Supporting Information: Impact of surface vacancies on the dynamics of metal atoms on reducible oxides: An enhanced sampling study with machine-learned potentials for Pt_1/TiO_2

Usama Saleem[†] and Shaama Mallikarjun Sharada[*,†,‡]

†*Mork Family Department of Chemical Engineering and Material Science, University of Southern California, Los Angeles, California 90089, USA*

‡*Department of Chemistry, University of Southern California, Los Angeles, California 90089, USA*

E-mail: ssharada@usc.edu

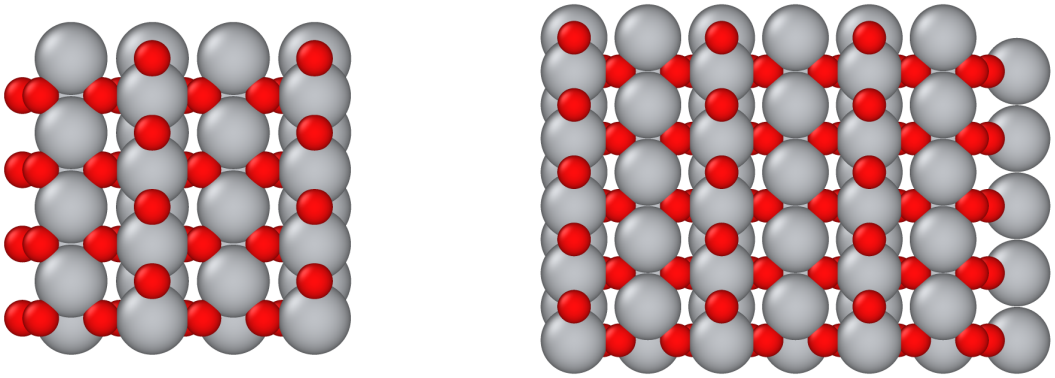
S1. Machine Learning Interatomic Potential (MLIP)

S1.1 Slab sizes

The differing slab sizes for TiO_2 rutile (110) are illustrated in Figure S1. For the first round of training, 2×4 slabs are used for all structures. All the production MD runs are performed with the 3×5 slabs.

S1.2 Validation

The parity plots for the atomic forces and atomic energies for each the two MLIP are illustrated in Figures S2 (a), (b), and S3 (a), (b). The distribution of these errors are illustrated



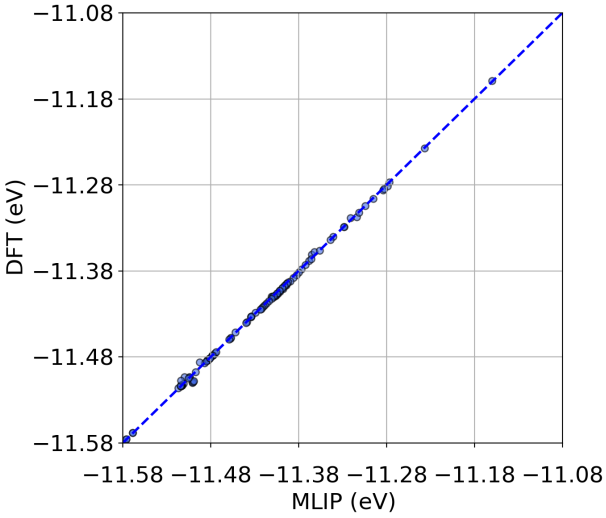
(a) Initial DFT 2×4 training slab (b) Production 3×5 simulation slab

Figure S1: TiO_2 rutile (110) slabs employed in this study: (a) The 2×4 slab is used to generate DFT data using VASP for the first iteration of training. (b) The 3×5 slab used for production molecular dynamics and metadynamics.

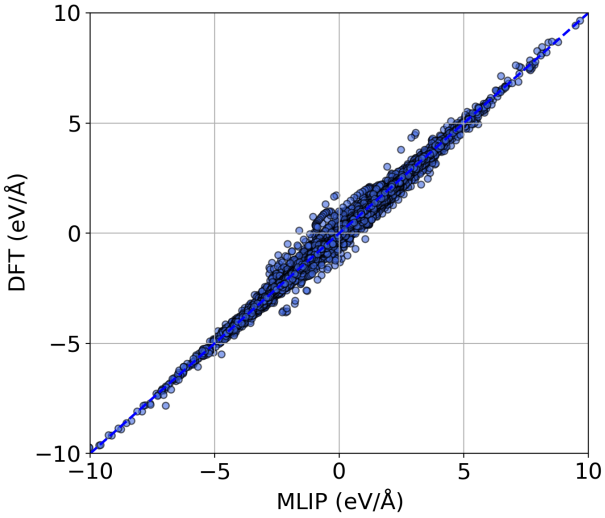
in S2 (c), (d), and S3 (c), (d). The structures are sampled from the third (final) and final iteration of the multi-generational training protocol. There are 110 structures validated for $\text{MLIP} - \text{O}_\text{v}$ and 34 structures validated for $\text{MLIP} - \text{Ti}_\text{v}$.

