

Hybrid sampling approach to machine-learning potentials for gas adsorption: Hydrogen adsorption in MOF-303

Kartik Sau ^{1,*} Ikutaro Hamada ^{2,†} Tamio Ikeshoji ³ Yiming Lu,¹ Susmita Roy ⁴ Shohichi Furukawa ¹
Linda Zhang ¹ Hung Ba Tran ¹ Takahiro Kondo ^{1,4,5,6} Hao Li ¹ and Shin-ichi Orimo ^{1,7,‡}

¹Advanced Institute for Materials Research (WPI-AIMR), Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan

²Department of Precision Engineering, Graduate School of Engineering,
The University of Osaka, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

³Mathematics for Advanced Materials Open Innovation Laboratory (MathAM-OIL),
National Institute of Advanced Industrial Science and Technology (AIST), c/o Advanced Institute for Materials Research (WPI-AIMR),
Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan

⁴Department of Materials Science, Institute of Pure and Applied Sciences,
University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan

⁵Hydrogen Boride Research Center, Tsukuba Institute for Advanced Research (TIAR),
University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan

⁶Tsukuba Research Center for Energy Materials Science, Institute of Pure and Applied Sciences and R&D Center for Zero CO₂
Emission with Functional Materials, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan

⁷Institute for Materials Research, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai, Miyagi 980-8577, Japan



(Received 15 January 2026; accepted 19 August 2026; published 5 October 2026)

Porous materials such as metal-organic frameworks (MOFs) are widely studied for applications in catalysis and gas adsorption, including carbon dioxide capture and hydrogen storage for energy and environmental challenges. Accurate simulation of gas adsorption in these materials remains challenging because most approaches rely on force fields (FFs), which lack first-principles accuracy. Here, we present a hybrid sampling approach, complemented by an active-learning-style refinement step, for generating machine-learning potentials (MLPs) tailored for grand canonical Monte Carlo (GCMC) simulations of gas adsorption. The training dataset combines snapshots from FF-based GCMC sampling, evaluated by single-point density functional theory (DFT) calculations, with trajectories from DFT-based molecular dynamics (MD). We demonstrate this approach for hydrogen adsorption in MOF-303, achieving accurate energy and force evaluation, as well as robust GCMC simulation including on-the-fly quantum corrections for nuclei. The MLP-based calculation also revealed four primary and four secondary adsorption sites formed near N, O, C, and –OH motifs, highlighting the material's structural heterogeneity and adsorption behavior. The resulting MLPs enable calculation of adsorption isotherms, isosteric heat values, and diffusion coefficients across temperatures. This hybrid-sampling framework bridges DFT-level accuracy and larger-scale simulation capability than *ab initio* MD, and can in principle be applied to other porous materials and gas types to support the design of materials for gas storage and separation.

DOI: [10.1103/d5b5-66b6](https://doi.org/10.1103/d5b5-66b6)

I. INTRODUCTION

The global energy sector is undergoing a significant shift driven by low-carbon emissions and the adoption of sustainable energy solutions. Within this changing landscape, hydrogen storage stands out due to hydrogen's high gravimetric energy density, renewability (via photo electrolysis), and environmentally friendly nature. The hydrogen

economy has witnessed substantial global growth over the past decade, fueling intensive research into efficient storage methods. Nonetheless, its volumetric energy density at ambient conditions is low, which represents the main challenge for practical storage. Porous materials mitigate this by concentrating hydrogen at pore surfaces to densities far exceeding those of the free gas at the same pressure, without incurring the substantial energy costs associated with compression or liquefaction. Metal-organic frameworks (MOFs) have emerged as highly promising materials for hydrogen storage owing to their crystalline nature formed by metal ions or clusters linked with organic ligands, creating extensive porous networks [1–6]. Notably, MOFs such as MOF-303 display outstanding features including high specific surface areas, large and adjustable pore volumes, intrinsic structural flexibility, and chemical tunability [7–9]. MOFs also offer significant advantages for hydrogen storage,

*Contact author: kartik.sau@gmail.com

†Contact author: ihamada@prec.eng.osaka-u.ac.jp

‡Contact author: shin-ichi.orimo.a6@tohoku.ac.jp

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

including their ability to undergo reversible and rapid hydrogen adsorption and desorption, features that are essential for practical applications. However, a major limitation of MOF-based hydrogen storage is its reliance on low temperatures to achieve efficient hydrogen uptake. This is largely due to the weak van der Waals forces between the MOF structure and hydrogen molecules, with adsorption energies typically ranging from 4–7 kJ mol^{−1} [10,11]. Such interaction strengths enable substantial hydrogen adsorption at cryogenic temperatures (around 77 K), but are insufficient for effective storage at ambient temperatures. To meet the U.S. Department of Energy’s targets for hydrogen storage at room temperature, adsorption energies within the range of 15–25 kJ mol^{−1} are necessary [12]. This gap highlights the need for developing advanced techniques to accurately predict and enhance hydrogen adsorption under ambient-temperature, and moderate-pressure conditions. Experimentally studying adsorption and desorption across a wide range of temperatures and pressures can be prohibitively expensive and time-consuming. Furthermore, atomistic details are often inaccessible by conventional experimental method, such as x-ray diffraction, especially for light hydrogen atoms. On the other hand, computational methods such as molecular dynamics (MD) and Monte Carlo (MC) simulations are powerful in determining and characterizing hydrogen adsorption sites in MOFs [13]. However, they must be combined with accurate energy and force calculations to predict adsorption properties reliably.

Simulations of gas adsorption in MOFs are typically performed using force fields (FFs), such as universal DREIDING FFs [14] and using Lorentz-Berthelot mixing rules [15–17]. However, the FFs often fail to reproduce the experimental energetics of MOFs [18]. To overcome this, polarizable FFs that explicitly consider adsorbate polarization effects at open metal sites via induced dipole moment have been developed. Yet, these attempts have still failed to fully capture the complexity of MOF adsorption behavior [19,20]. Recent progress in dispersion-inclusive exchange-correlation functionals in density functional theory (DFT) enables a high-fidelity description of the potential energy surface (PES) of molecules in the MOF framework, but their computational expense restricts their application to relatively small systems [21–23].

In recent years, machine-learning potentials (MLPs) have emerged as a promising approach that bridges the gap between accuracy and efficiency. By training on extensive *ab initio* MD (AIMD) datasets, MLPs can achieve near-DFT accuracy, while maintaining computational costs comparable to FF methods [24–26]. Applications of MLPs to porous materials, including MOF-5 [27], UiO-66 [28], MOF-47 [29], and zeolitic imidazolate frameworks (ZIFs) [29], have demonstrated their ability to perform larger-scale simulations than AIMD with high predictive accuracy. Although their adoption in MD simulations is widespread, their integration into the grand canonical MC (GCMC) simulations for gas adsorption presents unique challenges. Unlike MD simulations, which explore thermally accessible regions near energy minima of PES, GCMC simulations involve numerous trial insertions and deletions of molecules at random positions. This requires MLPs to be robust and accurate across a much wider range of PES, including high-energy configurations probed by GCMC trial insertions that must be evaluated reliably even when

ultimately rejected [29], and configurations spanning a broad range of H₂ loadings configuration that AIMD, which operates at a fixed number of molecules, cannot access. Several strategies can partially address this challenge, including balanced energy and force supervision [30], exclusion spheres that preserve detailed balance while preventing unphysical overlaps [29], and active-learning schemes that iteratively enrich training sets using configurations associated with high model uncertainty. Active learning is a powerful and well-established route to robust potentials, but it generally requires many repeated cycles of model training and sampling, and it has so far been applied mainly to MLP-based MD simulations [31]; its use for MLP-based GCMC, where the relevant configurational space is loading-dependent, remains comparatively underexplored. A complementary strategy is to build the required coverage into the initial training set: If the initial training set is constructed so that it already spans the loadings, energies, and local environments that GCMC will visit, the number of subsequent enrichment cycles needed to reach a robust potential is correspondingly reduced. This motivates a practical data-generation strategy that produces MLPs robust to GCMC trial moves while capturing configurations across a wide range of H₂ loadings, temperatures, and local chemical environments, with targeted enrichment reserved for the regions that the initial construction leaves undersupported.

