Supplemental Information

Grotthuss-type Proton Transport Governed by Oxygen Coordination Environment in VO_2 Polymorphs

Sunghyun Park[1], Shin-ichi Nishimura[1], Jinshi Li[1], Muyuan Li[1], and Atsuo Yamada[1,]*

[1]Department of Chemical System Engineering, School of Engineering, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8656, Japan.

*yamada@chemsys.t.u-tokyo.ac.jp

Methods

MD simulations were performed using an interatomic potential based on the Universal Model Architecture (UMA) of FAIRChem version 2, a graph neural network trained on the OMat24 dataset. The model preserves rotational and permutational symmetries and enables transferable simulations of proton dynamics in vanadium oxide frameworks.[1,2] The model adopts an equivariant graph neural network architecture pretrained on the OMat24 DFT+*U* dataset and is further fine-tuned with reference calculations tailored to vanadium dioxides (VO_2(A), VO_2(B), and VO_2(R))). For each polymorph, we sampled 450 structural configurations across a range of proton concentrations (H/V = 0, 1/8, 1/2, 11/12, 1) and temperatures (300, 600, 900 K). Spin-polarized density functional theory (DFT) calculations were performed using the Vienna *Ab initio* Simulation Package (VASP, version 5.4.4).[3,4] Energies and forces were computed using the PBE + *U* functional (*U* = 3.25 eV on V 3*d*) with D3 van der Waals corrections, as follows: The total dataset (1350 structures) was split evenly into training and validation subsets, and fine-tuning was performed using the training pipeline.

MD simulations were conducted using the Atomic Simulation Environment (ASE) with a fine-tuned potential as a calculator.[5] Periodic supercells of each polymorph were constructed to capture the long-range proton motion and minimize the image effects. For VO_2(A), VO_2(B), and VO_2(R), supercells consisting of 192 V and 384 O atoms were generated, and 48 protons were inserted at low-coordination oxygen sites, yielding a composition of $H_1V_4O_8$ (n_H/n_V = 0.25). The proton configuration was initialized to avoid short O–H bonds (<0.9 Å) and minimize electrostatic repulsion.

The validation of the pretrained MLIP and fine-tuned MLIP against DFT reference calculations was performed by comparing the mean absolute errors (MAEs) of the energy and force. (**Table S1**) The fine-tuned models exhibited meaningful improvements across all the vanadium oxide systems examined. For energy prediction, the MAEs decreased from approximately 18–25 meV

atom^{-1} in the pretrained model to 4–7 meV atom^{-1} after the fine-tuning. Likewise, the force MAEs were reduced from 0.080–0.120 eV Å^{-1} to 0.020–0.035 eV Å^{-1}. These results indicate that system-specific fine-tuning effectively captures local structural and energetic variations, yielding more accurate and physically consistent potential energy surfaces. **Table S2** compares the lattice parameters and unit cell volumes of the four vanadium oxides obtained from the fine-tuned MLIP-based structural relaxation with the corresponding experimental data. The MLIP-relaxed structures reproduced the experimental lattice parameters within approximately 1% and cell volumes within 4%, confirming that the fine-tuned potential accurately captured the equilibrium geometries and lattice anisotropy across all phases. Overall, these results demonstrate that the fine-tuned MLIPs not only reproduce DFT-level energies and forces with high fidelity, but also maintain structural consistency with experimental observations in VO_2 polymorphs.

Table S1 Comparison of mean absolute errors (MAEs) in energy and force between pretrained and fine-tuned machine-learning interatomic potentials (MLIPs) validated against DFT reference calculations for four vanadium oxide systems.

Materials	MAEs in energy (meV)		MAEs in force (eV Å^{-1})	
	Pretrained	Fine-tuned	Pretrained	Fine-tuned
VO_2(A)	0.747	0.698	0.118	0.090
VO_2(B)	2.619	0.523	0.101	0.093
VO_2(R)	3.147	1.099	0.122	0.108

Table S2 Comparison of lattice parameters and unit cell volumes of fine-tuned machine-learning interatomic potentials (MLIPs) validated against experimental data for four vanadium oxide systems.

