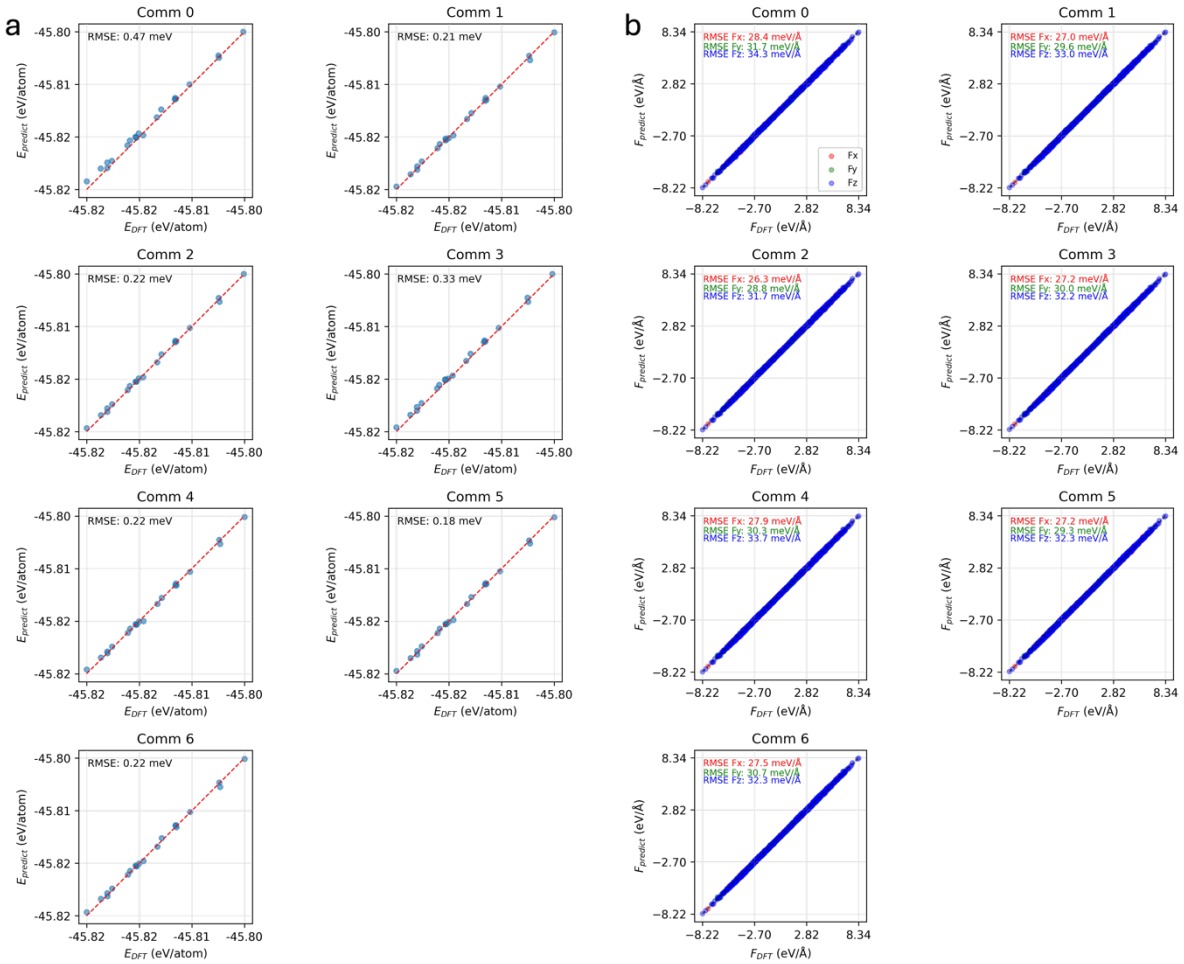


Figure S5. Interpolation performance of the 7-committee ML models. The training was conducted with the dataset with bulk electrolyte, confined electrolyte, with/without using a potentiostat with ΔΦ as 0.5 V and 1.0 V. The test dataset is comprised of confined electrolyte using a ΔΦ as 0.75 V. (**a**) The energy prediction and (**b**) the force prediction is compared against the reference DFT calculations, which were performed on the trajectories sampled from the QM/ML simulations.