S1.3 Performance in presence of two surface O vacancies

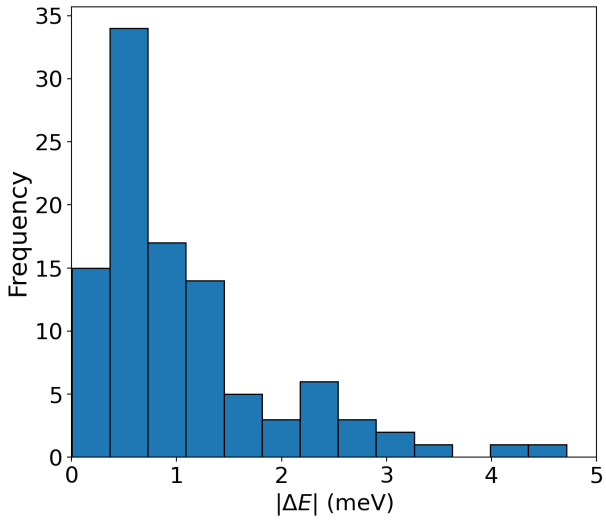
Structures are encountered for MD simulations and WT-MetaD for Pt at oxygen vacancies that contain more than one surface O vacancy on the metal-support interface. We sample 18 configurations from the final production simulations to assess the accuracy of MLIP. The computed MAE is 55.1 meV/Å for atomic forces and 8.8 meV/atom for potential energies. These metrics show that our $\text{MLIP} - \text{O}_\text{v}$ is reliable for simulations where additional oxygen vacancies are created.