Materials	Ratio between the calculated and experimental data.			
	a (%)	b (%)	c (%)	V (%)
VO_2(A)	+1.00	+1.00	+1.00	+3.01
VO_2(B)	-0.03	+1.21	+1.21	+3.68
VO_2(R)	+0.05	+1.01	+1.10	+2.02

All simulations were performed in the NVT ensemble using a Nose–Hoover thermostat. The temperatures were set at 550, 630, 700, and 780 K. The time step was 0.5 fs, and the trajectories were integrated for 1 ns under each condition. Each system was equilibrated for 200 ps prior to the production runs for relaxation. Proton self diffusion coefficients and related transport properties were extracted from the trajectories using the KINISI analysis package.[6,7]

Proton transport was evaluated from the mean squared displacement (MSD) of hydrogen atoms, defined as:

$$\mathrm{MSD}(t) = \frac{1}{N}\sum_{i=1}^{N}\left[\left(x_i(t) - x_i(0)\right)^2 + \left(y_i(t) - y_i(0)\right)^2 + \left(z_i(t) - z_i(0)\right)^2\right]$$

where N is the number of protons and $x_i(t)$, $y_i(t)$, and $z_i(t)$ are the Cartesian coordinates of atom i at time t. The self-diffusion coefficient D was determined from the slope of the MSD curve:

$$D = \frac{1}{6}\frac{d(\mathrm{MSD})}{dt}$$

Activation energies (E_a) were derived from Arrhenius fitting of the temperature-dependent diffusion coefficients according to

$$D = D_0 \exp\left(-\frac{E_a}{k_B T}\right)$$

where D_0 is the pre-exponential factor, k_B is the Boltzmann constant, and T is temperature.

The radial distribution function (RDF), $g_{AB}(r)$, was used to characterize local atomic environments:

$$g_{AB}(r) = \frac{1}{4\pi r^2 \rho_B N_A} \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \delta(r - r_{ij})$$

where ρ_B is the number density of species *B*, N_A is the number of reference atoms *A*, and r_{ij} is the interatomic distance between atoms *i* and *j*. The RDFs of the H–O and O–O pairs were analyzed to probe the hydrogen-bonding and structural order.[8] The coordination number $N_{AB}(r_c)$ was obtained by integrating the corresponding radial distribution function $g_{AB}(r)$ up to a cutoff distance r_c, typically defined at the first minimum of $g_{AB}(r)$:

$$N_{AB}(r_c) = 4\pi\rho_B \int_0^{r_c} g_{AB}(r)\ r^2\, dr$$

where $\rho_B = N_B/\mathrm{V}$ denotes the number density of species *B*, and this integral represents the average number of *B* atoms surrounding each *A* atom within a sphere of radius r_c, thereby quantifying the local atomic coordination.[8] The cumulative form $N_{AB}(r_c)$ provides the running coordination number as a function of distance, which allows identification of distinct coordination shells in the atomic structure.