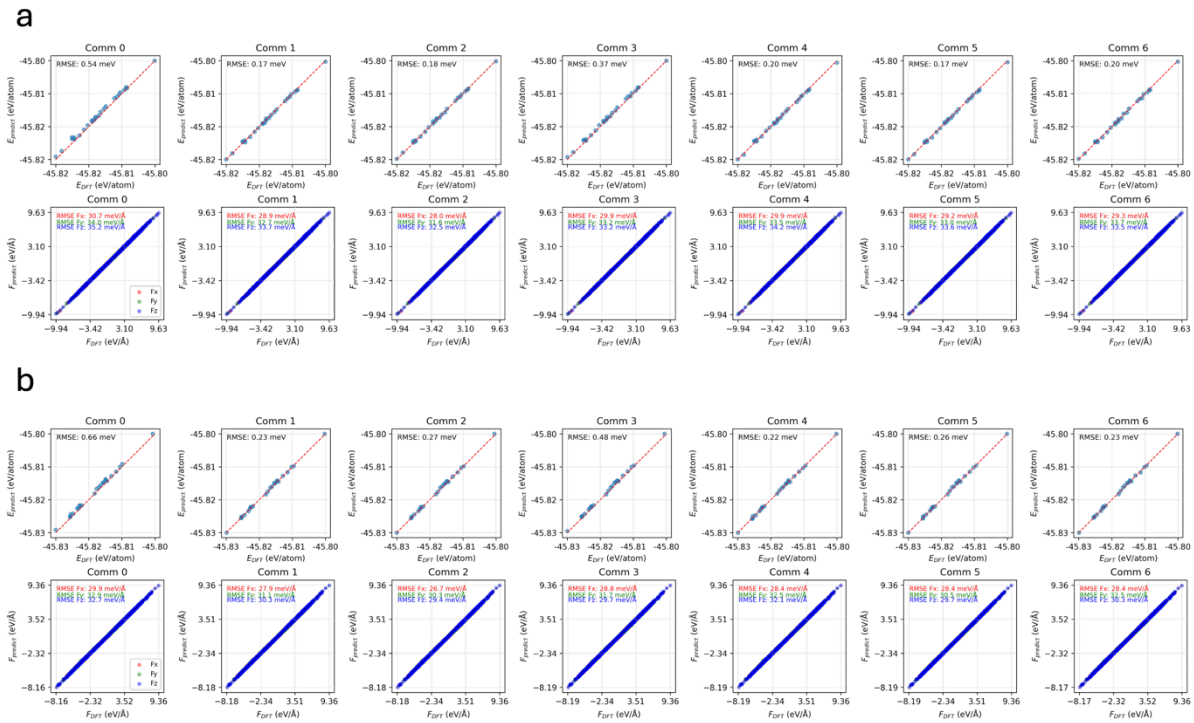


Figure S6. Extrapolation performance of the 7-committee ML models. The training was conducted with the dataset with bulk electrolyte, confined electrolyte, with/without using a potentiostat with ΔΦ as 0.5 V and 1.0 V. The test dataset is comprised of confined electrolyte using a ΔΦ as (**a**) 1.25 V and (**b**) 1.5 V. The energy prediction and the force prediction is compared against the reference DFT calculations, which were performed on the trajectories sampled from the QM/ML simulations.

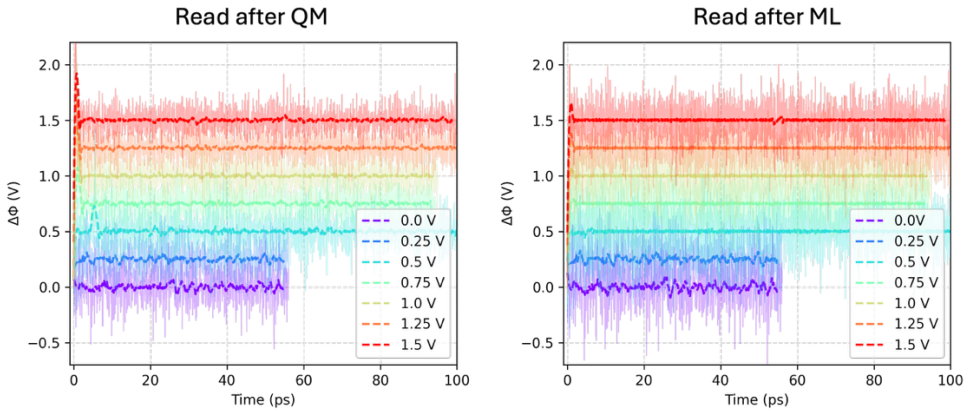


Figure S7. ΔΦ as a function of time. It is plotted for various target potential difference. Data corresponding to 0.25 V, 0.75 V, 1.25 V, and 1.5 V were excluded during the MLFF training. Averaged lines are displayed as a thicker curve after applying a 2.5 ps moving average. The data shown in Figure 3 corresponds to the potential obtained after the QM calculation, whereas the values measured after the ML step are also shown here to demonstrate consistency.