In this work, we develop a hybrid data generation flow, which combines the GCMC sampling based on the FFs followed by DFT energy and force evaluation, and AIMD sampling at varying temperatures. We demonstrate the validity of this hybrid sampling approach by applying it to hydrogen adsorption in MOF-303. The MOF-303 was chosen because its chemically heterogeneous pore environment, comprising Al-OH chains, pyrazole linkers, and polar functional groups, provides a stringent test of the MLP’s ability to capture diverse host-guest interactions, and because experimental adsorption data are available for validation [8,9]. Through MLP-GCMC simulations, hydrogen uptake is predicted under various pressures and temperatures, accurately reproducing the heat of adsorption at cryogenic temperatures. The MLPs further enable MD simulations for topological, structural, and dynamical analyses of hydrogen adsorption in MOF-303.

II. METHODOLOGY

The overall hybrid sampling workflow is schematically depicted in Fig. 1. We first perform GCMC simulations employing FFs to sample a wide-ranging configurational space. Subsequently, single-point DFT energies and forces are evaluated based on the configurations generated in the GCMC simulations. This DFT dataset is then augmented with data from AIMD simulations performed at various temperatures. All the combined datasets are used to generate MLPs based on a message-passing graph neural network architecture, which accurately models many-body interactions governing H₂ adsorption in the porous material. The trained MLPs are tested using a separate DFT data set, which was not included in the training. Finally, the MLPs are used in the production GCMC simulations to predict thermodynamic adsorption isotherms and MD simulations to analyze static and dynamical properties of H₂ in the porous material.

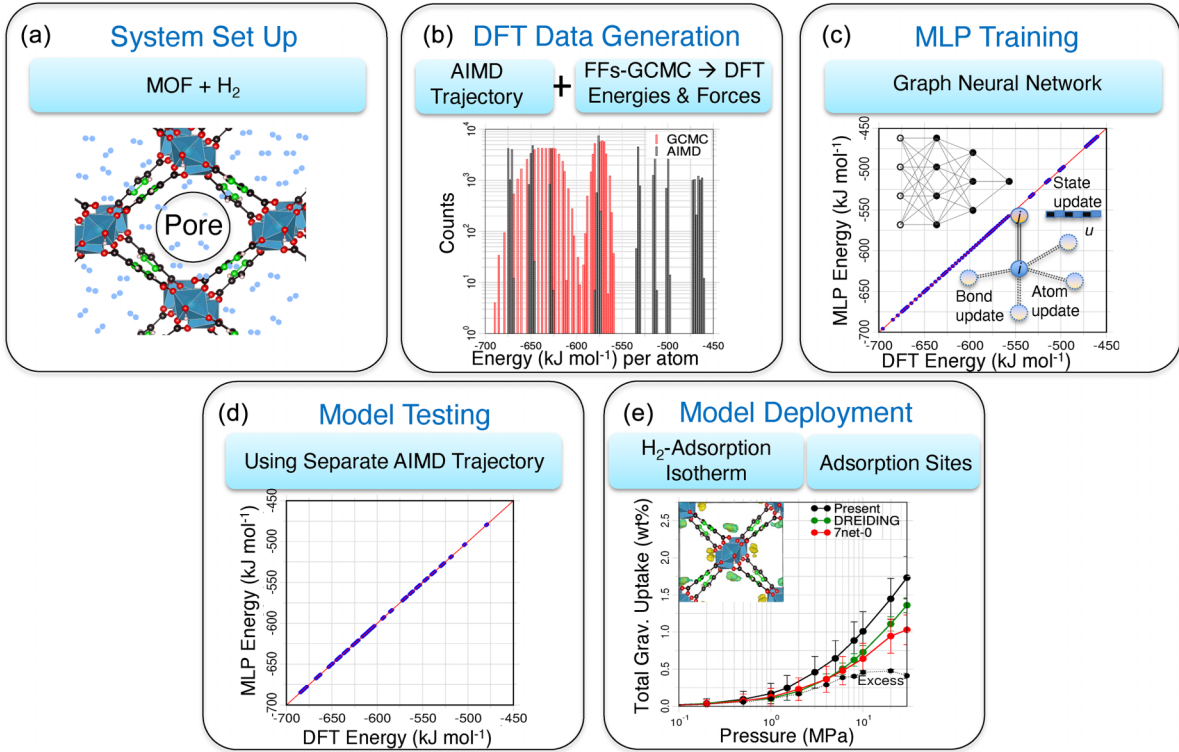


FIG. 1. *MLP development and deployment workflow.* (a) MOF-303 structure with H_2 molecules. (b) classical FF-based GCMC sampling across pressures to generate diverse H_2 configurations, followed by DFT energy and force evaluations. The data is augmented by AIMD sampling at various temperatures. $\approx 60\%$ DFT energies and forces based on the GCMC sampling; $\approx 40\%$ DFT energies and forces based on the AIMD sampling (77–900 K). (c) MLP-model training using graph neural network. (d) MLP-model testing against AIMD data, which were not used for the training. (e) Model deployment in both MLP-based GCMC for H_2 uptake and adsorption isotherms, and MLP-based MD for analyses.

A. Hybrid Training Data Generation and MLP development

We employed the DREIDING FFs (see Table SI in the Supplemental Material [32] for parameters) and performed GCMC simulations of H_2 adsorption in MOF-303 using the experimental crystal structure (Table I). The DREIDING model used here comprises Lennard-Jones (12-6) terms only, with no partial charges assigned to the framework or to H_2 , so electrostatic interactions are absent—a known limitation of universal FFs for polar frameworks. The MLP, being trained on self-consistent PBE-D3 data, implicitly captures electrostatics, polarization, and many-body dispersion without an explicit charge model. The wide range of H_2 -pressure ensured comprehensive coverage of H_2 densities, from low (0 wt %) to high (11 wt %), under pressures of 0–0.1 MPa at 77 K and 0–30 MPa at 298 K. It is also important to note

TABLE I. PBE-D3-optimized cell parameters (a , b , c , α , β , and γ) and equilibrium volume (V) of MOF-303 ($C_5H_3AlN_2O_5$), along with the experimental values taken from Ref. [7]. The unit cell contains eight formula units.

	a (Å)	b (Å)	c (Å)	α (deg)	β (deg)	γ (deg)	V (Å ³)
PBE-D3	12.006	15.340	13.806	90.0	101.85	90.00	2488.6
Expt.	12.081	15.347	13.873	90.0	102.17	90.00	2514.4

that all the sampled configurations lie within the interpolated region of the trained energy landscape. A complete adsorption isotherm was thereby generated. We used the LAMMPS software package [33] for the FF-based GCMC simulations with a tabular potential, such that nuclear quantum effects (NQE) can be integrated in LAMMPS GCMC. Snapshots are retained from the full GCMC trajectory, including preequilibrium states, so that the training data span from near-zero to target loading; this naturally covers a broader region of the PES [Fig. 1(b)] than trajectories at a fixed molecule number (Fig. S1) [32].

It is worth mentioning that the DREIDING FFs treat H_2 molecules as single atoms. After the GCMC simulations, we regenerated the H_2 molecules with random orientations and a bond length of 0.753 Å (DFT-D3 relaxed bond) for subsequent DFT calculations. We performed additional GCMC simulations using a two-point H_2 model with explicit rotational moves (see Fig. S5 in the Supplemental Material [32]). The FF-based GCMC simulations serve as a data generation tool to produce diverse H_2 configurations spanning a wide range of loadings and temperatures. The orientational degree of freedom is extensively sampled through the many random orientations assigned during regeneration and the AIMD trajectories at multiple temperatures which naturally explore all orientational configurations. In the production of MLP-GCMC simulations, a two-point H_2 representation with explicit rotational moves was used.

Using the FF-based GCMC configurations, single-point DFT calculations were performed. The projector augmented-wave (PAW) method [34] as implemented in the Vienna *ab initio* Simulation Package [35,36] with the standard PAW potentials was used for the DFT calculations. The exchange-correlation energy and potential were described using the Perdew-Burke-Ernzerhof [37] (PBE) generalized gradient approximation with Grimme's dispersion correction [38] (PBE-D3). Kohn-Sham orbitals were expanded in terms of a plane-wave basis with a kinetic energy cutoff of 480 eV, and the Brillouin zone was sampled using the Γ point. In addition to the configurations generated in the GCMC simulations, we performed AIMD simulations with varying numbers of H_2 molecules at different temperatures (77, 298, and 900 K). The training data consists of GCMC- and AIMD-based simulation data containing approximately 200 000 snapshots, which robustly represent a wide range of PES.

Figure S1 [32] compares the configurations obtained from FF-GCMC and AIMD sampling. FF-GCMC spans a substantially broader range of H_2 loadings and energies than AIMD [Figs. S1(a) and S1(b)] [32]. At representative loadings of 5, 10, and 15 H_2 molecules per unit cell, the GCMC-derived configurations also provide complementary coverage of moderately higher-energy regions of the potential-energy surface [Fig. S1(c)]. Together, these results show that FFs-GCMC and AIMD sample distinct but complementary configurational regions, motivating their combination in the hybrid training dataset.