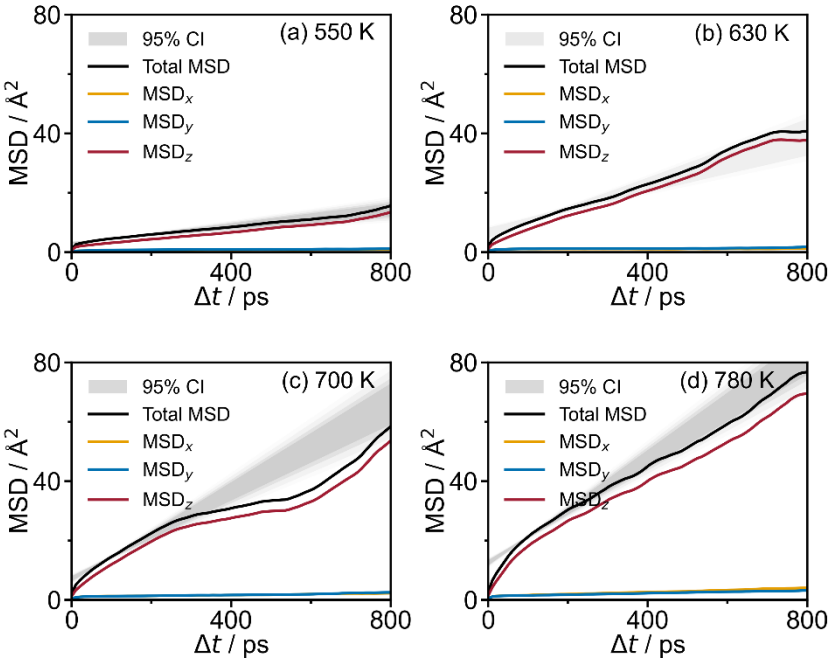

Figure S1. Mean-square displacement (MSD) profiles of VO_2(A) at (a–d) 550–780 K. The temperature dependence of the MSD slope reveals the acceleration of proton diffusion at high temperatures.

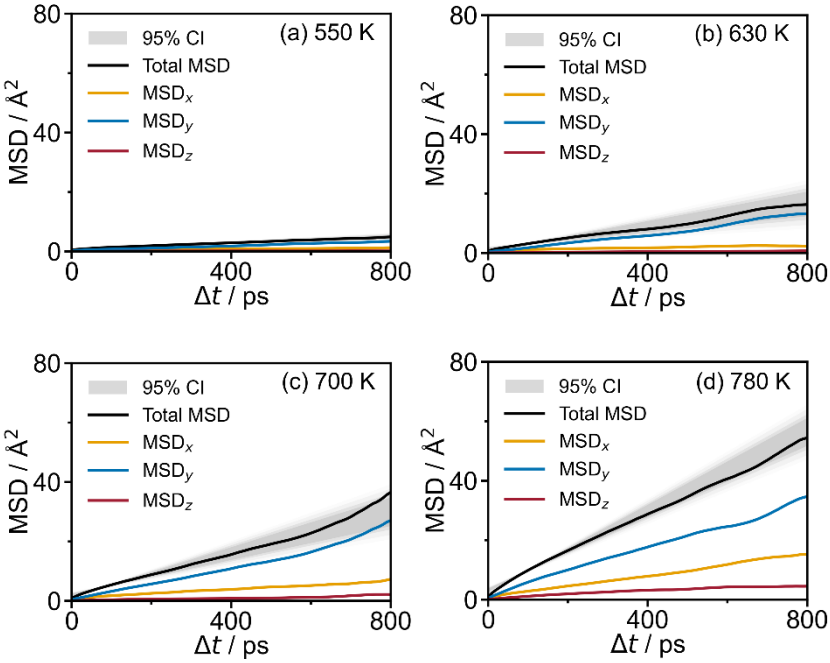

Figure S2. Mean-square displacement (MSD) profiles of VO_2(B) at (a–d) 550–780 K. The temperature dependence of the MSD slope reveals the acceleration of proton diffusion at elevated temperatures.

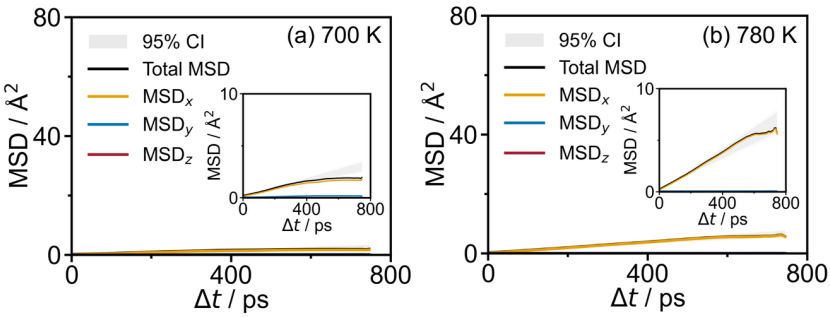

Figure S3. Mean-square displacement (MSD) profiles of VO_2(R) at (a) 700 and (b) 780 K. Data at 550 and 630 K are omitted because long-range diffusion was not observed within the simulation time window.

