High throughput sampling of phase space with deep learning potentials: δ-AlOOH at geophysical conditions

Chenxing Luo ⬥,[1] Yang Sun ⬥,[1, 2, *] and Renata M. Wentzcovitch ⬥[1, 3, 4, 5, †]

[1]*Department of Applied Physics and Applied Mathematics,*
Columbia University, New York, New York 10027, USA
[2]*Department of Physics and Astronomy, Iowa State University, Ames, Iowa 50011, USA*
[3]*Department of Earth and Environmental Sciences, Columbia University, New York, New York 10027, USA*
[4]*Lamont–Doherty Earth Observatory, Columbia University, Palisades, New York 10964, USA*
[5]*Data Science Institute, Columbia University, New York, New York 10027, USA*
(Dated: September 14, 2023)

Hydrous and nominally anhydrous minerals (NAMs) are a fundamental class of solids of enormous significance to geophysics. They are the water carriers in the deep geological water cycle. They impact structural, elastic, plastic, and thermodynamic properties and phase relations in Earth's forming aggregates (rocks). They play a critical role in the geochemical and geophysical processes that shape the planet. Their complexity has prevented predictive calculations of their properties, but progress in materials simulations ushered by machine learning potentials is transforming this state of affairs. Here, we adopt a hybrid approach that combines deep learning potentials (DP) with the SCAN meta-GGA functional to simulate a prototypical hydrous system. We illustrate the viability, success, and necessity of this approach to simulate δ-AlOOH (δ), a phase capable of transporting water down to near the core-mantle boundary of the Earth (~2,900 km depth and ~135 GPa) in subducting slabs. High-throughput sampling of phase space using molecular dynamics simulations with DP-potentials offers a panoramic view of the hydrogen-bond behavior and proton diffusion at geophysical conditions. These simulations provide a pathway for a deeper understanding of these crucial components that shape Earth's internal state.

I. INTRODUCTION

H-bearing mineral phases are responsible for water circulation from the Earth's surface to the interior via subduction and convection. Up to ten oceans could be stored in the Earth's interior as hydrous phases or nominally anhydrous minerals (NAMs), where H exists as substitutional or interstitial defects [1, 2]. Hydrogen bonds (H-bonds) transform under elevated pressures and temperatures in these solids. Like H-bonds in H_2O-ice, they symmetrize under pressure, producing ionic bonds [3, 4] or disorder before melting [5, 6]. Superionic diffusion of protons is also expected as a precursor to dehydration reactions responsible for melt production and volcanism. In NAMs, the H-concentration changes phase relations and plastic properties, causing seismological properties irregularities in the mantle (e.g., [7–10]). Fragile H-bonds and hydrous defects weaken the rock's rheological properties, facilitating plastic deformation and thermal convection in the mantle, a central process in Earth's evolution (e.g., [11–13]). Therefore, a detailed understanding of H-bond behavior and related mineral properties is essential for understanding the Earth's interior's geochemical activity, dynamic behavior, and seismological properties.

Because the pressures and temperatures (P, Ts) in the Earth's interior challenge the experimental determination of materials' properties, insights from *ab initio* studies have been indispensable. Standard LDA [14] and PBE [15] functionals combined with phonon calculations and the quasiharmonic approximation (QHA) [16–18] or with *ab initio* molecular dynamics (MD)

[19] have been used to address the physical properties of such phases. However, anharmonicity, small MD simulation size, inadequate DFT functionals, quantum nuclear effects, and other factors prevented the reliable description of H-bond behavior and property changes in these phases at relevant conditions.

Here, we focus on δ-AlOOH (δ) [20], a prototypical hydrous phase stable throughout the entire pressure range of the Earth's mantle (up to 135 GPa) [21]. It has been extensively studied experimentally (e.g., [21–23]) and computationally (e.g., [24–27]). Therefore, experimental uncertainties and *ab initio* calculations' limitations to reproducing observations are well established. This phase has a simple high-pressure $CaCl_2$-type structure [20, 28] and exhibits anomalous elastic properties under pressure [26], with proton diffusion expected at high temperatures [29]. δ, its isostructural siblings ε-FeOOH, $MgSiO_4H_2$ phase H [30], the NAM $CaCl_2$-type SiO_2, along with their multiple solid solutions, are the main H-bearing phases in the deep mantle [31]. The structural simplicity, the abundance of careful measurements and calculations of δ's properties, and its prototypical nature make it a most suitable model phase for testing the performance of *ab initio*-based machine learning (ML) methods in addressing these systems at high P, Ts.

According to measurements, H-bonds in δ symmetrize and form H-centered (HC) bonds at ~18 GPa [22, 23]. Quasiharmonic approximation (QHA) based methods (e.g., [16–18]), while widely used to describe finite-temperature properties of solids, rely on stable phonon modes. They cannot describe δ's properties in the pressure range of H-bond symmetrization, a transition accompanied by strong anharmonicity and an order-disorder phenomenon below ~40 GPa in static calculations [25, 27]. Besides, previous *ab initio* QHA studies have used the LDA and PBE/GGA functionals, making comparing *ab initio* predictions and experimental measurements at the same

---

pressures challenging. The relatively low experimental H-bond symmetrization pressure of ~8–18 GPa has not been correctly reproduced in these DFT calculations because of the process's complexity and these functional's inadequate description of H-bonds, giving a transition pressure in the 30–40 GPa pressure range [23, 27].

Superionic proton diffusion before the dehydration reaction (2 AlOOH → H_2O + Al_2O_3), is also anticipated at high temperatures. This phenomenon has also been found in other H-bearing systems like antigorite [32] or diopside with H defects [33] and has been extensively investigated in ice before melting (e.g., [34–39]). A recent Born-Oppenheimer molecular dynamics (BOMD) study using the PBE functional reported proton diffusion in δ at 2,700–3,000 K [29], a temperature range higher than the 1,600–2,500 K dissociation temperature measured in the 20–140 GPa range [21, 28, 40]. The PBE functional and the small MD simulation cell sizes with only a few hundred atoms are likely responsible for this discrepancy. In summary, accurately reproducing P, T stability fields of hydrous phases and their properties is fundamental for mapping out phase relations and dehydration sequences along a specific geotherm (e.g., [11]). These results provide the basis for interpreting irregular seismic properties caused by the non-homogeneous distribution of these H-bearing solids and melts in the mantle and estimating the water content in Earth's interior (e.g., [1, 2, 12]).