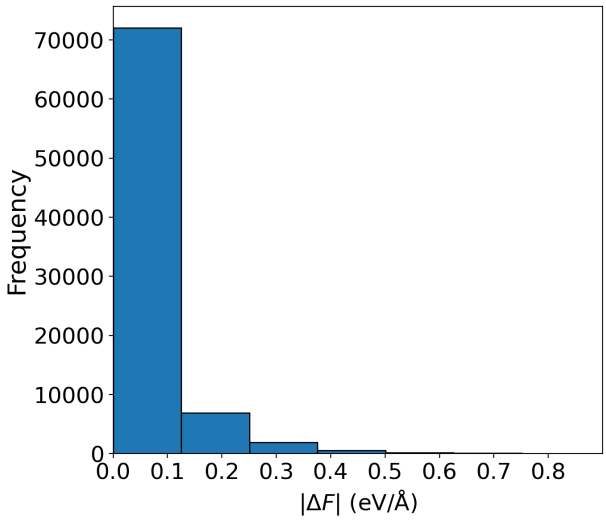
(a) Parity plot for atomic energies



(b) Parity plot for atomic forces

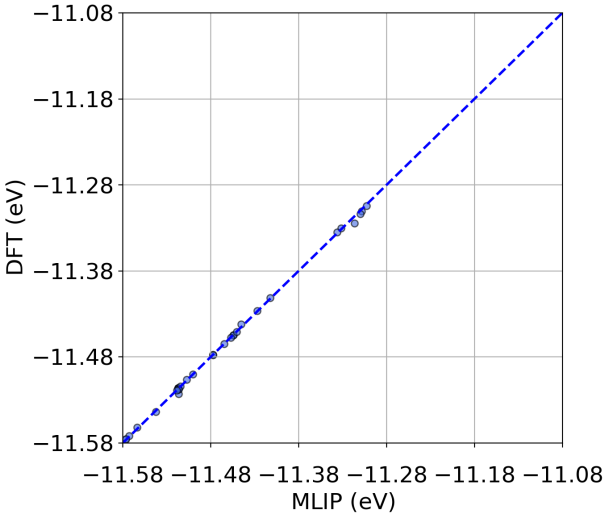


(c) Error distribution for atomic energies

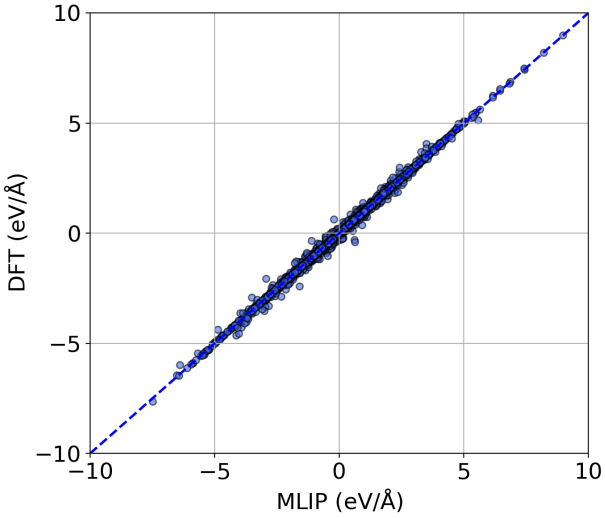


(d) Error distribution for atomic forces

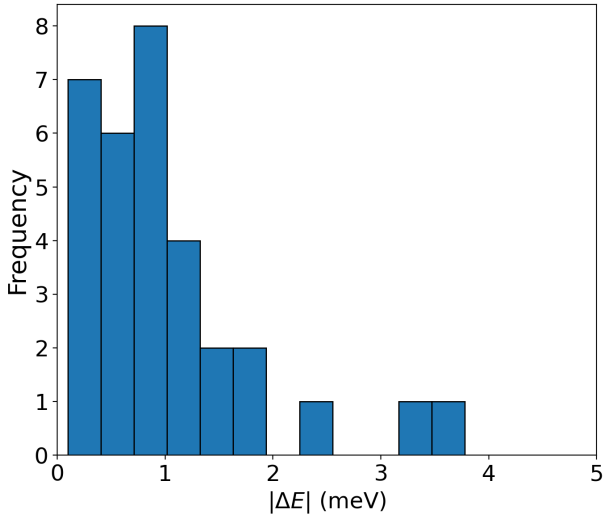
Figure S2: Performance metrics for MLIP − O_v. The parity plot for atomic energies is shown in (a) with MAE of **1.58 meV/atom**. The parity plot for atomic forces in (b) corresponds to a MAE of **51.1 meV/Å**. The histograms in (c) and (d) map the distribution of these absolute errors.



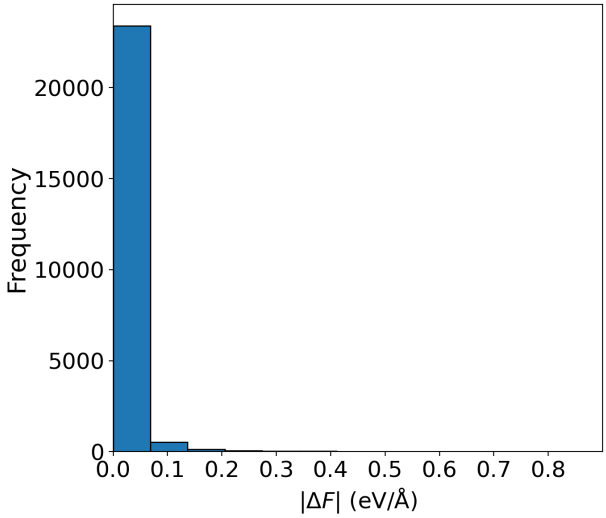
(a) Parity plot for atomic energies



(b) Parity plot for atomic forces

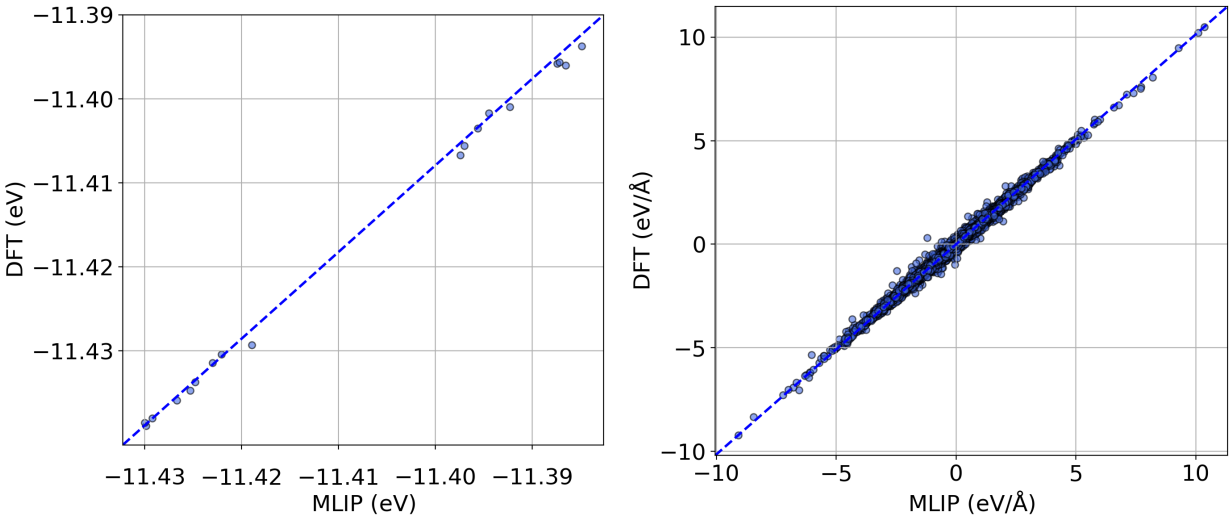


(c) Error distribution for atomic energies



(d) Error distribution for atomic forces

Figure S3: Performance metrics for MLIP − Ti_v. The parity plot for atomic energies is shown in (a) with MAE of **1.33 meV/atom**. The parity plot for atomic forces in (b) corresponds to a MAE of **18.2 meV/Å**. The histograms in (c) and (d) map the distribution of these absolute errors.



(a) Parity plot for atomic energies

(b) Parity plot for atomic forces

Figure S4: Performance of MLIP − O_v on structures with additional O vacancies.

S1.4 Performance on key structures encountered in production simulations

To verify the accuracy of important/rare motifs observed in our simulations, we sample structures from the production simulations that correspond to the following:

- Migration of oxygen vacancies.
- The rupturing of O-Pt-O motif in presence of oxygen vacancy.
- Formation of PtO_4 motif in presence of Ti vacancy.
- Pt atom burying itself beneath the first layer of TiO_2 rutile (110).