The data was then used to train MLPs, for which we employed the neural equivariant interaction potential (NequIP), a model based on an E(3)-equivariant neural network [39]. In this work, SevenNet [40] was used, which is an efficient implementation of message-passing graph-neural network MLPs with NequIP. We used a cutoff radius of 5 Å, four interaction layers, 32 features, and a maximum rotation order of one. Both energies and forces were included in the training process, while stress tensors were excluded, as the MOF framework was treated as rigid in this study.

B. MLP-GCMC simulations

The trained MLP was deployed to perform GCMC simulations of H_2 adsorption in a fixed MOF-303 framework using software developed by Goeminne *et al.* [29] via the Atomic Simulation Environment [41] (ASE). For the present MLP, the vibrational free-energy correction was evaluated on the fly for each adsorbed H_2 molecule and included in the GCMC acceptance probability. The normal-mode treatment, and the soft translational, librational, and intramolecular stretching modes are described in the Supplemental Material, Appendix A [32]. This correction was not applied to the DREIDING and SevenNet-0 reference models because their vibrational descriptions were not assessed for this purpose. All three models (present MLP, DREIDING, 7net-0) were evaluated with the Feynman-Hibbs (FH) effective-potential correction, which accounts for the translational quantum delocalization of H_2 . The FH approach replaces the potential $U(r)$ with a temperature (T)-dependent effective potential [42,43]:

$$U_{\text{FH}}(r) = U(r) + \frac{\hbar^2}{24 \mu_{\text{H}_2} k_B T} \nabla^2 U(r), \quad (1)$$

where μ_{H_2} is the molecular mass of the H_2 molecule (the mass appropriate for center-of-mass motion against the rigid framework; because each $\text{H}_2 - \text{H}_2$ pair enters the Laplacians of both partners, the per-molecule sum with the molecular mass reproduces the reduced-mass pair form once per pair), k_B is the Boltzmann constant, and $\hbar$ is the reduced Planck constant. The quantum corrections by FH and vibrations account for the nuclear delocalization of H_2 , effectively enlarging the molecule at low temperatures and reducing the predicted uptake compared to without NQE-corrected GCMC. For the present MLP, the FH term is evaluated on the fly from central finite differences of the MLP forces and applied to the center-of-mass (translational) degrees of freedom only; the librational and intramolecular stretching modes are instead treated by the vibrational free-energy correction described below and in Appendix A of the Supplemental Material [32].

In the MLP-based GCMC simulations, insertion, deletion, translation, and rotation moves were attempted with equal probability (25% each). To improve low-temperature sampling, translation and rotation amplitudes were adaptively adjusted every fixed number of steps toward a target acceptance rate of 0.3–0.5, increasing or decreasing by 10% within prescribed bounds. Similarly, at low temperature, where random insertions are rarely accepted, Rosenbluth configurational-bias Monte Carlo (CBMC) [44] was used to accelerate pore filling. The decision to accept or reject these moves was made by calculating the Boltzmann probability of different configurations. This move ratio ensured efficient sampling of the MC ensemble. GCMC naturally allows the number of H_2 molecules fluctuates through exchange with a reservoir at a fixed chemical potential (μ), determined by the gas pressure. The simulation consists of 10^6 steps for the production run to ensure convergence of H_2 uptake, with a few hundred thousand preproduction steps. 5×10^4 configurations were taken from the equilibrium windows for ensemble averaging. For comparison, we also performed GCMC simulations by employing FFs and pretrained graph-neural network interatomic potentials (SevenNet-0). We benchmarked the computational cost of the MLP and the FFs under identical conditions (MOF-303 with H_2 , 2000 GCMC steps, 77 K, 0.4 bar), in both cases without the quantum correction. The FF-based simulation, running on a single CPU core (AMD EPYC 9124, 3.0 GHz), was completed in ~ 44 s (~ 0.022 s/step), while the present MLP, GPU-accelerated on an NVIDIA RTX A6000 (48 GB, 2.1 GHz) required ~ 1200 s (~ 0.6 s/step), a speed advantage for the classical force field. Despite this additional computational cost, the MLP-based GCMC remains computationally practical; a full isotherm calculation completes within a few days, with the pressure points run concurrently.

C. H_2 uptake calculation

Based on the GCMC simulations at constant μ , we determined the average number of adsorbed H_2 molecules in the MOF framework at a particular temperature and H_2 -pressure.

The total gravimetric H_2 uptake (θ_T) was calculated as

$$\theta_T = \frac{m_{\text{H}_2} \langle N \rangle}{m_{\text{MOF}} + m_{\text{H}_2} \langle N \rangle}, \quad (2)$$

where m_{H_2} is the atomic mass of H_2 , N is the number of H_2 molecules in the unit cell, and m_{MOF} is the mass of the MOF-303 unit cell. The excess H_2 uptake (θ_{Exc}) was derived from the total uptake by subtracting the amount of bulk gas present in the pore volume [45]:

$$\theta_{\text{Exc}} = \theta_{\text{T}} - \frac{m_{\text{H}_2} \rho_{\text{bulk}} V_{\text{P}}}{m_{\text{MOF}} + m_{\text{H}_2} \langle N \rangle}, \quad (3)$$

where ρ_{bulk} represents the number density of bulk H_2 gas at the corresponding temperature and pressure obtained from the National Institute of Standards and Technology (NIST) database, and V_{P} is the specific pore volume of MOF-303, calculated using the PoreBlazer [46] software.

D. Isostatic heat calculation

Here, we calculated the isosteric heat of adsorption from GCMC simulation data using the fluctuation formula defined as

$$Q_{\text{st}} = -\frac{\langle UN \rangle - \langle U \rangle \langle N \rangle}{\langle N^2 \rangle - \langle N \rangle^2} + RT, \quad (4)$$

where R is the gas constant, and $\langle \dots \rangle$ indicates the ensemble average. Alternatively, the isosteric heat of adsorption at infinite dilution $Q_{\text{st,dilution}}$ can also be calculated using the Widom insertion method [47], by performing a set of random insertions of the adsorbate in the framework and computing the interaction energy ΔE . The expression is given as

$$Q_{\text{st,dilution}} = -\frac{\langle \Delta E e^{-\frac{\Delta E}{k_{\text{B}}T}} \rangle}{\langle e^{-\frac{\Delta E}{k_{\text{B}}T}} \rangle} + RT. \quad (5)$$

We employ two complementary NQE approaches for Q_{st} and compare their effects: (i) on the fly FH + vib effective potential correction applied during GCMC sampling (Sec. II B), which accounts for the quantum delocalization of the H_2 center of mass and yields a quantum-corrected Q_{st} directly from the GCMC; and (ii) the vibrational free energy correction described below, which captures the change in H_2 vibrational energy upon adsorption and is applied as a *post-hoc* correction to the computed $Q_{\text{st,dilution}}$. FH is a semiclassical correction: It involves only the Laplacian of the potential with respect to the molecular center-of-mass coordinates and is therefore blind, by construction, to the orientational (librational) and intramolecular (stretching) degrees of freedom, whose quanta ($\sim 30 - 45$ meV and ~ 0.52 eV, respectively) greatly exceed $k_{\text{B}}T$ at 77 K. To calculate the vibrational frequencies of the adsorbed H_2 molecule, we placed one H_2 molecule at the primary adsorption site and performed relaxation of H_2 molecule, keeping the MOF-303 structure fixed. The six modes of each molecule are classified by eigenvector projection into translational, librational, and stretching character, and each group is treated by the appropriate method (FH for translations; full quantum harmonic free energy for librations, with a librational floor; and the adsorbed-vs-gas quantum free-energy difference for the stretch); the complete scheme and its validation against DFT-D3 frequencies are given in Appendix A of the Supplemental Material [32].