This challenge can be overcome by adopting ML-based potentials (e.g., [41–46]) in MD simulations with DFT predictive power but significantly reduced computational cost [47]. This approach only invokes DFT while training the potential to reproduce interatomic forces and energies using ML-based descriptors. The ML-based MD simulations are then performed at a cost and scaling close to empirical force-field simulations. Active learning schemes (e.g., [48, 49]) help create compact reference datasets by iteratively discovering and actively adding "missing" necessary atomic configurations, reducing the DFT computation cost in the potential training process. In particular, Deep Potential (DP) is a neural-network (NN) type ML potential [41] with a neural network's flexible fitting capability, allowing NN-potentials to represent chemical systems of varied nature. Recent benchmarks have shown that DP forces and energies in solids and liquids structures trained on a few thousand reference configurations can be highly accurate [47, 50, 51]. The adoption of ML-based potentials allows the use of the strongly constrained and appropriately normed (SCAN) meta-generalized gradient approximation (meta-GGA) functional's more accurate description of the H-bonded systems [52] regardless of being computationally more expensive. This method successfully calculated the water/ice phase diagram [53] and proton diffusion in liquid water [54].

Combining deep learning potentials with the SCAN functional is a promising path to simulate H-bearing systems at extreme geophysical conditions accurately. Here, we apply this hybrid approach to δ, a prototypical H-bearing system, to understand how well it overcomes the limitations presented in conventional purely DFT studies. This method enables us to perform long and large simulations that densely cover a wide (P, T) or (V, T) range with changing H-bond or proton diffusion

behavior. We benchmark the potential against SCAN calculations and experimental measurements to assess the accuracy level achieved. We address δ's compression curve in its high-P H-bond symmetrization regime and its high-T proton diffusion regime, which were not accurately described in previous QHA- or BOMD-based studies.
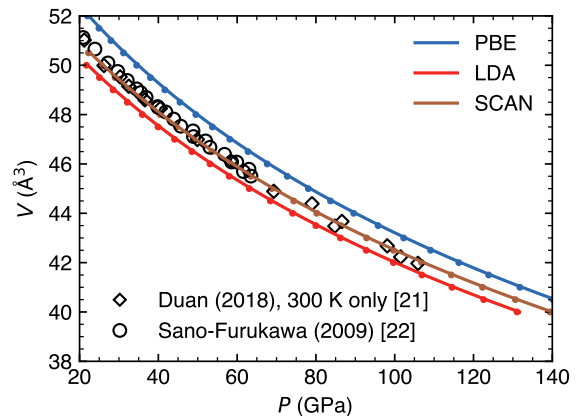
II. RESULT

A. NN-potential benchmark

FIG. 1. LDA, PBE, and SCAN static compression curves compared to 300 K measurements [21, 22].

We employ the SCAN functional [52] to describe δ's potential energy (Born-Oppenheimer) surface and interatomic forces. Compared to LDA's underestimation and PBE's overestimation of the pressure, SCAN's static compression curve is the closest to the 300 K measured one (see Fig. 1). The slight underestimation of the SCAN volume is eliminated when thermal effects at 300 K are included, as shown in the subsequent discussion. Quantum nuclear effects disregarded in the present study might spoil this good agreement [55], and this effect must be further inspected. We are interested in the temperature range of typical subducting tectonic plates (slabs), i.e., along "slab geotherms" [56, 57], where this phase is stable in the mantle, i.e., ~ 1,000 K < T < 3,000 K [21, 58]. At these temperatures, classical MD should adequately describe the ionic dynamics in δ.

Our SCAN-DP interatomic potential trained on just a few thousand reference configurations can reproduce SCAN-DFT's forces and energies for configurations sampled from SCAN-BOMD simulations with 128 atoms, as illustrated in Fig. 2. The root-mean-square error (RMSE) [59] of the SCAN-DP predictions for these tests at various (V, T) states are listed in Table SI. Configurations generated at higher temperatures produce larger RMSEs in general. At 3,000 K, the RMSEs for potential energy and force predictions are ~2 meV/atom, and ~0.12 eV/Å, respectively. Such accuracy is similar to previous benchmarks (e.g., [47]) in similar DP studies.
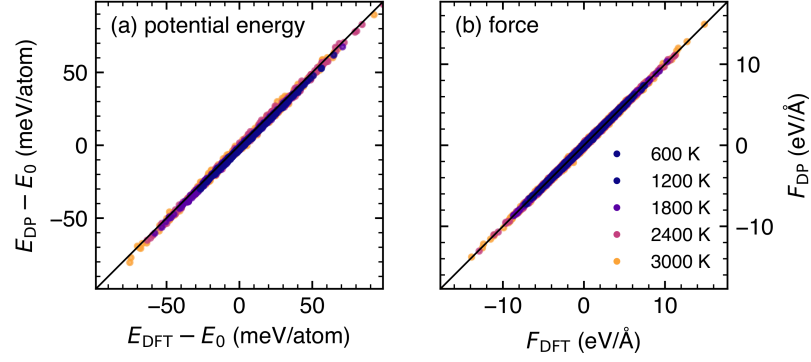
FIG. 2. Comparison between SCAN-DP and SCAN-DFT predictions of (a) 1,600 potential energies and (b) 10,240 atomic forces randomly sampled from 30,128 configurations derived from 20 SCAN-BOMD NVT trajectories at 5 temperatures (T = 600, 1,200, 1,800, and 3,000 K) and 4 volumes (V = 40, 44, 48, 54 Å^3). Symbol colors denote temperature.
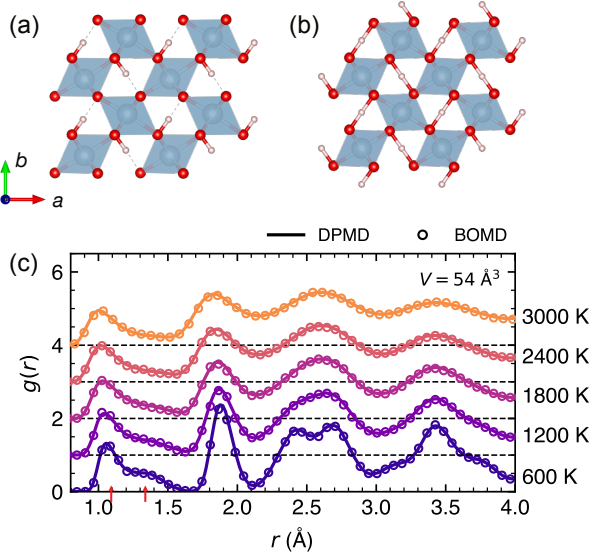

FIG. 3. (a, b) The crystal structures of δ-AlOOH with (a) asymmetric and (b) symmetric H-bonds. (c) Comparison between SCAN-BOMD and SCAN-DPMD pair-distribution function, $g(r)$, at various (T, V)'s. The conventional-cell volume 54 Å^3 corresponds to 9.42, 13.67, 17.90, 22.73, and 28.34 GPa at 600–3,000 K. Red arrows show the H-bond and OH ionic bond lengths in the first $g(r)$ peak.