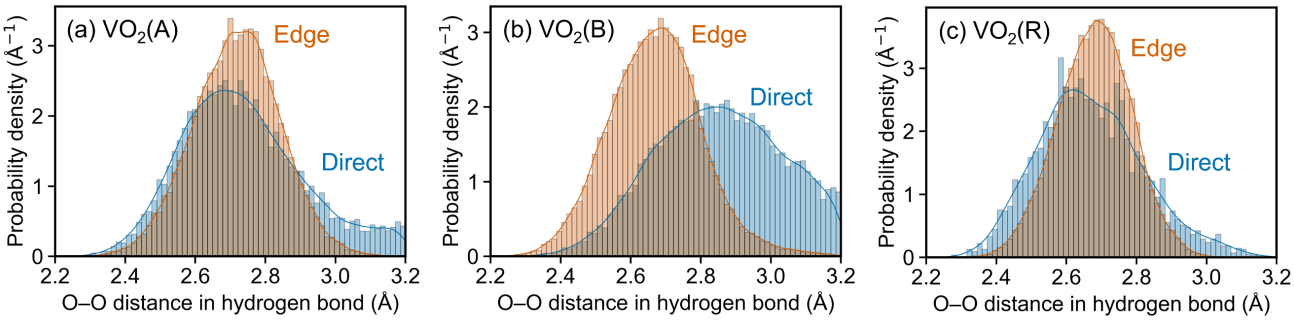

Figure S4. Statistical distributions of O–O distances involved in hydrogen-bond formation. Probability density distributions of O–O distances ($d_{O-O,HB}$) for direct-hop (blue) and edge-hop (orange) pathways in (a) VO_2(A), (b) VO_2(B), and (c) VO_2(R) extracted from MLIP-MD trajectories at 700 K. The histograms represent the raw statistical frequency, whereas the solid lines denote the Gaussian kernel density estimation (KDE) used to visualize the characteristic geometric preferences of each polymorph. The calculated mean distances for each migration motif are as follows: $d_{O-O,HB}$ = 2.74 Å for VO_2(A) direct-hop; 2.85 Å for VO_2(B) direct-hop and 2.67 Å for VO_2(B) edge-hop; and 2.88 Å for VO_2(R) edge-hop.

1 L. Barroso-Luque, M. Shuaibi, X. Fu, B. M. Wood, M. Dzamba, M. Gao, A. Rizvi, C. L. Zitnick and Z. W. Ulissi, *arXiv*, 2024, preprint, arXiv:arXiv:2410.12771, DOI: 10.48550/arXiv.2410.12771.
2 B. M. Wood, M. Dzamba, X. Fu, M. Gao, M. Shuaibi, L. Barroso-Luque, K. Abdelmaqsoud, V. Gharakhanyan, J. R. Kitchin, D. S. Levine, K. Michel, A. Sriram, T. Cohen, A. Das, A. Rizvi, S. J. Sahoo, Z. W. Ulissi and C. L. Zitnick, *arXiv*, 2025, preprint, arXiv:arXiv:2506.23971, DOI: 10.48550/arXiv.2506.23971.
3 G. Kresse and J. Furthmüller, *Comput. Mater. Sci.*, 1996, **6**, 15–50.
4 G. Kresse and J. Furthmüller, *Phys Rev B*, 1996, **54**, 11169–11186.
5 A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng and K. W. Jacobsen, *J. Phys. Condens. Matter*, 2017, **29**, 273002.
6 A. R. McCluskey, S. W. Coles and B. J. Morgan, *J. Chem. Theory Comput.*, 2025, **21**, 79–87.
7 A. R. McCluskey, A. G. Squires, J. Dunn, S. W. Coles and B. J. Morgan, *J. Open Source Softw.*, 2024, **9**, 5984.
8 J. G. Kirkwood, V. A. Lewinson and B. J. Alder, *J. Chem. Phys.*, 1952, **20**, 929–938.