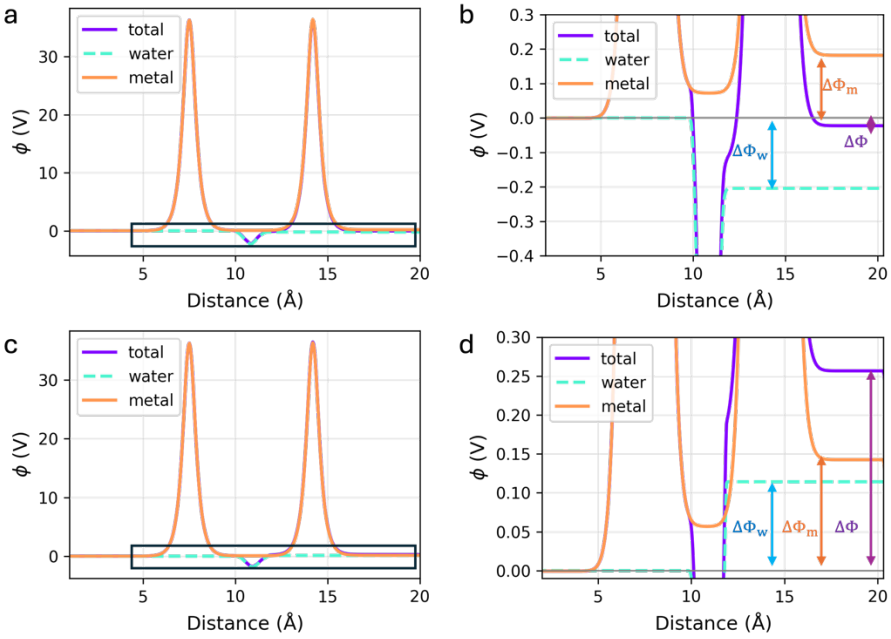


Figure S8. Electrostatic potential, ϕ, profile. (**a**) Electrostatic potential profile at ΔΦ = 0 V and its component analysis with a metal electrode and water electrolyte. (**b**) A zoomed-in view of the interfacial region from (**a**). (**c**) Potential profiles obtained at ΔΦ = 0.25 V, and its (**d**) zoomed-in image.

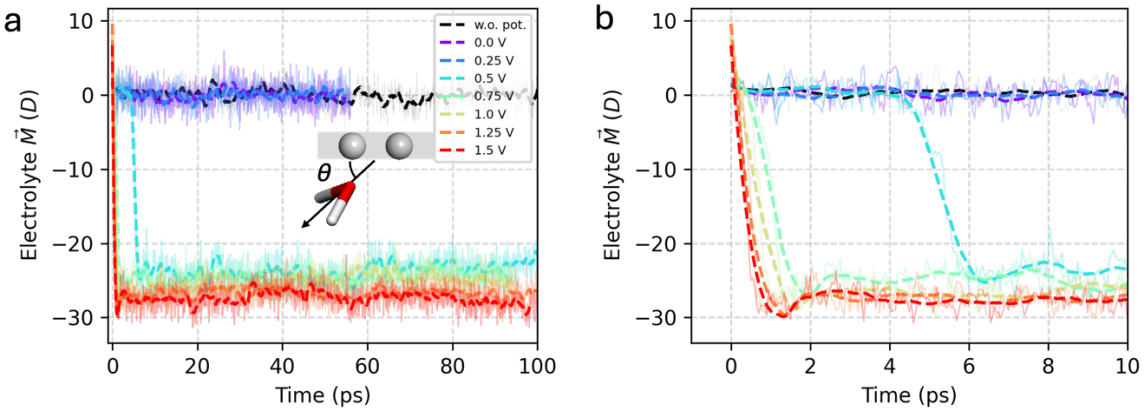


Figure S9. Analysis of the electrolyte dipole moment during the QM/ML simulations. **(a)** It is shown as a function of time, performed at various $\Delta\Phi$. The angle is defined as that between the vector connecting the oxygen atom to the midpoint of the hydrogen atoms and the electrode surface. **(b)** Zoomed-in plot focusing on the transient response.