SCAN-DPMD also reproduces SCAN-BOMD's structure and bonding properties at high-P, T, as illustrated by the comparison of SCAN-BOMD's and SCAN-DPMD's radial distribution functions, $g(r)$, in Fig. 3. The results are for 54 Å^3 volume, corresponding to 9.5–28.3 GPa at 600–3,000 K obtained in 128-atom simulations. Plots detailing the contribution from each type of atomic pair at other P, T's are shown in Fig. S1.

The $g(r)$ at V = 54 Å^3 shown in Fig. 3 is particularly significant because this volume approximately corresponds to

the critical volume of the experimentally observed H-bond disorder transition at 300 K [22]. At 600 K, the first peak at ~1.1 Å corresponds to the ionic OH bond length. This peak is asymmetric at this volume and at all sampled temperatures due to asymmetric proton motion. At 600 K, $g(r)$ shows a near double-peak distribution of OH ionic bonds and H-bonds; the broad shoulder centered at ~1.4 Å corresponds to the distribution of H-bond lengths. The increased overlap of the two peaks with increasing temperature suggests the onset of a disordered state, a precursor to H-bond symmetrization [23]. A similar OH bond-length distribution was observed in 300 K classical BOMD simulations with a quantum thermostat [60]. Some features on $g(r)$ disappear at elevated temperatures. For example, the 2.4 Å peak, a combination of Al-O and O-H contributions, and the 2.7 Å peak associated with O-O bonds (see Fig. S1) are split at 600 K but merge above 1,200 K. This phenomenon results from increased atomic vibration amplitudes, proton dynamic disorder, and diffusion onset at elevated temperatures. We will elaborate further on the diffusion process in Sec. II C. Under pressure, the first peak at ~1.1 Å and the shoulder at 1.4 Å evolve into a single symmetric peak due to H-bond disorder and finally symmetrization [23, 27] (see Fig. S1).

SCAN-DPMD with *ab initio*-level accuracy and improved efficiency will help predict the thermoelastic properties of hydrous solids, a property of first-order importance in geophysics [18, 51, 61]. Here, we compare its predictions for δ's high-temperature compressive behavior with *ab initio* SCAN-BOMD predictions and measurements in the 300–3,000 K and 20–115 GPa range. Fig. 4(a) shows excellent agreement between SCAN-DPMD's high-temperature compression curves fitted to third-order Birch-Murnaghan equations of state (EoS) and the SCAN-BOMD's predicted P, V data points. Comparison with measurements [21, 22] in Fig. 4(b) shows a promising 3 GPa maximum error in pressure prediction at the same V, Ts. The next section will present a more detailed analysis of the

300 K compression curve at lower pressure and compare it with measurements.

These good agreements between the SCAN-DFT and SCAN-DP predictions of forces, potential energies, $g(r)$, and high-P,T EoSs suggest that the deep-learning potentials have reached satisfactory accuracy and are predictive at least in the 20–120 GPa pressure range and up to 3,000 K.

B. The 300 K compression curve and dynamic stabilization of the HC phase

The 300 K compression curve is one of δ's best-determined properties. Experimentally, δ exhibits an anomalous change in compressive behavior at ~10 GPa, which is usually attributed to a change in H-bond structure, most often with H-bond symmetrization [24, 64–66], but also with H-bond disorder [22, 23, 67]. However, it has been impossible to reproduce this compressive behavior using QHA-based DFT calculations. This is because (a) static *ab initio* calculations without multiconfiguration analysis do not account for H-bond disorder; (b) the HC phase atomic configuration is dynamically stable above ~35 GPa only in static GGA/PBE calculations [25]. Strong anharmonicity produces unstable phonon modes at lower pressures, hindering the application of the QHA for high-temperature calculations [25, 27]. Using SCAN-DPMD, we optimize and equilibrate the cell shape at several pressures to obtain the static and the 300 K compression curves shown in Figs. 5(a,b). With these, we can analyze the effect of atomic motion on the H-bond state, its effect on the volume, and compare the latter with measurements [21, 22, 62]. As shown in Fig. 5(b), extrapolations of a Birch-Murnaghan EoS fit to results below and above ~10 GPa produce divergent curves (dotted blue and red dotted lines), indicating a change in the OH-bond state at ~10 GPa.

The 300 K isothermal bulk modulus, $K_T = V(\partial P/\partial V)_T$, shown in Fig. 5(c), amplifies the subtle change in compressibility across this critical point. Across this state change in the 0–20 GPa range, the OH bond-length distribution, $g_{OH}(r)$, evolves from a double peak to a single modal distribution, though not symmetric (Fig. 5(d)). In this pressure range, the H-bond compresses quickly while the ionic OH bond stretches, a well-known behavior of these bonds (e.g., [55, 68]). δ's compressive behavior obtained from SCAN-DP calculation confirms previous PBE/GGA results [26] that this change in compressive behavior occurs in the 30–40 GPa range in static calculations (dashed line in Fig. 5(c)).

SCAN-DPMD results at 300 K account for the thermal expansion and reproduce well the 300 K experimental compression curve [21, 22, 62]: in the 0–15 GPa range, the difference between results and measurements is less than the differences between measurements; at higher pressures, the predicted SCAN-DPMD volume is slightly overestimated, with an error smaller than 0.25 Å³/f.u. Although predicted [24], it is striking to see that the inclusion of *classical* ionic motion lowers the transition pressure by ~27 GPa, giving a transition pressure that agrees quite well with the experimentally measured one at ~9 GPa (e.g., [22, 66]). This 300 K result is surprising and

puzzling. At this low temperature, quantum proton motion, particularly tunneling [69], should impact the H-bond behavior (e.g., [70]).

Above 10 GPa, V and K_T display a smooth monotonic dependence on pressure. Table I summarizes the Birch-Murnaghan EoS parameters resulting from the fitting of measured and calculated 300 K P, V data above 10 GPa. Fitting the SCAN-DPMD results within the 10–64 GPa pressure range with $K_0' = 4$ gives $V_0 = 55.5$ Å³, $K_0 = 225 \pm 2$ GPa (red dotted line in Fig. 5(b)). The predicted V_0 is in excellent agreement with the measured V_0 [22] fit to the same Birch-Murnaghan EoS, 55.47 Å³. The predicted bulk modulus differs by ~3% from the experimental one, 219 GPa [22], an impressive agreement for a result extrapolated to 0 GPa.

Our SCAN-DPMD results are significantly more accurate than previous PBE-BOMD results, which reported a $V_0 = 58.5$ Å³ and $K_0 = 183.4$ GPa using the Vinet EoS fit to results above 10 GPa [60]. This improvement can be attributed to (a) SCAN's more accurate description of the P-V relation, (b) the denser sampling of (P, V) states over a broader pressure range, and (c) larger, longer, and better-converged simulation runs enabled by DPMD's performance leap compared to BOMD. Nevertheless, this level of agreement between *classical* MD results and measurements at 300 K is unexpected.