The finite temperature vibrational energy and vibrational entropy contribution were estimated as the difference between

those of adsorbed and gas-phase H_2 molecules. We also applied quantum vibrational free energy correction to the isosteric heat (in the case without NQE correction in the GCMC simulation) of adsorption as follows [48]:

$$Q_{\text{st}} \rightarrow Q_{\text{st}} + \Delta U_{\text{vib}} - T \Delta S_{\text{vib}}, \quad (6)$$

$$\Delta U_{\text{vib}} = U_{\text{vib}}^{\text{gas}} - U_{\text{vib}}^{\text{ads}}, \quad (7)$$

$$U_{\text{vib}} = \sum_{i=1}^n \left[\frac{1}{2} \hbar \omega_i + \frac{\hbar \omega_i}{e^{\hbar \omega_i / k_{\text{B}}T} - 1} \right], \quad (8)$$

where $U_{\text{vib}}^{\text{ads}}$ and $U_{\text{vib}}^{\text{gas}}$ correspond to the vibrational energies of the adsorbed and gas-phase H_2 molecules, respectively. Here, ω_i represents the frequency of the i -th vibrational mode. $\Delta S_{\text{vib}} = S_{\text{vib}}^{\text{gas}} - S_{\text{vib}}^{\text{ads}}$ is the change in vibrational entropy upon adsorption, where $S_{\text{vib}}^{\text{ads}}$ and $S_{\text{vib}}^{\text{gas}}$ are the vibrational entropies for the gas phase and adsorbed H_2 , respectively. The vibrational entropy was calculated using the harmonic approximation, summed over the same n discrete vibrational modes as U_{vib} as [49]

$$S_{\text{vib}}(p, T) = \sum_{i=1}^n k_{\text{B}} \ln \left[2 \sinh \left(\frac{\hbar \omega_i}{2k_{\text{B}}T} \right) \right] + \sum_{i=1}^n \frac{\hbar \omega_i}{2T} \coth \left(\frac{\hbar \omega_i}{2k_{\text{B}}T} \right). \quad (9)$$

E. Basin-hopping and Monte Carlo search

We investigated stable adsorption sites for a single H_2 molecule by employing the trained MLP and the basin-hopping (BH) global optimization search [50,51] using our in-house script based on ASE. The H_2 molecule was treated as a rigid body, and the MOF-303 framework was kept fixed. To efficiently explore the potential energy surface, we performed on the order of several hundred thousand steps at 77 K, each step consisting of a random translation and rotation of the H_2 molecule followed by local energy minimization. An exclusion sphere of 1.5 Å was used to prevent unphysical overlaps between the H_2 molecule and framework atoms. For comparison, we also performed standard MC simulations for a single H_2 molecule at 77 K using the Metropolis algorithm. In these simulations, trial moves consisting of random translations and rotations were accepted or rejected based on the Boltzmann probability, without the local minimization step used in BH.

F. MLP-MD simulations

The trained MLP was deployed to perform molecular dynamics (MD) simulations in the NVT ensemble using LAMMPS [40] via the SevenNet interface, enabling the investigation of the structural and dynamical properties of H_2 within the MOF-303 framework. The MD simulations were performed for several hundred picoseconds with a time step of 0.5 fs at temperatures of 77, 100, 150, 200, 250, and 298 K, with the framework kept fixed. A Langevin thermostat was employed to control the temperature. To ensure sufficient statistical sampling for the diffusion analysis, we used a $(2 \times 2 \times 2)$ supercell containing 64 H_2 molecules (corresponding to 8 H_2 molecules per primitive cell).

TABLE II. Accessible pore volume (V_{acc}), accessible surface area (ASA), pore limiting diameter (PLD), and largest cavity diameter (LCD) of MOF-303 calculated using PoreBlazer [46].

Probe	V_{acc} ($\text{cm}^3 \text{g}^{-1}$)	ASA ($\text{m}^2 \text{g}^{-1}$)	PLD ($\text{\AA}$)	LCD ($\text{\AA}$)
H ₂	0.55	1518	6.1	5.53
N ₂	0.46	1120	6.1	5.53

Kinetic properties of H₂ molecules within MOF-303 were investigated using mean-squared displacement (MSD) analysis from MLP-MD trajectories. The self-diffusion coefficient D along the cylindrical pore was obtained from the one-dimensional Einstein relation applied to the MSD of a hydrogen atom in the H₂ molecule. We used the one-dimensional form because H₂ diffusion in MOF-303 predominantly occurs along the cylindrical channel (a direction). The activation energy for diffusion (E_a) was extracted from an Arrhenius fit of the temperature-dependent diffusion coefficients. The diffusion coefficient along the cylindrical pore was calculated using the Einstein relation for a one-dimensional system as $D = \lim_{\Delta t \rightarrow \infty} \text{MSD}/2\Delta t$, where Δt is the time interval.

III. RESULTS AND DISCUSSION

A. Structure and topological properties of MOF-303

MOF-303 is a robust Al-based MOF with composition $[\text{Al}(\text{OH})(\text{C}_5\text{H}_2\text{N}_2\text{O}_2)]$, in which corner-sharing AlO_6 octahedra form rodlike structures along a axis and pyrazole linkers connect these structures, forming a three-dimensional framework with one-dimensional cylindrical channel [Fig. 4(a)]. Within each pore, there are two symmetric “pockets” formed by the O atom of the hydroxyl ($-\text{OH}$) group, and N and C atoms from the pyrazole linkers. The one-dimensional channels are lined with the OH and N atoms from the pyrazole linkers, which can interact with H₂ molecules through weak van der Waals forces. The presence of these functional groups enhances the H₂ adsorption capacity of MOF-303. DFT-optimized cell parameters are shown in Table I, which agree well with the experimental data within the deviation of 1.0% (Table I). Throughout this work, however, we used the experimental cell parameters because the differences are insignificant and to preserve the crystal symmetry. We calculated the topological details of MOF-303 using PoreBlazer [46] with H₂ and N₂ (Table II). The calculated surface area ($1120 \text{ m}^2 \text{g}^{-1}$), pore volume ($0.46 \text{ cm}^3 \text{g}^{-1}$), and pore diameter ($6.1 \text{ \AA}$) obtained using N₂ as probe are in good agreement with the respective experimental values of $1280\text{--}1360 \text{ m}^2 \text{g}^{-1}$, $0.48\text{--}0.51 \text{ cm}^3 \text{g}^{-1}$ [7,52], and $\approx 6 \text{ \AA}$ [53]. The high surface area and moderate pore size suggest that MOF-303 is a promising candidate for H₂ storage. The topological properties were used to calculate the excess H₂ uptake.

B. MLP testing

The developed MLP using the workflow outlined in Sec. II achieves a mean absolute error (MAE) of 0.02 kJ mol^{-1} for energies per atom and $0.5 \text{ kJ mol}^{-1} \text{ \AA}^{-1}$ for forces on the training dataset (parity plots of energies and forces are shown

in Fig. S2 in the Supplemental Material [32]). To further test the MLPs, we evaluated their performance on a separate test set of an AIMD trajectory and the MLP-trained model-based GCMC followed by DFT single-point energy that were excluded from the training dataset [Fig. 1(d)]. Parity plots comparing predicted and reference energies and forces are provided in Fig. S3 in the Supplemental Material [32]. Our MLPs achieve an MAE of 0.04 kJ mol^{-1} for energies per atom and $0.4 \text{ kJ mol}^{-1} \text{ \AA}^{-1}$ for forces in the test dataset, demonstrating their ability to reproduce DFT-level accuracy across a diverse set of configurations [Fig. 1(b)]. The agreement of the MLP and DFT-D3 vibrational frequencies at the relaxed adsorption minimum is reported in Table SII of the Supplemental Material [32]. The accuracy of the energies and forces indicates that our MLPs can reliably capture the complex interactions between H₂ and the MOF framework, making it suitable for further analyses. To further validate the dynamic reliability of the MLP, we compare partial pair distribution functions (PDFs) between the MLP-MD and AIMD trajectories at 298 K, using a single unit cell with eight H₂ molecules (Fig. S4 in the Supplemental Material [32]). The MLP-MD PDFs show excellent agreement with the AIMD reference, confirming that the MLP faithfully reproduces the structural correlations between H₂ and the MOF framework during dynamics, beyond static energy and force accuracy.

C. H₂ adsorption isotherms

We performed MLP-GCMC simulations up to 30 MPa H₂ pressure to assess the predicted adsorption isotherms. For comparison, we performed simulations using DREIDING FFs and SevenNet-0, a pretrained graph-neural network MLP based on SevenNet. Figure 2 presents the total gravimetric H₂ adsorption isotherms, calculated according to Eq. (2), at 77 and 298 K. At 77 K, H₂ uptake is significantly high, with a steep isotherm, indicating abrupt filling of adsorption sites as binding energetics overcome thermal fluctuations [Fig. 2(a)]. Our particular interest is adsorption at ambient temperature (298 K) for practical storage. At 10 MPa and 298 K, the present model predicts an uptake of 0.86 wt% versus 0.75 wt% (parameter-based) and 0.65 wt% (pretrained). The excess H₂ uptake at 298 K [Eq. (3)] saturates below 20 MPa, indicating that the adsorption capacity of MOF-303 is effectively reached under these conditions. For completeness, we also show the uptake without NQE corrections for all model cases (Fig. S5 in the Supplemental Material [32]; confirming that the NQE correction consistently reduces the predicted uptake. In addition, FFs-based GCMC simulations using a two-point H₂ model yield an isotherm nearly identical to that of the single-point model, validating the use of the single-site FFs model for data generation. We also use a supercell to ensure the convergence of the predicted isotherm with respect to the cell size using the present model (Fig. S6 in the Supplemental Material) [32].