The resulting parity plots are shown in Figure S5. A total of 25 oxygen deficient structures and 26 titanium deficient structures are labeled with DFT. The MAE for atomic forces is 59.6 meV/Å for MLIP − O_v and 93.8 meV/Å for MLIP − Ti_v. The MAE for relative atomic energies is 1.6 meV/atom for MLIP − O_v and 3.0 meV/atom for MLIP − Ti_v.

We point out that the atomic energies in Figures S5 (a) and (c) are relative to the first

structure in each list, for DFT and MLIP each based on Equation 1:

$$E_{relative,i} = E_i - E_0 \tag{1}$$

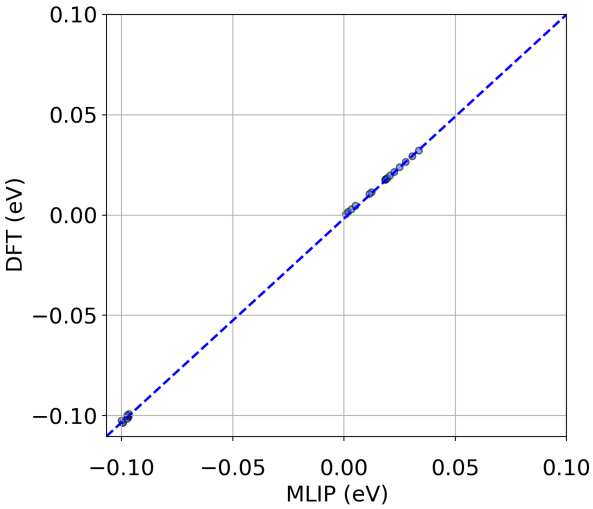
E_0 refers to the atomic energy for the first structure. For computing the MAE for atomic energies, these reference structures are excluded from the list. The reason for adopting this approach is the systematic shift encountered in potential energy predictions for production slabs (3×5) compared to training slabs (2×4). MACE decomposes potential energy of the structure as a sum of contributions of atomic energies.[1,2] In the process, it estimates the energy of isolated atoms by averaging over the structures in training dataset. The resulting systemic deviations in isolated atom energies cause a systematic shift in potential energy predictions for structures corresponding to sizes differing from training dataset when compared to DFT, since DFT does not compute potential energy as a sum of atomic energy contributions. Since the relative energies are the meaningful quantities in atomistic simulations, the MLIPs developed here are capable of accurately describing atomic interactions for the relevant structures.

S2. Nanosecond trajectories

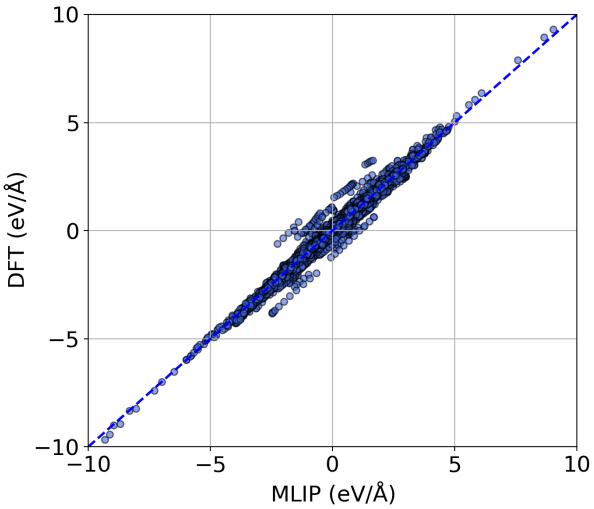
For Pt bound to the three vacancies, we extend the 250 ps simulation to 1 ns. The mean-squared displacement (MSD) versus time is depicted in Figure S6.

S3. Basal oxygen vacancy migration

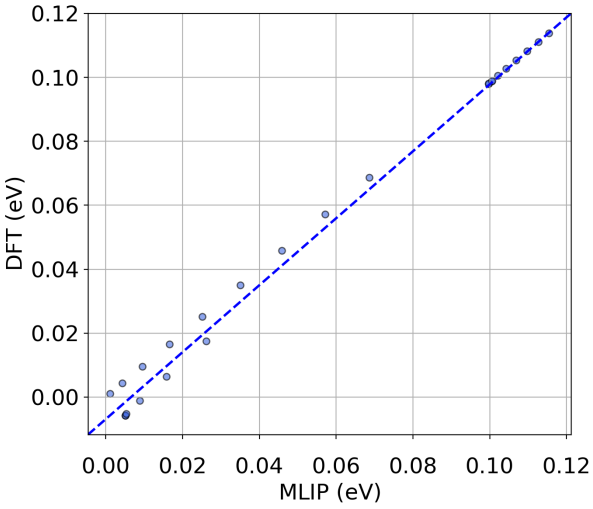
We find that the basal oxygen vacancy created on the bare, metal-free surface migrates to a bridge vacancy at 0 K and 1000 K, confirmed by DFT optimization and AIMD using the r^{2}SCAN functional. This observation is in agreement with the work presented by Humphrey and co-workers, in which (SCAN-based) DFT optimization showed that creating an isolated