This good agreement between SCAN-DPMD results and experimental data above 10 GPa suggests that our simulations describe well the H-bond state above this pressure. As indicated in Fig. 5(d), not even at 20 GPa, the first two peaks in the $g_{OH}(r)$ pair distribution function have merged, indicating that H-bond symmetrization has not yet been achieved in the simulations. Therefore, the change in compressive behavior at ~10 GPa cannot be attributed to H-bond symmetrization in this classical simulation. Fig. 6 shows in detail the evolution of $g_{OH}(r)$ and $g_{HH}(r)$ pair distribution functions from 0 to 10 GPa. Both show two clear peaks at 0 GPa and tend to merge with increasing pressure. However, only the $g_{HH}(r)$ peaks are fully merged at 10 GPa. $g_{OH}(r)$ still displays two superposing broad peaks. This indicates that H-bonds are not symmetric in these classical simulations at 10 GPa (or 20 GPa). The change in the H-bonds' state at 10 GPa seems related to a change in H-ordering, likely into a disordered state. Our previous multi-configuration PBE-QHA study [27] suggested several short- to medium-range ordered H-bonded configurations remain dynamically stable beyond the pressure where the compressive behavior changes (~10 GPa in experiments, ~12 GPa in PBE-QHA calculations). This result supported the notion of a more disordered H-bond state. The current results showing an asymmetric first peak in $g_{OH}(r)$ and the fully merged broad peak in $g_{HH}(r)$, and the previously identified multi-configuration equilibrium state [27] suggest that the change in compressive behavior at ~10 GPa is associated with the emergence of a disordered state in the simulations. Static or dynamic, disorder involves proton hopping and is a precursor to proton diffusion at higher temperatures. This 300 K behavior might change significantly if the proton dynamics is addressed quantum mechanically using path integrals [70].
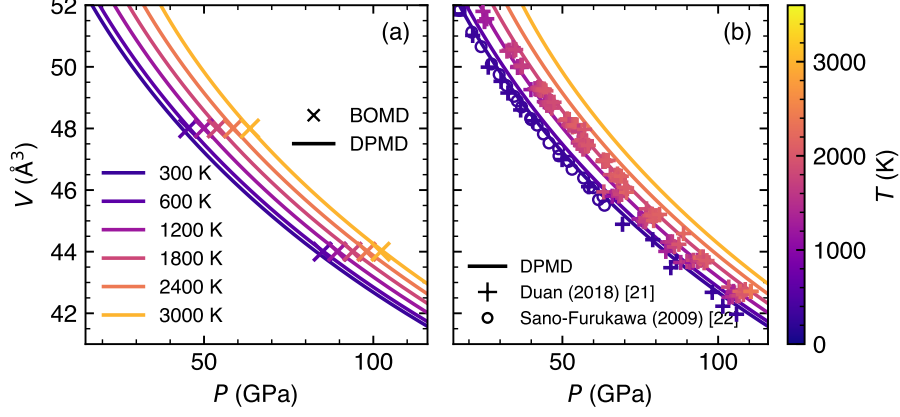
FIG. 4. δ's SCAN-DPMD compression curves at various temperatures compared to (a) SCAN-BOMD's PVT data and (b) measurements [21, 22]. Volumes correspond to the conventional unit cell containing two f.u. of δ-AlOOH.
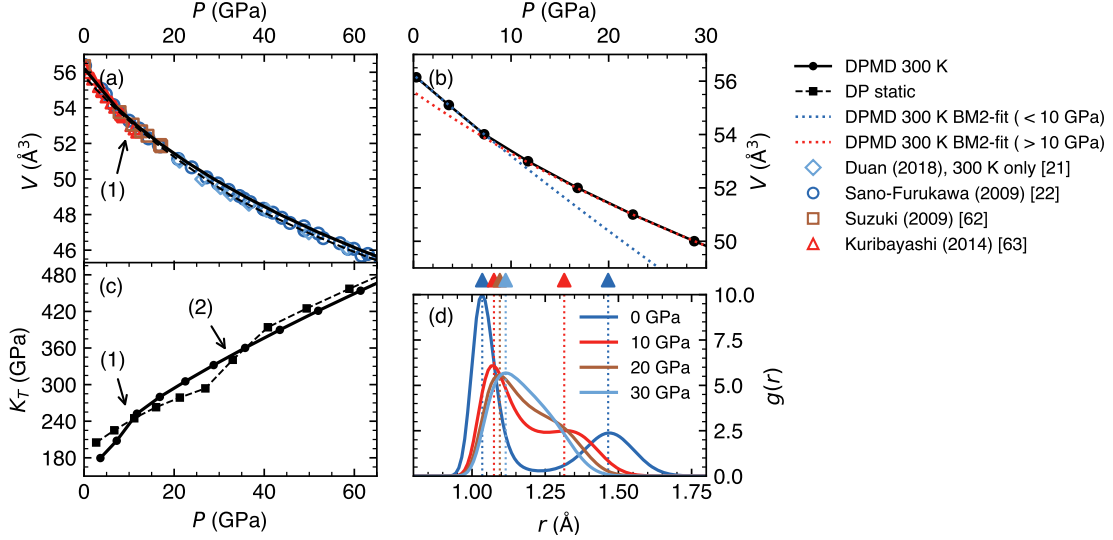

FIG. 5. (a) 300 K and static compression curve results are compared to 300 K measurements [21, 22, 62, 63]. (b) Static and 300 K SCAN-DPMD results and 300 K Birch-Murnaghan EoS fitted for 10–64 GPa data. (c) 300 K and static bulk modulus ($K = -V\, \partial P/\partial V$). Arrows (1) and (2) highlight the change in compressibility at (1) 300 K at ~10 GPa and (2) static conditions at ~35 GPa. (d) The evolution of r(OH) bond length distributions, $g_{OH}(r)$, at 0 (blue), 10 (red), and 20 GPa (brown). Color dashed lines and filled triangles indicate the r of the peaks.

C. Proton diffusion at high temperatures

The DP potential allows us to systematically perform large-scale DPMD simulations to investigate proton diffusion in δ over a wide P, T range. We perform 96 NVT DPMD simulations covering the pressure range of 35–140 GPa and temperature range of 1,500–3,000 K (see Fig. S2). They contain 8,192 atoms and run for 2 ns. We do not observe diffusion, structural change, or melting of the Al and O sub-lattices in this P, T range.

Proton diffusion is quantitatively characterized by protons' mean-square displacement (MSD). Protons' MSD over a simu-

lation run time t, $L_H^2(t)$, is defined by [71, 72]

$$L_H^2(t) = \frac{1}{N_H}\left\langle \sum_{i=1}^{N_H} \left| R_{H,i}(t+\tau) - R_{H,i}(\tau) \right|^2 \right\rangle, \quad (1)$$

where N_H is the total number of protons, $R_{H,i}(\tau)$ denotes the position of i-th proton at moment τ, and "$\langle \cdot \rangle$" denotes the ensemble average over the start time τ. A linear (non-linear) dependence of the MSD on the simulation run time, t, reflects a continuous (irregular) diffusive behavior. The self-diffusion coefficient, D_H, or the diffusivity, is related to the slope of the MSD (L^2) vs. t line,

$$D_H = \lim_{t \to \infty} \frac{L_H^2}{6t}. \quad (2)$$

TABLE I. Comparision of 300 K EoS parameters for HC-δ.