We compared our MLP-predicted isotherms with the experimental data of Wang *et al.* [8] at 77 K [Fig. 2(a)] up to 1 bar. The present MLP closely resembles the experimentally reported pressure-dependent uptake. We note that the central aim of this work is to develop an MLP that faithfully reproduces DFT-level energetics at computational costs

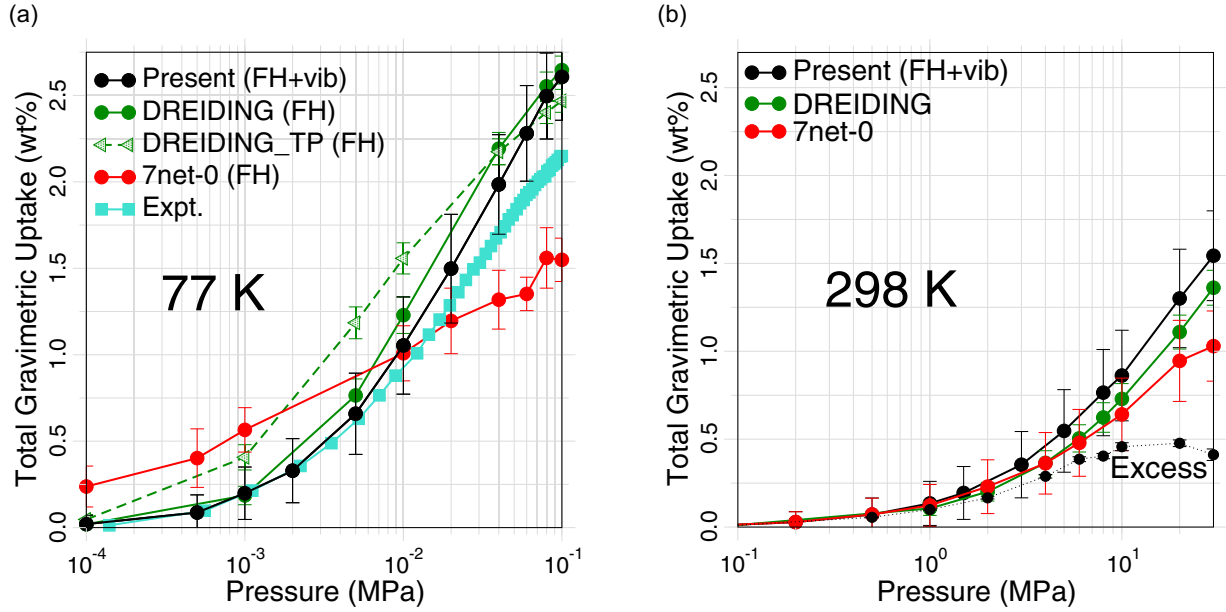


FIG. 2. Calculated H_2 uptake for MOF-303. Calculated gravimetric H_2 uptake for MOF-303 from GCMC simulations employing different potential models (present model, DREIDING FF, and pretrained 7net-0) at (a) 77 K and (b) 298 K. Experimental data from Wang *et al.* [8] at 77 K are included for comparison. Excess H_2 uptake at 298 K is also shown in (b). The FH and vibrational quantum corrections is implemented in the present MLP at both the temperature, whereas only FH correction is applied to DREIDING, and 7net-0 here at 77 K (without at 298 K). Error bars denote the statistical uncertainty of the present MLP calculations.

far below direct DFT, offering a practical alternative to both classical force fields and prohibitively expensive AIMD-based simulations. Comparison with experiment serves primarily to confirm that the PBE-D3 reference is physically reasonable and that the MLP reproduces it accurately. To rationalize the trend of the predicted adsorption isotherm, we calculated Q_{st} , which quantifies the average H_2 -framework interaction strength and governs H_2 uptake. Table III summarizes isosteric heat at 77 K obtained using the fluctuation formula (Q_{st}) and that using the Widom insertion method ($Q_{st,dilution}$), both with and without NQE taken into account. Calculated Q_{st} 's using FFs and pretrained SevenNet-0 deviate significantly compared to the reported experimental Q_{st} value; furthermore, the vibrational energy corrections have not been incorporated. By taking into account NQE, the Q_{st} predicted by our MLP

shows reasonable agreement with the reported experimental value of 6.44 kJ mol^{-1} [8]. For the present MLP system, the $\Delta U_{vib} - T\Delta S_{vib}$ correction evaluated at the adsorption site using the Widom insertion method is in reasonable agreement with the FH + vib correction (6.0 vs 6.4 kJ mol^{-1}), since in the Widom approach the inserted H_2 is only sampled in the vicinity of the adsorption site. The MLP-predicted Q_{st} decreases with increasing H_2 pressure at 77 K [Fig. 3], consistent with the progressive filling of less energetically favorable adsorption sites at higher loadings, in agreement with the experimental trend reported by Wang *et al.* [8]. In contrast, the DREIDING FF predicts a nearly pressure-independent Q_{st} that is systematically lower than the experimental values, reflecting its inability to capture the heterogeneous, many-body nature of the $H_2 - H_2$, and H_2 -MOF interactions. The Q_{st}

TABLE III. Calculated isosteric heat of H_2 adsorption from GCMC simulations at 77 K for H_2 pressure 0.001 MPa using the fluctuation theorem (Q_{st}) and Widom insertion method ($Q_{st,dilution}$), employing different interatomic potentials. “w/o NQE” denotes values without any quantum correction. The remaining columns report two distinct quantum treatments. (i) “NQE”: quantum corrections included in the GCMC sampling itself, namely the on-the-fly, mode-resolved vibrational free-energy correction together with the FH effective potential for the present MLP (FH+vib), and the FH correction alone for DREIDING and SevenNet-0 (FH). (ii) $\Delta U_{vib} - T\Delta S_{vib}$: a *post-hoc* correction [Eq. (6)] evaluated from the vibrational frequencies of a single H_2 molecule at the primary adsorption site (Table SII of the Supplemental Material) [32] and added to the classical value. Uncertainties denote the statistical error of the GCMC averages; $Q_{st,dilution}$ is a single-value estimate and is quoted without an uncertainty.

Potential model	$Q_{st} \text{ (kJ mol}^{-1}\text{)}$		$Q_{st,dilution} \text{ (kJ mol}^{-1}\text{)}$	
	w/o NQE	w/ NQE	w/o NQE	w/ NQE $\Delta U_{vib} - T\Delta S_{vib}$
Present	8.4 ± 0.1	6.4 ± 0.1 (FH+vib)	11.8	6.0
DREIDING	5.7 ± 0.0	5.3 ± 0.0 (FH)	10.0	4.2
SevenNet-0	7.9 ± 0.1	6.4 ± 0.1 (FH)	9.5	3.7

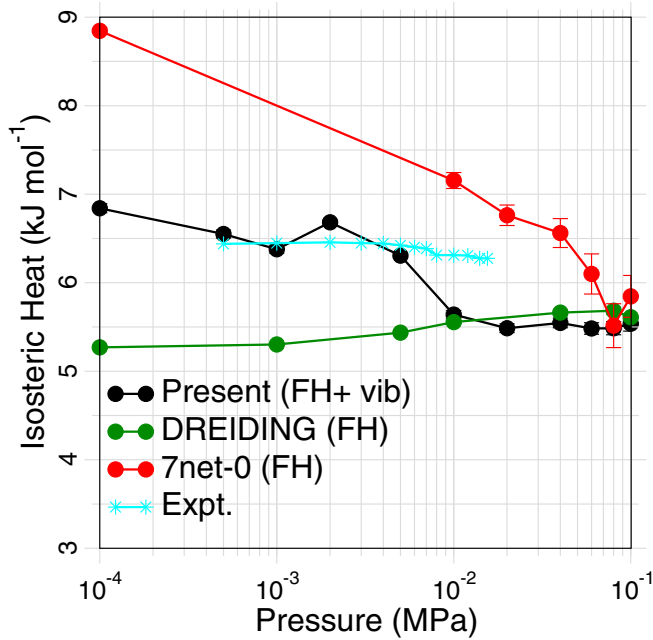


FIG. 3. *H₂-loading dependence of binding strength.* Isosteric heat of adsorption (Q_{st}) as a function of H_2 pressure at 77 K, comparing the present MLP, DREIDING, and SevenNet-0 (7net-0) predictions with experimental data from Wang *et al.* [8]. The FH and vibration-free energy correction is applied in the present MLP, and only FH is applied for the other two model cases. Error bars denote statistical uncertainties.

values prior to the FH correction are expected to be higher, as shown in Fig. S7 in the Supplemental Material [32], with minor deviations likely arising from numerical errors. We note that the present MLP inherently captures polarization and many-body effects, being trained on DFT data where electronic polarization is included self-consistently, thereby going beyond both nonpolarizable and pairwise-additive polarizable FFs [19,20].