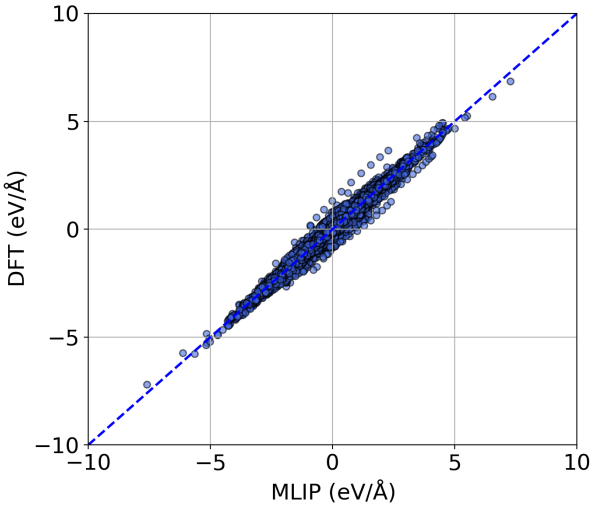
(a) Parity plot for relative atomic energies on Ti-deficient structures



(b) Parity plot for atomic forces on Ti-deficient structures



(c) Parity plot for relative atomic energies on O-deficient structures



(d) Parity plot for atomic forces on O-deficient structures

Figure S5: Performance of MLIP-O_v and MLIP-Ti_v on key structures sampled from production simulations

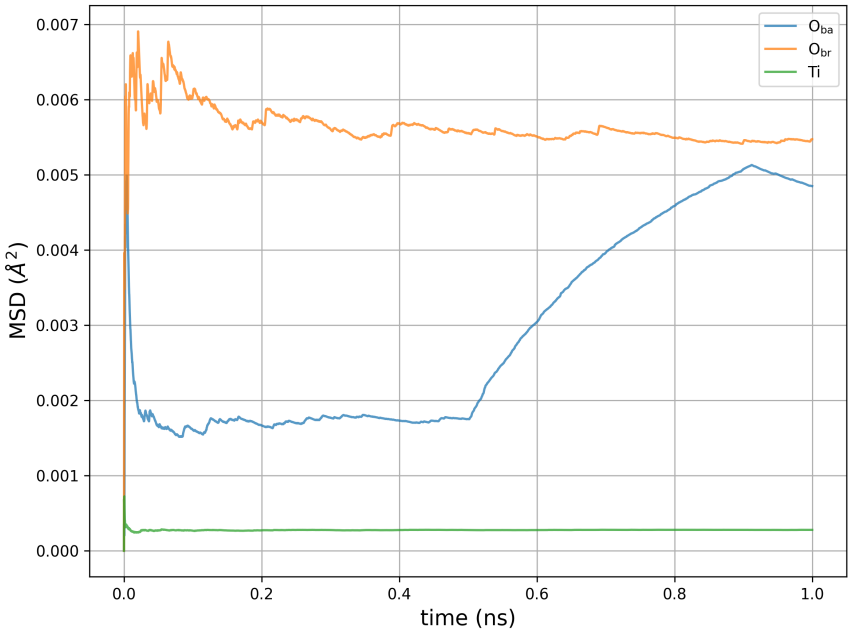
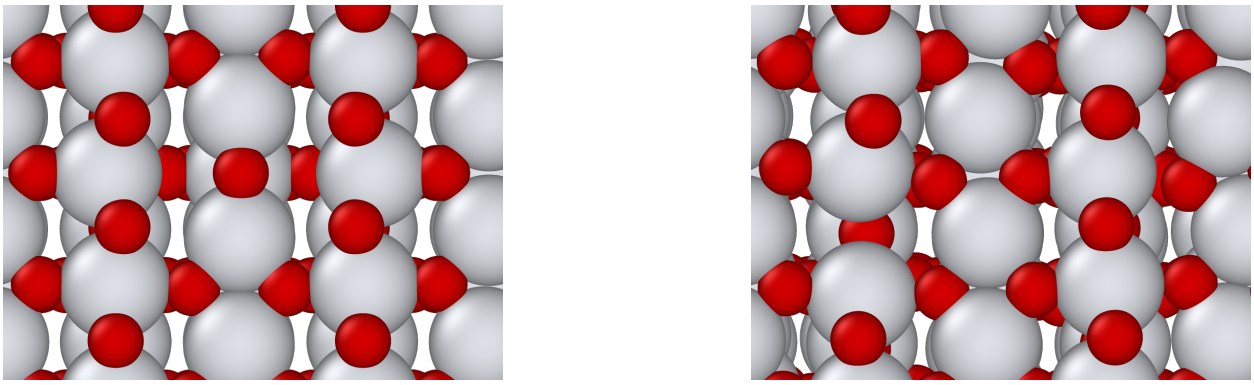


Figure S6: Mean-squared displacement (MSD) for Pt atom bound at the three vacancy sites, O_{br}, O_{ba}, and Ti5c, at 1000 K for 1 ns MD post thermal equilibration.