	V_0 (Å^3)	K_0 (GPa)	K_0'	Notes
Sano-Furukawa *et al.* [22]	55.47	219	4 (fixed)	10–63.5 GPa
Duan *et al.* [21]	55.3	223	4 (fixed)	Mie-Grüneisen, up to 142 GPa
Mashino *et al.* [66]	56.374	190	3.7	
Simonova *et al.* [67]	55.56	216	4 (fixed)	
Bronstein *et al.* [60]	58.5	183.4		PBE-BOMD
Kang *et al.* [64]	57.6	196	4.0	PBE-BOMD, 20–35 GPa
This study	55.5	225 ± 2	4 (fixed)	SCAN-DPMD, 10–64 GPa
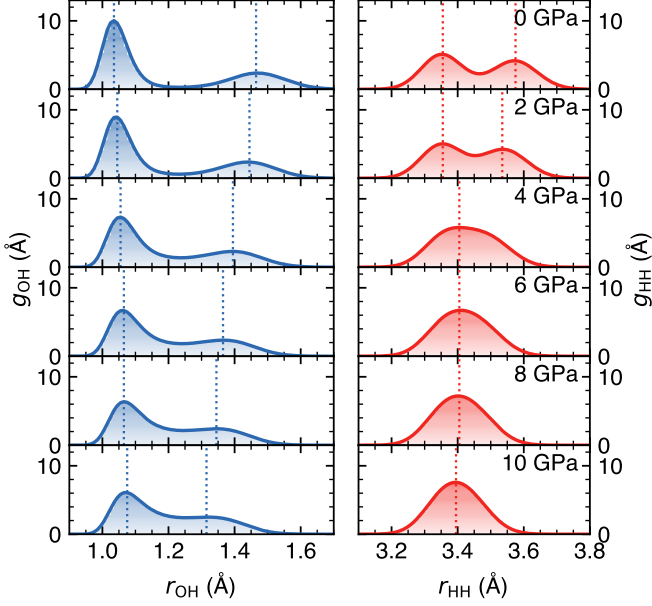
FIG. 6. Pair correlation function $g_{OH}(r)$ and $g_{HH}(r)$ in 0–10 GPa at 300 K. Color dashed lines indicate peak extremes.

FIG. 7. Proton diffusion at high temperatures for $V = 47$ Å^3 conventional cell which corresponds to 60.9 and 67.5 GPa at 1,500 K and 2,400 K. (a) Proton mean-square displacement (MSD) at 1,500 K and (b) 2,400 K. (c) Proton trajectories at 1,500 K and (d) 2,400 K; blue, red, and gray dots denote Al, O, and H ions.

Figs. 7(a,b) show the protons' MSD over a time span of 1 ns for $V = 47$ Å^3 and $T = 1,500$ K (60.9 GPa) and 2,400 K (67.5 GPa), respectively. MSD at several other V, T conditions are available in Fig. S3. Increasing the temperature significantly increases protons' diffusivity (see Fig. S3). Diffusion is not observed at 1,500 K. It starts between 1,800 K and 2,100 K and becomes steady at 2,400 K. Figs. 7(c,d) show the corresponding proton trajectories. Fig. 7(b), shows that the protons' MSD along the c direction is approximately twice as large as those along the a and b axes, indicating they move more freely through the interstitial channels along the c direction (see Figs. 3(a,b)). A similar phenomenon has also been reported in the NAM phase hydrous Al-bearing stishovite, which has a similar CaCl$_2$-type Si-O framework [73].

At 1,800–3,000 K, the relationship between log D_H and $1/T$ is linear at all volumes and can be fitted using the Arrhenius equation (see Fig. S4),

$$D_H(T) = D_0 \exp\left(\frac{-E_a}{k_B T}\right), \qquad (3)$$

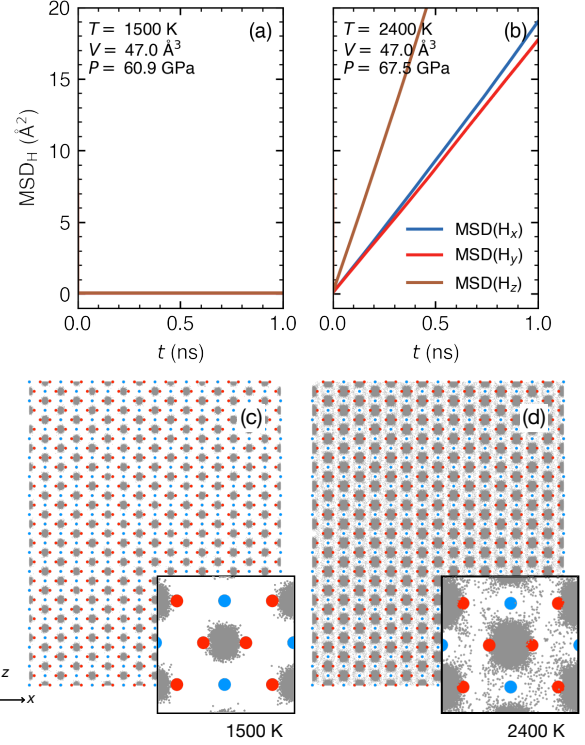where k_B is the Boltzmann constant and E_a denotes the activa-