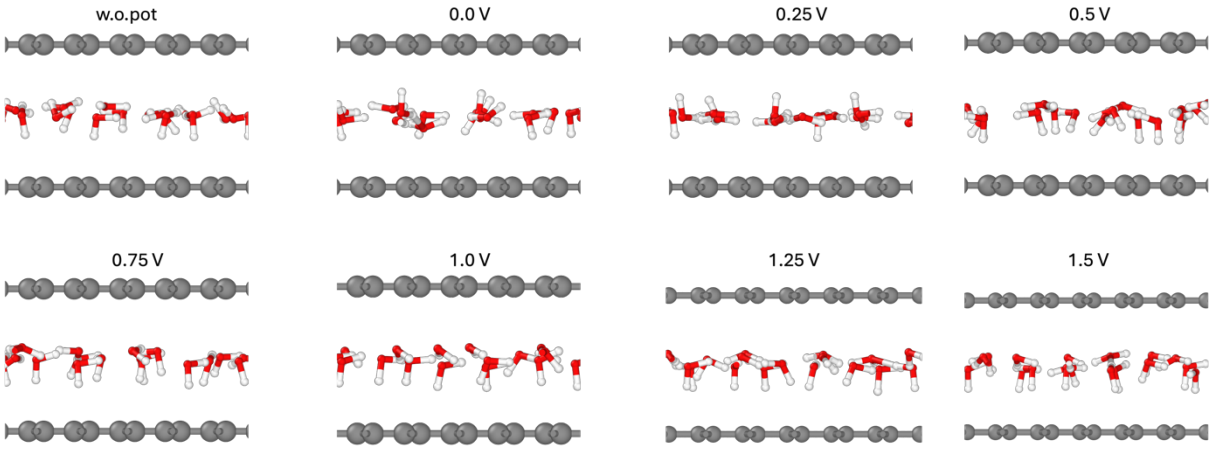


Figure S10. Snapshots of QM/ML simulations at different ΔΦ. The C, H, and O atoms are colored grey, white, and red, respectively.

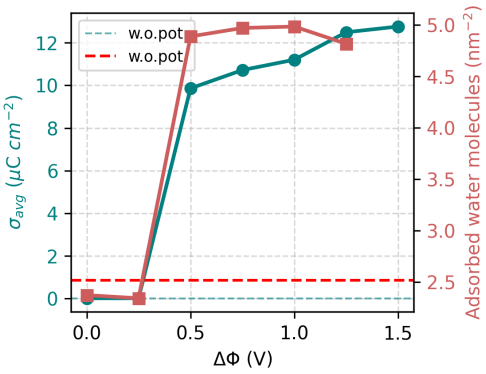


Figure S11. Average surface charge density (σ_{avg}) and the number of adsorbed water molecules as a function of $\Delta\Phi$. Values obtained without using a potentiostat (w.o. pot.) are indicated by dashed lines. A distance criterion is established to count the number of adsorbed water molecules, defined as those whose oxygen atom is within 2.9 Å of the electrode surface.

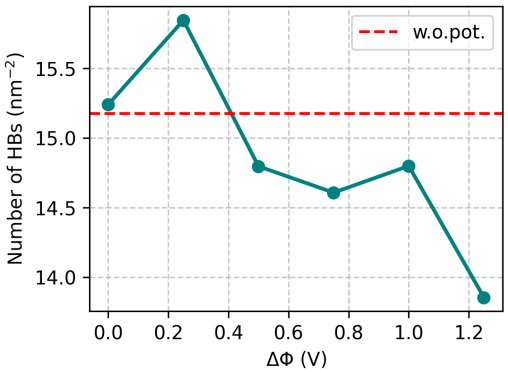


Figure S12. Average number of hydrogen bonds as a function of ΔΦ. Dashed lines indicate values obtained without a potentiostat (w.o. pot.) for comparison. The criteria used for hydrogen bond network analysis are a distance of 2.5 Å between the H atom and acceptor O atom, and an angle of 120° formed by the donor O, H, and acceptor O atoms.