basal oxygen vacancy induces migration of the neighboring basal oxygen,[3] resulting in a structure identical to that shown in Figure S7 (a). At finite temperature, the migrated basal oxygen atom heals the basal oxygen vacancy at temperatures as low as 50 K, creating a bridge oxygen vacancy. The lack of stability of a basal oxygen vacancy can be directly linked to the coordination environment of the basal oxygen atoms, each oxygen atom coordinated to three Ti atoms, making it less reactive compared to the bridge oxygen atoms, where each bridge oxygen atom is coordinated to two Ti atoms. Although we show that Pt at the basal oxygen vacancy site is highly stable, a basal oxygen vacancy transforms to a bridge oxygen vacancy during molecular dynamics. This makes it impossible to model impact of O_{ba} vacancy on neighboring Pt site. It is important to note that the vacancy migration does not occur if the O_{ba} vacancy is occupied by a Pt atom. In addition, the migration of basal oxygen is also observed upon geometry optimization with the MLIP $- O_v$. This is followed by basal vacancy healing at finite temperature simulation; serving as an additional validation for the MLIP.



(a) Geometry-optimized (b) Finite temperature

Figure S7: Basal O vacancy induced restructuring of TiO_2

S4. O-Pt-O near Ti5c vacancy

In the presence of a Ti vacancy, the O-Pt-O motif undergoes reconstruction to form a square planar PtO_4 motif. For O-Pt-O near O_{br} vacancy, we noted a rupture of the motif upon performing molecular dynamics at 1000 K. However, in the presence of Ti vacancy, the reconstruction takes place during structure optimization using the BFGS algorithm with the MLIP, shown in Figure S8. This indicates that O-Pt-O in presence of Ti5c vacancy is not stable at 0 K.

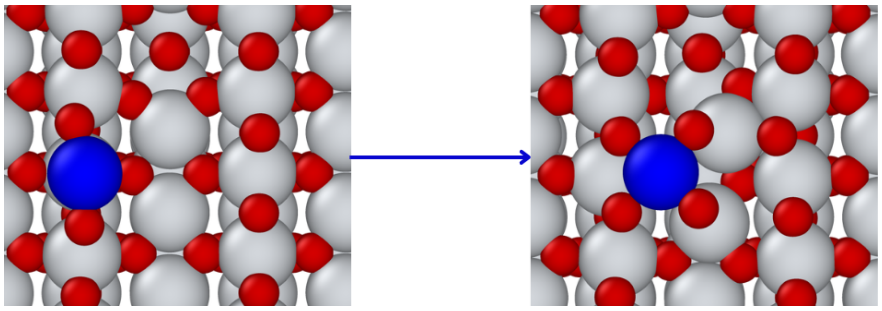


Figure S8: Formation of PtO_4 upon BFGS geometry optimization of O-Pt-O in vicinity of Ti vacancy using MLIP.

S5. Motion of Pt@Ti5c vacancy

We note that the motion of Pt atoms doped at the Ti5c site is similar to the motion of neighboring Ti5c atoms. We illustrate this by plotting the time evolution of the z coordinate of the Pt atom and the neighboring Ti5c atom in Figure S9.

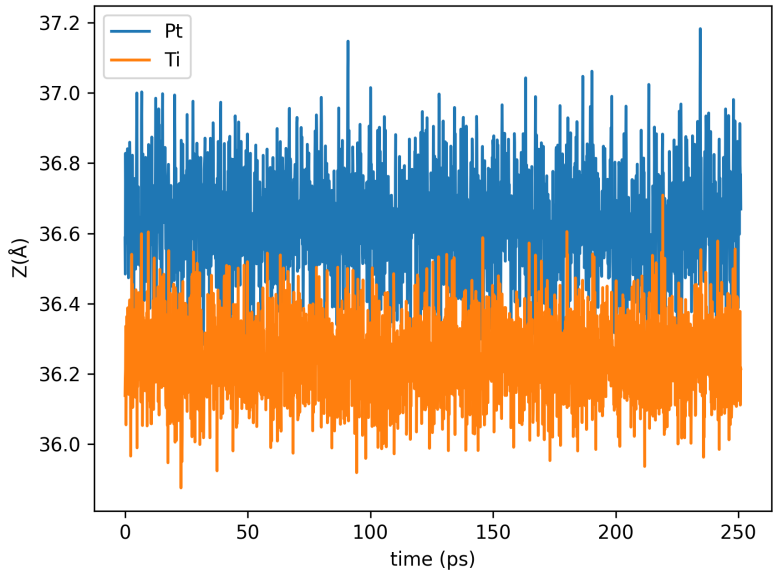


Figure S9: Motion of the doped Pt and adjacent Ti5c atom perpendicular to the TiO_2 surface at 1000 K for one of the 20 replica MD trajectories. The standard deviation in the z-coordinate of Pt is 0.16 Å and 0.11 Å for the z-coodinate of Ti5c atom. The mean z-coordinate of Pt atom is 0.43 Å higher that of Ti5c.