D. H_2 adsorption and distribution

Having established the thermodynamic properties of H_2 adsorption in MOF-303, we used a single H_2 molecule to precisely identify its adsorption sites. We used the BH as described in Sec. II to find adsorption sites. The BH search revealed several sites in the cylindrical pore. Among them, two are prominent adsorption sites within the symmetric pockets: the primary site (four) with an adsorption energy of 13.89 kJ mol⁻¹ and the secondary site (four) with an adsorption energy of 12.85 kJ mol⁻¹, approximately 1 kJ mol⁻¹ smaller than that for the primary site. Even though the difference between the primary and secondary sites is marginal, they are different in their local environment. The adsorption sites are illustrated in Fig. S8(a) in the Supplemental Material [32]. It is worth noting that MC simulations performed with a single H_2 molecule resulted in the molecule remaining confined within a single cylindrical pore (Fig. S8(b) in the Supplemental Material [32]), unable to overcome the high energy barriers between adjacent pores. In contrast, the BH approach successfully overcomes these barriers, exploring the entire unit

cell and identifying the energy minima across all pores. The primary adsorption site lies between two symmetric triangles formed with the O atom of the OH group and N atoms, while the secondary adsorption site is between triangles with the O atom of the OH group and N atoms and the O atom of the OH group and C atoms, both near the OH groups.

To further understand the molecular distribution in the presence of multiple H_2 molecules at a finite temperature, we performed MLP-MD simulations of MOF-303 containing eight H_2 molecules in a primitive unit cell at 77 and 298 K, and mapped the three-dimensional H_2 occupancy (Fig. 4). Figure 4(a) displays the H_2 probability densities obtained from MLP-MD simulations at 77 K, where four H_2 molecules were initially placed at the primary and secondary adsorption sites. The compact H_2 distributions indicate that H_2 molecules are highly localized in the symmetric “pockets”, consistent with those obtained for a single H_2 molecule using BH and MC search. Our calculated spatial distributions are consistent with the H_2 adsorption probability maps reported by Wang *et al.* [8] based on classical FF simulations. However, their study did not resolve individual adsorption sites, likely due to the limitations of the classical FFs. In contrast, our MLP identifies both primary and secondary sites with atomistic precision.

At 298 K [Fig. 4(c)], the H_2 occupancy becomes significantly delocalized along the pore surface. The broadly distributed low H_2 occupancy indicated by the isosurface with the low isovalue (1.1×10^{-6} , yellow) suggests the delocalized nature of H_2 in the pore, while maintaining the localized occupation near the primary and secondary sites (isosurface with the high isovalue of 2.5×10^{-6}). This temperature evolution suggests a transition from an enthalpy-dominated, localized regime at cryogenic temperatures to an entropy-driven regime at ambient temperature, where molecules explore a larger portion of the pore surface. Partial pair distribution functions (PDFs) between the H atoms of H_2 and the framework atoms were calculated from the same MLP-MD trajectories (Fig. S9). At 77 K [Fig. S9(a)] [32] they show sharp peaks for O2, N1, and N2, and for the OH and pyrazolate hydrogens (H1 and H2), consistent with strong interactions at the polar motifs. At 298 K [Fig. S9(b)] [32] all the peaks in PDFs broaden, and shift slightly outward, reflecting increased thermal motion and reduced adsorption strength. However, the peaks in PDFs for framework-H (H2), O, and N persist, indicating that the primary adsorption at the polar motif dominates even at 298 K, albeit with spatial delocalization. Taken together, the PDFs and occupancy maps provide a coherent picture across temperature and site type: Primary pockets are governed by O/N-centered polar interactions and are highly localized at 77 K; secondary pockets involve additional sampling near C with broader distributions; and thermal broadening at 298 K weakens and delocalizes the same motif-centered interactions without changing their qualitative hierarchy.

E. H_2 diffusion

To estimate the diffusional activation energy (E_a) between the pores/sites, we first calculated the diffusion coefficients (D) of H_2 in MOF-303 using $(2 \times 2 \times 2)$ supercell for different temperatures (77–298 K) using the mean squared displacement (MSD) from the MLP-MD trajectories (Fig. 5).

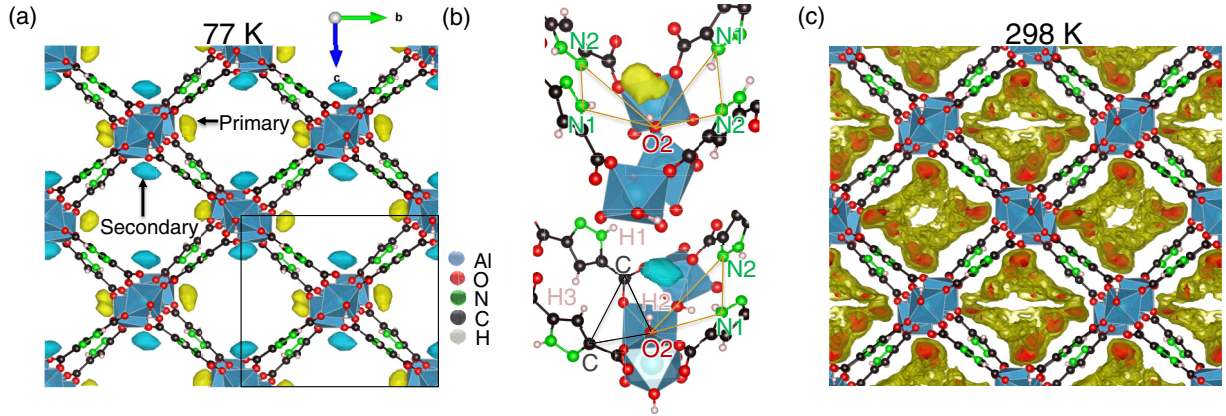


FIG. 4. *Spatial distribution of H_2 in MOF-303.* (a) Isosurface plot of the three-dimensional H_2 probability density in MOF-303 at 77 K, calculated from MLP-MD trajectories with eight H_2 molecules per unit cell (four initially at primary sites, four at secondary sites). The structure is replicated to a $2 \times 2 \times 2$ supercell for visualization. The isovalue is $1.1 \times 10^{-5} \text{ \AA}^{-3}$. The density corresponding to the secondary sites is colored cyan to distinguish it from the primary sites (yellow). The black solid line outlines the unit cell. (b) Close-up view of the primary and secondary adsorption sites, illustrating the local “pockets” formed by neighboring Al, O, C, and N atoms. The primary adsorption site is surrounded by two symmetric N1–N2–O2, while the secondary site lies between O2–C–C and O2–N1–N2. (c) Isosurface plot of the H_2 distribution at 298 K. The yellow isosurface represents a lower density ($1.1 \times 10^{-6} \text{ \AA}^{-3}$), indicating delocalization, while the red isosurface represents a higher density ($2.5 \times 10^{-6} \text{ \AA}^{-3}$), showing residual localization.

(See also Fig. S10(a) in the Supplemental Material [32]). We observed that the diffusion of H_2 occurs along the pore direction (a axis) as shown in Fig. S10(b) in the Supplemental Material [32], which was also reflected in the MC simulations due to high energy barriers between the pores.

The calculated D 's are 3.50×10^{-9} , 3.15×10^{-8} , and $12.7 \times 10^{-8} \text{ m}^2 \text{ s}^{-1}$ at 77, 150, and 298 K, respectively. To the best of our knowledge, no experimental self-diffusion coefficients for H_2 transport in MOF-303 have been reported in the literature. Our calculated D at 77 K is lower than those experimentally (theoretically) measured (estimated) for MIL-47(V) and MIL-53(Cr) of $1\text{--}5 \times 10^{-8} \text{ m}^2 \text{ s}^{-1}$ at 77 K [54], but is comparable to the value estimated for Al-soc-MOF-1d from MLP-MD [30]. D 's at elevated temperatures also indicate that H_2 diffusion is significantly faster at room temperature

compared to cryogenic conditions within the framework. The diffusion behavior is consistent with the observed adsorption characteristics, where stronger binding at lower temperatures leads to reduced mobility. Finally, we calculated the diffusion coefficient as a function of inverse temperature and then extracted the slope to obtain E_a of approximately 2.8 kJ mol^{-1} (Fig. 5).