tion energy. This behavior indicates that proton diffusion in δ is a common thermally activated process. The volume dependence of the activation energy, E_a, is nearly linear (Fig. S5). Using these relationships, we interpolate D_H vs. V and T using the finite temperature EoS displayed in Fig. 4. The magnitude of protons' diffusivity is shown as the background color in Fig. 8. Depending on the pressure, for $1,800$ K $< T < 2,100$ K, D_H approaches 1 Å^2/ns and starts deviating from the high-T Arrhenius behavior (see Fig. S4). This region of the diffusivity diagram is transiting from a blueish to a whitish background. At these P, T conditions, one can recognize in Fig. S3 a change in the MSD's behavior from non-linear to linear. We designate this regime as the onset of the fully superionic behavior in the P, T phase space and represent this $D_H = 1$ Å^2/ns diffusion
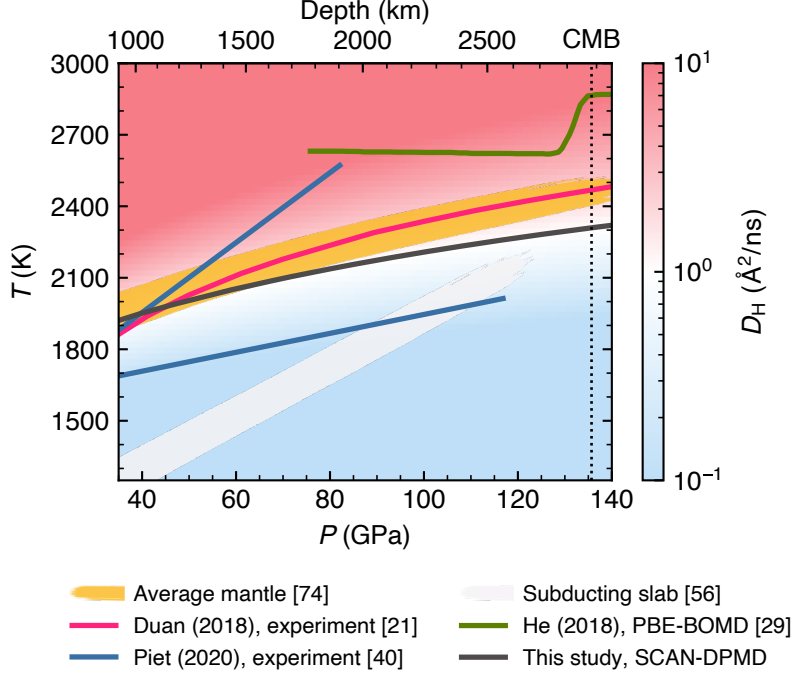
FIG. 8. The diffusion phase diagram of δ. The background color represents the diffusivity D_H. The solid black curve indicates the $D_H = 1$ Å²/ns boundary. The diffusion boundary from Ref. [29] and dissociation boundaries from Refs. [21, 40] are also shown. Adiabatic mantle geotherm [74] and slab geotherm [56].

boundary in Fig. 8 with a solid black curve.

The experimental dehydration boundary reported in Ref. [21] resembles our diffusion boundary. Similar measurements in Ref. [40] pinpointed two δ dehydration temperatures in two distinct pressure ranges: 1,850–1,900 K at 28–68 GPa and ~1,959 K at 100–110 GPa. Two divergent boundaries were reported in the mid-lower mantle (68–100 GPa) due to uncertainties related to unidentified X-ray diffraction peaks and uncertain experimental conditions (Fig. 8). Their lower boundary directly connects the low-pressure and high-pressure boundaries and runs parallel to our boundary with a 200 K downward shift. It has been argued [40] that the disagreement between these studies could be due to uncertainties in granularity or the preferred orientation of samples in the high-pressure cell, which could affect the dehydration product detection sensitivity. The proximity of our diffusion boundary and the experimental dehydration boundary, particularly their similar pressure gradients, suggests that continuous proton diffusion controls the dehydration process as in antigorite [32] or diopside with H defects [33].

The diffusion boundary previously obtained with PBE-BOMD [29] largely overestimates the superionic transition temperature. It is known that supercell size could affect the superionic diffusivity calculations, often leading to overestimation of the transition temperature [75, 76]. The short simulation time and the PBE's inaccurate description of the H-bond could also contribute to the overestimation of the diffusion boundary.

It is also possible to investigate the change in protons' diffusivity by inspecting the constant-volume heat capacity, C_V, through the P, T range shown in Fig. S6. The constant-volume heat capacity, C_V, can be calculated from MD ensemble aver-

ages [77] as

$$NC_V = \frac{\text{var}(E)}{k_B T^2} = \frac{1}{k_B T^2}\left[\langle E^2 \rangle - \langle E \rangle^2\right], \quad (4)$$

where k_B denotes the Boltzmann's constant, N denotes the number of atoms, $\langle E^2 \rangle$ and $\langle E \rangle^2$ denotes the run average of energy E and its square E^2. C_V computed using this method does not exhibit a size effect for simulation cell sizes varying from 128–27,648 atoms in both the solid and diffusive states (Fig. S6). C_V of an approximately harmonic solid is expected to plateau at the Dulong-Petit limit, $3NK_B$. Deviation from this value suggests strongly anharmonic behavior, in the present case, entering the diffusive regime.

Fig. 9 shows δ's C_V vs. P, T as background color. For $T < 1,500$ K (blueish background), C_V generally follows the Dulong-Petit limit with at most 5% excess. For $T > 1,800$ K, C_V increases quickly. Depending on the pressure, the background color turns from blue to white then red between 1,800 K and 2,100 K, indicating > 10% deviation from $3Nk_B$ within this color change range. The $D_H = 1$ Å²/ns diffusion boundary corresponds to C_V ~6% above the Dulong-Petit limit. Measurements of C_V could be useful to determine the onset of δ's superionic state.

III. DISCUSSION

This study demonstrates the necessity of adopting a hybrid *ab initio* description aided by deep-learning potential at extreme P, T mantle conditions to address the properties of a hydrous system. The DP-GEN active learning scheme used to develop the potential is highly efficient, requiring only a few thousand massively parallelized DFT calculations on reference
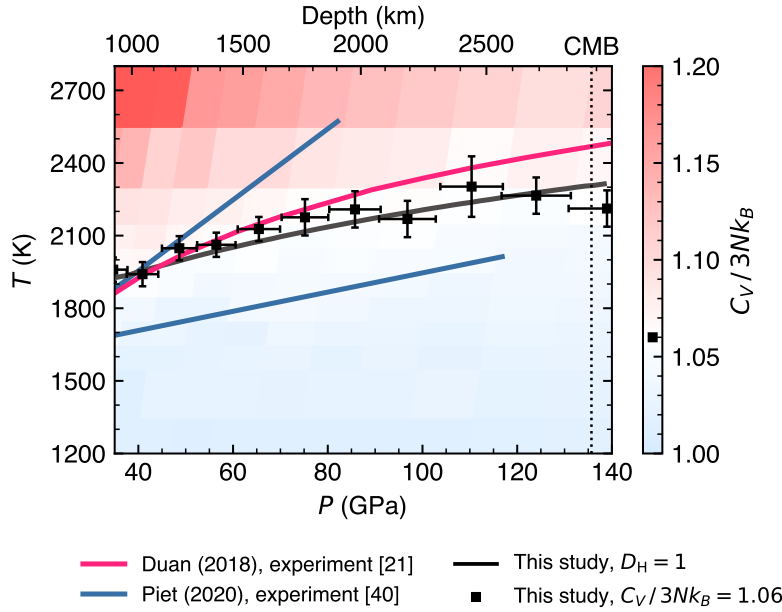
FIG. 9. The ratio between C_V and the Dulong-Petit limit, $3Nk_B$, at different pressures and temperatures. Black symbols indicate a 6% excess above the limit. The error bar is determined based on the P, T grid size. Results are compared to the $D_H = 1$ boundary and measured dehydration phase boundaries [21, 40].

configurations, taking only a few days. The number of *ab initio* calculations needed to prepare the interatomic potential for a wide range of P, T conditions, ~3,000 in this case, is far fewer than that of a typical BOMD run necessary for a single P, T sampling, ~10^4–10^5. This efficiency allows us to use more accurate functionals and perform more accurate simulations using larger simulation cells, longer run times, and denser sampling of points in P, T phase space. Yet, it is still desirable to further develop these potentials to perform under reactive conditions and detect the dehydration process in a simulation.