S6. Pt@Ti5c vacancy in the presence of displaced oxygen

For one of the 10 walkers used for Pt@Ti5c vacancy WT-MetaD, a bridge oxygen migrates to the top of the nearest Ti5c atom. This configuration is shown in Figure S10. Upon geometry optimization, Pt atom in this configuration acquires a z-coodinate 0.09 Å higher than the unperturbed Pt@Ti5c vacancy. This difference in z-coordinate is smaller than variation we

see in Figure S9. As a result, we assume that this displaced O_{br} has negligible impact on the collective variable and results of our metadynamics simulations for Pt@Ti5c vacancy.

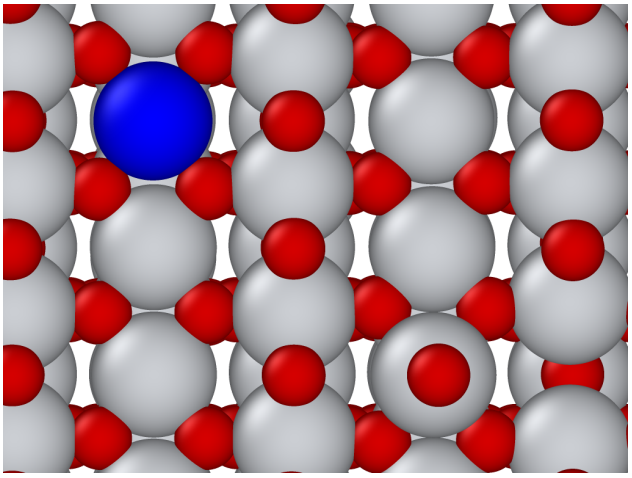
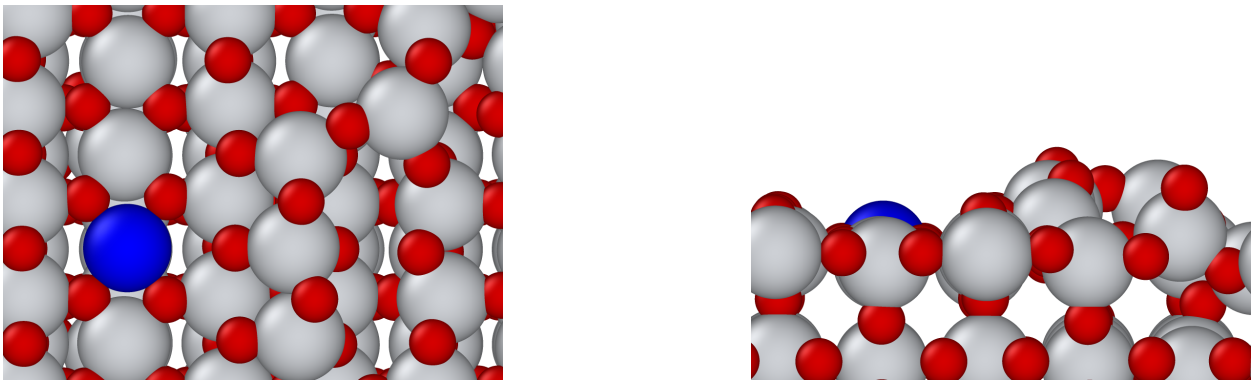


Figure S10: Pt@Ti5c vacancy with a displaced O_{br}.

S7. Formation of TiO_x overlayers

For 1 of 20 replica MD simulations for Pt@Ti5c vacancy, we encounter formations of TiO_x layers as the surface atoms are displaced. Upon DFT optimization, we find that these restructured layers are stable as shown in Figure S11. Formation of these TiO_x motifs under annealing conditions is a well documented phenomenon for TiO_2 rutile (110).[4–6] We note that the reconstruction in Figure S11 comes from an MD simulation that has been executed for only 250 ps. As a result, the atoms beneath the first layer do not have enough time to diffuse to the top layers. We also take into account the fact that MLIP $-$ Ti_v is not explicitly trained on structures containing TiO_x reconstructions. Upon validating on 50 structures containing the reconstruction in Figure S11, we find



(a) Top view (b) Side view

Figure S11: DFT optimized TiO_x formed at 1000 K for Pt@Ti5c vacancy.

S8. Uncertainty in Free Energy

We estimate the uncertainty in metadynamics computed free energies based on the work by Laio and coworkers.[7] The method expresses uncertainty based on the parameters used to perform metadynamics by arriving at the following relationship:

$$\bar{\epsilon} = C(d)\sqrt{\frac{S\delta_s}{D\tau_G}\frac{w}{\beta}} \tag{2}$$

Where $\bar{\epsilon}$ is the uncertainty in free energy, $C(d)$ is coefficient dependent on the dimensionality of the FES, S is the width of the metadynamics collective variable grid, δ_s is the variation in the width of applied Gaussian hills, w is the height of the Gaussian hills, D is the diffusivity of the collective variables computed using autocorrelation functions, τ_G is the time-interval between successive hill-depositions, and β is $1/k_BT$.

The coefficient $C(d)$ is set equal to 0.3 if 2 CVs are used and 0.5 if 1 CV is used. For the diffusivity of CVs, D, the smallest value, rounded down to the order of magnitude, across all replica trajectories is employed. This ensures we obtain the largest estimate for $\bar{\epsilon}$. Additionally, it is also worth noting that the uncertainty $\bar{\epsilon}$ does not depend on the number of walkers in multi-walker metadynamics.[8] Upon obtaining an estimate of the uncertainty in $\Delta G^{\ddagger}$, the uncertainty in the intrinsic rate coefficient k is estimated using the following equations:

Table S1: Detailed uncertainty estimation for free energy.