Our calculated E_a is higher than the values experimentally measured (theoretically estimated) for MIL-47(V) and MIL-53(Cr) of 0.60 (0.50–0.80) and 1.60 (1.05–1.70) kJ mol^{-1} , respectively [54], and comparable to that of 3.3 kJ mol^{-1} estimated for Al-soc-MOF-1d based on MLP-MD simulations [30]. The calculated E_a aligns with the physisorption nature of H_2 , where weak interactions with the framework allow for relatively facile diffusion. All GCMC and MD simulations in this work employ a rigid MOF-303 framework. Framework flexibility can in principle affect both adsorption uptake and diffusion: Flexible frameworks can alter pore geometry and accessible volume, thereby shifting adsorption isotherms and Q_{st} , and can also couple phonon modes to guest motion, modifying diffusion barriers and coefficients. For small-molecule physisorption such as H_2 in a relatively stiff Al-based MOF, however, these effects are expected to be modest. Extending the present MLP to a fully flexible framework, we leave it for future research.

IV. CONCLUSION

In this work, we propose a hybrid-sampling approach to overcome a key challenge in the application of MLPs to GCMC simulations of gas adsorption. In contrast to MD simulations at fixed composition, GCMC requires an accurate description of configurations spanning a continuously varying H_2 loading. The training set was therefore constructed by combining FF-based GCMC configurations, which cover this loading range, with AIMD trajectories at several

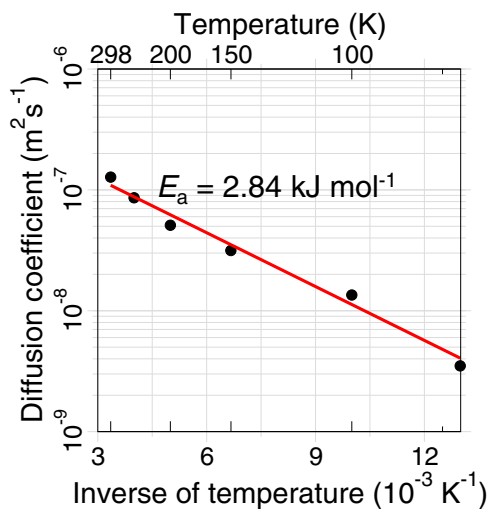


FIG. 5. *H_2 diffusion in MOF-303.* Diffusion coefficient as a function of inverse temperature, along with the activation energy (E_a) for diffusion estimated from the Arrhenius plot.

temperatures. Because the resulting dataset already spans the relevant loadings and energies, a few additional reference calculations at different loadings, rather than repeated retraining cycles, were sufficient to obtain an MLP that remains accurate and numerically stable throughout the GCMC sampling. The MLP was applied to H₂ adsorption in MOF-303. An accurate potential alone, however, was found to be insufficient for quantitative agreement with experiment at 77 K. Nuclear quantum effects were treated by a mode-resolved vibrational free-energy correction, evaluated on the fly for each adsorbed molecule and incorporated directly into the GCMC acceptance probability: Feynman-Hibbs corrections for the translational modes, quantum harmonic free energies for the hindered rotations, and the difference between the adsorbed and gas-phase free energies for the intramolecular stretch. With this treatment, the calculated 77 K isotherm reproduces both the pressure-dependent experimental uptake, and the isosteric heat (6.4 ± 0.1 kJ mol⁻¹) is in good agreement with the calorimetric value of 6.44 kJ mol⁻¹ at 0.001 MPa.

Basin-hopping search and MLP-MD simulations identify four primary and four secondary adsorption sites near the N, O, C, and -OH motifs lining the one-dimensional channels, and reveal a temperature-dependent transition from enthalpy-driven localization of H₂ at 77 K to entropy-driven delocalization across the pore surface at 298 K, along with a relatively low activation energy for diffusion, indicating facile transport of H₂ essential for practical cycling. The hybrid sampling approach proposed in this work is general and can be applied to other MOF systems and adsorbates. Although

the present MLPs are specific to MOF-303 and H₂, the overall workflow itself is readily transferable. We anticipate that combining computational screening with this approach will enable more reliable *in-silico* design of porous materials for gas adsorption.

ACKNOWLEDGMENTS

A part of the calculations was performed using the super-computer facilities of the Institute for Solid State Physics, the University of Tokyo, and Center for Computational Materials Science of Institute for Materials Research, Tohoku University (Project No. 202412-SCKXX-0202). K.S. thanks Professor Kazuto Akagi for helpful discussions. This work was supported by JSPS KAKENHI Grant No. 25K17822 and GteX Grant ID: JPMJGX23H1.

K.S., I.H., and S.O. conceptualized the project; K.S. performed research; K.S., I.H., T.I., Y.L., S.R., S.F., L.Z., H.T., T.K., H.L., and S.O. analyzed data; K.S. wrote the original draft; KS and IH edited the manuscript; S.O. and I.H. supervised the project; all the authors reviewed the manuscript.

All authors declare no financial or nonfinancial competing interests.

DATA AVAILABILITY

The data that support the findings of this article are openly available at [55].

- [1] H. Li, M. Eddaoudi, M. O’Keeffe, and O. M. Yaghi, Design and synthesis of an exceptionally stable and highly porous metal-organic framework, *Nature (London)* **402**, 276 (1999).
- [2] O. M. Yaghi, M. O’Keeffe, N. W. Ockwig, H. K. Chae, M. Eddaoudi, and J. Kim, Reticular synthesis and the design of new materials, *Nature (London)* **423**, 705 (2003).
- [3] H. Furukawa, K. E. Cordova, M. O’Keeffe, and O. M. Yaghi, The chemistry and applications of metal-organic frameworks, *Science* **341**, 1230444 (2013).
- [4] M. Eddaoudi, J. Kim, N. Rosi, D. Vodak, J. Wachter, M. O’Keeffe, and O. M. Yaghi, Systematic design of pore size and functionality in isoreticular MOFs and their application in methane storage, *Science* **295**, 469 (2002).
- [5] G. Maurin, C. Serre, A. Cooper, and G. Férey, The new age of MOFs and of their porous-related solids, *Chem. Soc. Rev.* **46**, 3104 (2017).
- [6] W. Temmerman, R. Goeminne, K. S. Rawat, and V. Van Speybroeck, Computational modeling of reticular materials: The past, the present, and the future, *Adv. Mater.* **37**, 2412005 (2025).
- [7] N. Hanikel, X. Pei, S. Chheda, H. Lyu, W. Jeong, J. Sauer, L. Gagliardi, and O. M. Yaghi, Evolution of water structures in metal-organic frameworks for improved atmospheric water harvesting, *Science* **374**, 454 (2021).
- [8] Z. Wang, F. Wang, Y. Luo, Y. Liu, Y. Wang, F. Wang, Y. Jin, and C. Xia, The Ag (I) coordination in Ag-MOF-303 enhanced the adsorption and separation capabilities for H₂/D₂, *Separat. Purific. Technol.* **360**, 131219 (2025).
- [9] H. Kim, S. Jee, J. Park, M. Jung, R. Muhammad, K. Choi, and H. Oh, High D₂/H₂ selectivity performance in MOF-303 under ambient pressure for potential industrial applications, *Sep. Purif. Technol.* **325**, 124660 (2023).
- [10] E. Knight, A. Gillespie, M. Prosniewski, D. Stalla, E. Dohnke, T. Rash, P. Pfeifer, and C. Wexler, Determination of the enthalpy of adsorption of hydrogen in activated carbon at room temperature, *Int. J. Hydrogen Energy* **45**, 15541 (2020).
- [11] L. J. Murray, M. Dinca, and J. R. Long, Hydrogen storage in metal-organic frameworks, *Chem. Soc. Rev.* **38**, 1294 (2009).
- [12] K. Cousins and R. Zhang, Highly porous organic polymers for hydrogen fuel storage, *Polymers* **11**, 690 (2019).
- [13] H. Daglar, H. C. Gulbalkan, G. O. Aksu, and S. Keskin, Computational simulations of metal-organic frameworks to enhance adsorption applications, *Adv. Mater.* **37**, 2405532 (2025).
- [14] S. L. Mayo, B. D. Olafson, and W. A. Goddard, DREIDING: A generic force field for molecular simulations, *J. Phys. Chem.* **94**, 8897 (1990).
- [15] L.-C. Lin, K. Lee, L. Gagliardi, J. B. Neaton, and B. Smit, Force-field development from electronic structure calculations with periodic boundary conditions: Applications to gas adsorption in metal-organic frameworks, *J. Chem. Theory Comput.* **10**, 1477 (2014).
- [16] R. Mercado, B. Vlaisavljevich, L.-C. Lin, K. Lee, Y. Lee, J. A. Mason, D. J. Xiao, M. I. Gonzalez, M. T. Kapelewski, J. B. Neaton, and B. Smit, Polarizable force fields for CO₂ and CH₄ adsorption in M-MOF-74 (M = Mg, Mn, Fe, Co, Ni, Zn), *J. Phys. Chem. C* **120**, 12590 (2016).