The current DP-SCAN approach combined with classical MD reproduces measurements of δ's room temperature compression curve surprisingly well. It paves the way for adopting the path-integral approach to quantum ionic dynamics [78, 79] in low-temperature simulations of δ. Zero-point motion effects on the EoS of solids at low temperatures, particularly in H_2O-ice (e.g., [55]), are well-known. Previous studies [3, 27] have identified significant quantum nuclear effects, e.g., tunneling, that could affect H-bond disordering, symmetrization, and details of P, T phase diagrams of hydrous phases (e.g., [80]). Such effects have yet to be fully explored in hydrous minerals. At typical mantle temperatures of thousands of Kelvin, classical MD combined with the SCAN functional seems to reproduce closely the thermal EoS of a prototypical lower mantle hydrous phase changing the H-bonds' state.

At mantle and subducting slab conditions (Fig. 8), it remains challenging to predict the dehydration process, particularly the transition from the superionic state to the dehydration point. The lack of detailed experimental data in this critical range of conditions aggravates understanding of this process. Electrical conductivity measurements commonly identify diffusion. δ's electrical conductivity has only been measured up to 1,200 K and up to 20 GPa, and has shown exponential dependence vs. reciprocal temperature [81] similar to ours. Extending such measurements to more extreme conditions is desirable to

understand better the relation between the superionic state and the dehydration process.

Simulations of these processes will be fundamental to understanding water circulation in Earth's interior. Subducting slabs carry hydrous phases into the mantle. Describing and predicting the behavior of such phases at extreme conditions is key to understanding mantle dynamics, e.g., volcanism, melt generation, etc. While here we focus on the superionic state, we see a relationship between the onset of the superionic state and the dehydration boundary. Our predicted "diffusion boundary" (see Fig. 8), the dehydration boundary reported by Ref. [21], and Ref. [40]'s lower boundary, all share a similar slope that is less steep than the temperature profiles along the subducting slab. These diffusion and dissociation boundaries intercept the subducting slab geotherm [56, 57] at a depth of ~2,400–2,700 km near the bottom of the lower mantle, and the normal mantle geotherm [74] at ~1,200–1,500 km, i.e., approximately mid-mantle. These results confirm previous suggestions that δ could remain stable up to near the bottom of the mantle and only then release water. H_2O released from δ should react with its environment to form other H-bearing phases or melts, e.g., ϵ-FeOOH and FeH$_x$ [40, 82] in the deep mantle. Understanding the entire diffusion to dehydration process and the P, T conditions for these transitions will help clarify the deep water cycle, the potential signature of water presence in seismic tomography, and the effects of water on mantle properties.

METHOD

Machine learning potentials

The NN potential for δ was developed based on the Deep Potential Smooth Edition (DEEPPOT-SE) model [46] implemented in DEEPMD-KIT v2.1 [83, 84]. Two-body embedding with co-

ordinates of the neighboring atoms (se_e2_a) was used for the descriptor. The embedding network shape is (25, 50, 100). The fitting network shape is (128, 128, 128). The cut-off radius is 6 Å, and the smoothing parameter is 0.5 Å.

The model was trained using the ADAM optimizer [85] for 1×10^6 training steps, with the learning rate exponential decaying from 1×10^{-3} to 3.51×10^{-8} throughout the training process. The loss function $\mathcal{L}(p_e, p_f)$ is [83]

$$\mathcal{L}(p_e, p_f) = p_e |\Delta e|^2 + \frac{p_f}{3N} |\Delta f_i|^2, \tag{5}$$

where p_e decays linearly from 1.00 to 0.02, and p_f increases linearly from 1×10^0 to 1×10^3 throughout the training process.

Active learning scheme

The DP-GEN concurrent learning scheme [86] was employed to create the training data set and to generate the potential. We randomly extracted 118 labeled configurations from 59 BOMD runs at various V, Ts to generate the initial potentials and to kickstart the DP-GEN training process. 6 DP-GEN iterations were performed to explore the configuration space and to eventually generate a potential reaching satisfactory accuracy requirement for DPMD for a temperature range of $300 < T < 3,000$ K. 4 candidate DP potentials initialized with different random seeds were trained in each iteration. They were used to perform NVT DPMD simulations at various V, Ts for a few thousand timesteps. After the simulations, the error estimator (model deviation), ϵ_t, is calculated every 50 MD steps based on the force disagreement between the candidate DPs [49, 86]:

$$\epsilon_t = \max_i \sqrt{\langle \|F_{w,i}(\mathcal{R}_t) - \langle F_{w,i}(\mathcal{R}_t) \rangle \|^2 \rangle}, \tag{6}$$

where $F_{w,i}(\mathcal{R}_t)$ denotes the force on the i-th atom predicted by the w-th potential for the $\mathcal{R}_t$ configuration. For a particular configuration, if ϵ_t satisfies $\epsilon_{\min} \leq \epsilon_t \leq \epsilon_{\max}$, the corresponding configuration is collected, then labeled with DFT forces and total energy then added into the training dataset; if $\epsilon_t < \epsilon_{\min}$, these configurations are considered "covered" by the current training dataset; if $\epsilon_t > \epsilon_{\max}$ are considered failed and were discarded. After a few iterations, almost no new configurations are collected according to this standard (> 99% are "accurate" for a few iterations), the DP-GEN process is then complete. The parameters for these DP-GEN iterations are listed in Table SII in detail.

After these DP-GEN iterations, our training dataset consists of 3,487 configurations labeled with *ab initio* force and energy. DPMD simulations were performed using the LAMMPS code [87] with 0.5 fs timesteps.

DFT calculations

Training and testing data sets were based on *ab initio* calculations performed with the VIENNA AB INITIO SIMULATION PACKAGE (VASP) v6.3 [88]. The strongly constrained and appropriately normed (SCAN) [89] meta-GGA functional with PAW basis sets were adopted. The cutoff energy for the plane-wave-basis set was set to 520 eV. The Brillouin zone sampling for the $2 \times 2 \times 4$ supercells (with 16 f.u. of δ, or 128 atoms) used a shifted $2 \times 2 \times 2$ Monkhorst-Pack k-point mesh. BOMD simulations were performed with a timestep of 0.5 fs.