System	$C(d)$	S (Å)	δs (Å)	w (eV)	D ($Å^2 fs^{-1}$)	τ_G (fs)	$\bar{\epsilon}$ (eV)
Pt@O_{br}	0.3	40	0.10	0.01	0.0001	1000	0.06
Pt@O_ba	0.3	40	0.10	0.01	0.0001	1000	0.06
Pt@Ti5c	0.5	20	0.10	0.01	0.0001	1000	0.07

$$k(UL) = \frac{k_B T}{h} e^{\frac{\Delta G^{\ddagger}(UL)}{k_B T}} \tag{3}$$

$$k(LL) = \frac{k_B T}{h} e^{\frac{\Delta G^{\ddagger}(LL)}{k_B T}} \tag{4}$$

$$\Delta k = \frac{k(UL) - k(LL)}{2} \tag{5}$$

Where $k(UL)$ is the upper limit on the rate coefficient, computed using the upper limit of free energy barrier $\Delta G^{\ddagger}(UL)$, $k(LL)$ is the lower limit on the rate constant, computed using the lower limit of free energy barrier $\Delta G^{\ddagger}(LL)$, and Δk is the estimated uncertainty in the rate coefficient. The results are shown in Table S1.

S9. Uncertainties in Diffusion Parameters

We recall the equation used for modeling anomalous diffusion in this study:[9]

$$\langle \Delta r(t)^2 \rangle = \frac{4 K_\alpha t^\alpha}{\Gamma(1+\alpha)} \tag{6}$$

For the linearized model, the variance in slope, m, is computed as:

$$\mathrm{Var}(m) = \frac{\mathrm{Var}(y_{true} - y_{model})}{\sum(x_i - \bar{x})^2} \tag{7}$$

Similarly the variance in the intercept, b, is calculated using:

$$\mathrm{Var}(b) = \frac{\mathrm{Var}(y_{true} - y_{model})}{n} + \frac{\mathrm{Var}(y_{true} - y_{model})\bar{x}^2}{\sum(x_i - \bar{x})^2} \tag{8}$$

where n is the number of data points fitted using ordinary least squares. The errors reported for α in Table 4 are the square roots of the variances in the reported slopes. To compute error in K_α, the standard deviation in intercept, Δb is used to obtain an upper and lower limit for the values of K_α as following:

$$K_\alpha(UL) = \frac{\Gamma(1+\alpha)e^{b+\Delta b}}{4} \tag{9}$$

$$K_\alpha(LL) = \frac{\Gamma(1+\alpha)e^{b-\Delta b}}{4} \tag{10}$$

The error in K_α is then computed as:

$$\Delta K_\alpha = \frac{K_\alpha(UL) - K_\alpha(LL)}{2} \tag{11}$$

References

(1) Batatia, I.; Kovacs, D. P.; Simm, G.; Ortner, C.; Csanyi, G. MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields. Advances in Neural Information Processing Systems. 2022; pp 11423–11436.

(2) Batatia, I.; Batzner, S.; Kovács, D. P.; Musaelian, A.; Simm, G. N. C.; Drautz, R.; Ortner, C.; Kozinsky, B.; Csányi, G. The design space of E(3)-equivariant atom-centred interatomic potentials. *Nature Machine Intelligence* **2025**, *7*, 56–67.

(3) Humphrey, N.; Bac, S.; Sharada, S. M. Ab Initio Molecular Dynamics Reveals New Metal-Binding Sites in Atomically Dispersed Pt/TiO_2 Catalysts. *The Journal of Physical Chemistry C* **2020**, *124*, 24033–24542.

(4) Diebold, U. The surface science of titanium dioxide. *Surface Science Reports* **2003**, *48*, 53–229.

(5) Onishi, H.; Iwasawa, Y. Reconstruction of TiO_2(110) surface: STM study with atomic-scale resolution. *Surface Science* **1994**, *313*, L783–L789.

(6) Yuan, W.; Chen, B.; Han, Z.-K.; You, R.; Jiang, Y.; Qi, R.; Li, G.; Wu, H.; Ganduglia-Pirovano, M. V.; Wang, Y. Direct in-situ insights into the asymmetric surface reconstruction of rutile TiO_2 (110). *Nature Communications* **2024**, *15*, 1616.

(7) Laio, A.; Rodríguez-Fortea, A.; Gervasio, F. L.; Ceccarelli, M.; Parrinello, M. Assessing the Accuracy of Metadynamics. *The Journal of Physical Chemistry B* **2005**, *109*, 6714–6721.

(8) Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics. *The Journal of Physical Chemistry B* **2006**, *110*, 3533–3539, PMID: 16494409.

(9) Oliveira, F. A.; Ferreira, R. M. S.; Lapas, L. C.; Vainstein, M. H. Anomalous Diffusion: A Basic Mechanism for the Evolution of Inhomogeneous Systems. *Frontiers in Physics* **2019**, *7*, 18.