- [17] J. Borycz, L.-C. Lin, E. D. Bloch, J. Kim, A. L. Dzuba, R. Maurice, D. Semrouni, K. Lee, B. Smit, and L. Gagliardi, A theoretical study of hydrogen storage in metal-organic frameworks: H₂ binding in various types of pores, *J. Phys. Chem. C* **118**, 12230 (2014).
- [18] M. Fischer, J. R. Gomes, and M. Jorge, Computational approaches to study adsorption in MOFs with unsaturated metal sites, *Mol. Simul.* **40**, 537 (2014).
- [19] J. L. Belof, A. C. Stern, M. Eddaoudi, and B. Space, On the mechanism of hydrogen storage in a metal-organic framework material, *J. Am. Chem. Soc.* **129**, 15202 (2007).
- [20] D. M. Franz, Z. E. Dyott, K. A. Forrest, A. Hogan, T. Pham, and B. Space, Simulations of hydrogen, carbon dioxide, and small hydrocarbon sorption in a nitrogen-rich *rht*-metal-organic framework, *Phys. Chem. Chem. Phys.* **20**, 1761 (2018).
- [21] A. Samanta, T. Furuta, and J. Li, A theoretical study of hydrogen storage in metal-organic frameworks: H₂ binding in various types of pores, *J. Chem. Phys.* **125**, 084714 (2006).
- [22] M. Dixit, T. A. Maark, and S. Pal, *Ab initio* and periodic DFT investigation of hydrogen storage on light metal-decorated MOF-5, *Int. J. Hydrogen Energy* **36**, 10816 (2011).
- [23] M. Fischer, DFT-based evaluation of porous metal formates for the storage and separation of small molecules, *Microporous Mesoporous Mater.* **219**, 249 (2016).
- [24] J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, Recent advances and applications of machine learning in solid-state materials science, *npj Comput. Mater.* **5**, 83 (2019).
- [25] O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K. R. Müller, Machine learning force fields, *Chem. Rev.* **121**, 10142 (2021).
- [26] J. Behler, Four generations of high-dimensional neural network potentials, *Chem. Rev.* **121**, 10037 (2021).
- [27] M. Eckhoff and J. Behler, From molecular fragments to the bulk: Development of a neural network potential for MOF-5, *J. Chem. Theory Comput.* **15**, 3793 (2019).
- [28] S. K. Achar, J. J. Wardzala, L. Bernasconi, L. Zhang, and J. K. Johnson, Combined deep learning and classical potential approach for modeling diffusion in UiO-66, *J. Chem. Theory Comput.* **18**, 3593 (2022).
- [29] R. Goeminne, L. Vanduyfhuys, V. Van Speybroeck, and T. Verstraelen, DFT-quality adsorption simulations in metal-organic frameworks enabled by machine learning potentials, *J. Chem. Theory Comput.* **19**, 6313 (2023).
- [30] S. Liu, R. Dupuis, D. Fan, S. Benzaria, M. Bonneau, P. Bhatt, M. Eddaoudi, and G. Maurin, Machine learning potential for modelling H₂ adsorption/diffusion in MOFs with open metal sites, *Chem. Sci.* **15**, 5294 (2024).
- [31] J. Klarbring and A. Walsh, Na vacancy-driven phase transformation and fast ion conduction in W-doped Na₃SbS₄ from machine learning force fields, *Chem. Mater.* **36**, 9406 (2024).
- [32] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/d5b5-66b6> for containing the DREIDING force-field parameters, the training-data distributions and parity plots of MLP against AIMD, and additional uptake and isosteric-heat results. It further contains the adsorption-site searches, the DFT-D3 and MLP vibrational frequencies, structural and dynamical analyses, and Appendix A on modifications to the GCMC algorithm.
- [33] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, *Comput. Phys. Commun.* **271**, 108171 (2022).
- [34] P. E. Blöchl, Projector augmented-wave method, *Phys. Rev. B* **50**, 17953 (1994).
- [35] G. Kresse and J. Hafner, *Ab initio* molecular dynamics for liquid metals, *Phys. Rev. B* **47**, 558 (1993).
- [36] G. Kresse and J. Furthmüller, Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set, *Comput. Mater. Sci.* **6**, 15 (1996).
- [37] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [38] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *J. Chem. Phys.* **132**, 154104 (2010).
- [39] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky, (E3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials, *Nat. Commun.* **13**, 2453 (2022).
- [40] Y. Park, J. Kim, S. Hwang, and S. Han, Scalable parallel algorithm for graph neural network interatomic potentials in molecular dynamics simulations, *J. Chem. Theory Comput.* **20**, 4857 (2024).
- [41] A. Hjorth Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Duřak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, *et al.*, The atomic simulation environment—A Python library for working with atoms, *J. Phys.: Condens. Matter* **29**, 273002 (2017).
- [42] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
- [43] R. P. Feynman, *Statistical Mechanics: A Set of Lectures* (W. A. Benjamin, Reading, Massachusetts, 1972), Sec. 3.5.
- [44] J. I. Siepmann and D. Frenkel, Configurational bias Monte Carlo: A new sampling scheme for flexible chains, *Mol. Phys.* **75**, 59 (1992).
- [45] S. E. Bambalaza, H. W. Langmi, R. Mokaya, N. M. Musyoka, J. Ren, and L. E. Khotseng, Compaction of a zirconium metal-organic framework (UiO-66) for high density hydrogen storage applications, *J. Mater. Chem. A* **6**, 23569 (2018).
- [46] L. Sarkisov, R. Bueno-Perez, M. Sutharson, and D. Fairen-Jimenez, Materials informatics with PoreBlazer v4.0 and the CSD MOF database, *Chem. Mater.* **32**, 9849 (2020).
- [47] B. Widom, Some topics in the theory of fluids, *J. Chem. Phys.* **39**, 2808 (1963).
- [48] D. R. Galimberti and J. Sauer, Chemically accurate vibrational free energies of adsorption from density functional theory molecular dynamics: Alkanes in zeolites, *J. Chem. Theory Comput.* **17**, 5849 (2021).
- [49] M. E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation* (Oxford University Press, Oxford, UK, 2023).
- [50] D. J. Wales and J. P. K. Doye, Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters, *J. Phys. Chem. A* **101**, 5111 (1997).
- [51] D. J. Wales, *Energy Landscapes: Applications to Clusters, Biomolecules and Glasses* (Cambridge University Press, Cambridge, UK, 2003).

- [52] A. E. Abu EL-Maaty, M. A. Abdalla, M. Essalhi, M. M. Abdelnaby, M. M. Mahmoud, M. A. Habib, M. Antar, and R. Ben-Mansour, Kinetics of water adsorption in metal-organic framework (MOF-303) for adsorption cooling application, *Energy Convers. Manage.: X* **24**, 100694 (2024).
- [53] O. M. Yaghi and J. Yang, Porous aluminum pyrazoledicarboxylate frameworks (2021), US Patent 11,014,068, https://patentscope2.wipo.int/search/en/detail.jsf?docId=WO2019010102&_cid=JP1-MTMINR-64949-1.
- [54] F. Salles, D. I. Kolokolov, H. Jobic, G. Maurin, P. L. Llewellyn, T. Devic, C. Serre, and G. Ferey, Adsorption and diffusion of H₂ in the MOF type systems MIL-47(V) and MIL-53(Cr): A combination of microcalorimetry and QENS experiments with molecular simulations, *J. Phys. Chem. C* **113**, 7802 (2009).
- [55] K. Sau, Hybrid sampling approach to machine-learning potentials for gas adsorption: Hydrogen adsorption in MOF-303 [Dataset], Zenodo, 2026, <https://doi.org/10.5281/zenodo.22030358>.