MD simulations

A dataset for validating the DP potential is created by performing both BOMD and DPMD NVT canonical ensemble simulations on $2 \times 2 \times 4$ supercells (128 atoms) with the Nóse-Hoover (NH) thermostat/barostat [90]. These simulations are performed at 6 Ts ($T = 300, 600, 1,200, 1,800, 2,400$, and 3,000 K) and 4 Vs (the conventional cell volume $V = 40, 44, 48$, and 54 Å³). This mesh covers the pressure range of ~20–150 GPa. These simulations start from configurations produced after 10^4 DPMD equilibration timesteps at each given (V, T) and run for a minimum of 1 ns at each (V, T). The DPMD simulation time scales linearly with the number of atoms (see Fig. S7), similarly with previous benchmarks [91].

The investigation of the 300 K compression curve involves DPMD simulations with 1,024-atom (or $4 \times 4 \times 8$) supercells. We perform NPT simulations for 50 ps to equilibrate the structure. Then, perform NVT simulations for another 50 ps to obtain the pressure at the given volume. The investigation of high P, Ts, involved systematic DPMD simulations with 8,192-atom (or $8 \times 8 \times 16$) supercells. We perform NVT simulations. 96 DPMD simulations at different (P, T)'s cover the P, T range from 1,500 K to 3,000 K and 10 GPa to 180 GPa (see Fig. S2). Each MD simulation at equilibrated V, T conditions runs for 2 ns, or 4×10^6 timesteps. These simulations were performed concurrently on 96 GPUs.

Workflow management

Our workflows, including the DP-GEN active learning iterations and subsequent MD simulations, were implemented and managed using the Snakemake [92, 93] workflow management system.

ACKNOWLEDGMENTS

[1] Ohtani, E. The role of water in Earth's mantle. *National Science Review* **7**, 224–232 (2020).

[2] Ohtani, E. Hydration and Dehydration in Earth's Interior. *Annual Review of Earth and Planetary Sciences* **49**, 253–278 (2021).

[3] Benoit, M. & Marx, D. The Shapes of Protons in Hydrogen Bonds Depend on the Bond Length. *ChemPhysChem* **6**, 1738–1741 (2005).

[4] Zhang, L. *et al.* Deep neural network for the dielectric response of insulators. *Physical Review B* **102**, 041121 (2020).

[5] Umemoto, K., Wentzcovitch, R. M., de Gironcoli, S. & Baroni, S. Order–disorder phase boundary between ice VII and VIII obtained by first principles. *Chemical Physics Letters* **499**, 236–240 (2010).

[6] Hernandez, J.-A. & Caracas, R. Superionic-Superionic Phase Transitions in Body-Centered Cubic H 2 O Ice. *Physical Review Letters* **117**, 135503 (2016).

[7] Jacobsen, S. D. *et al.* Effects of hydration on the elastic properties of olivine. *Geophysical Research Letters* **35** (2008).

[8] Wang, W., Zhang, H., Brodholt, J. P. & Wu, Z. Elasticity of hydrous ringwoodite at mantle conditions: Implication for water distribution in the lowermost mantle transition zone. *Earth and Planetary Science Letters* **554**, 116626 (2021).

[9] Wang, W. & Wu, Z. A first-principles study of water in wadsleyite and ringwoodite: Implication for the 520 km discontinuity. *American Mineralogist* **107**, 1361–1368 (2022).

[10] Han, G. *et al.* Pervasive low-velocity layer atop the 410-km discontinuity beneath the northwest Pacific subduction zone: Implications for rheology and geodynamics. *Earth and Planetary Science Letters* **554**, 116642 (2021).

[11] van Keken, P. E., Hacker, B. R., Syracuse, E. M. & Abers, G. A. Subduction factory: 4. Depth-dependent flux of H2O from subducting slabs worldwide. *Journal of Geophysical Research: Solid Earth* **116** (2011).

[12] Ohtani, E., Yuan, L., Ohira, I., Shatskiy, A. & Litasov, K. Fate of water transported into the deep mantle by slab subduction. *Journal of Asian Earth Sciences* **167**, 2–10 (2018).

[13] Deng, X., Luo, C., Wentzcovitch, R. M., Abers, G. A. & Wu, Z. Elastic anisotropy of lizardite at subduction zone conditions. *Geophysical Research Letters* **n/a**, e2022GL099712 (2022).

[14] Perdew, J. P. & Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Physical Review B* **23**, 5048–5079 (1981).

[15] Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical Review Letters* **77**, 3865–3868 (1996).

[16] Qin, T., Zhang, Q., Wentzcovitch, R. M. & Umemoto, K. Qha: A Python package for quasiharmonic free energy calculation for multi-configuration systems. *Computer Physics Communications* **237**, 199–207 (2019).

[17] Wu, Z. & Wentzcovitch, R. M. Quasiharmonic thermal elasticity of crystals: An analytical approach. *Physical Review B* **83**, 184115 (2011).

[18] Luo, C. *et al.* Cij: A Python code for quasiharmonic thermoelasticity. *Computer Physics Communications* **267**, 108067 (2021).

[19] Wentzcovitch, R. M. & Martins, J. First principles molecular dynamics of Li: Test of a new algorithm. *Solid State Communications* **78**, 831–834 (1991).

[20] Suzuki, A., Ohtani, E. & Kamada, T. A new hydrous phase δ-AlOOH synthesized at 21 GPa and 1000 °C. *Physics and Chemistry of Minerals* **27**, 689–693 (2000).

[21] Duan, Y. *et al.* Phase stability and thermal equation of state of δ-AlOOH: Implication for water transportation to the Deep Lower Mantle. *Earth and Planetary Science Letters* **494**, 92–98 (2018).

[22] Sano-Furukawa, A. *et al.* Change in compressibility of δ-AlOOH and δ-AlOOD at high pressure: A study of isotope effect and hydrogen-bond symmetrization. *American Mineralogist* **94**, 1255–1261 (2009).

[23] Sano-Furukawa, A. *et al.* Direct observation of symmetrization of hydrogen bond in δ-AlOOH under mantle conditions using neutron diffraction. *Scientific Reports* **8**, 15520 (2018).

[24] Tsuchiya, J., Tsuchiya, T., Tsuneyuki, S. & Yamanaka, T. First principles calculation of a high-pressure hydrous phase, δ-AlOOH. *Geophysical Research Letters* **29**, 15–1–15–4 (2002).

[25] Tsuchiya, J., Tsuchiya, T. & Wentzcovitch, R. M. Vibrational properties of δ-AlOOH under pressure. *American Mineralogist* **93**, 477–482 (2008).

[26] Tsuchiya, J. & Tsuchiya, T. Elastic properties of δ-AlOOH under pressure: First principles investigation. *Physics of the Earth and Planetary Interiors* **174**, 122–127 (2009).

[27] Luo, C., Umemoto, K. & Wentzcovitch, R. M. *Ab Initio* investigation of H-bond disordering in δ-AlOOH. *Physical Review Research* **4**, 023223 (2022).

[28] Ohtani, E., Litasov, K., Suzuki, A. & Kondo, T. Stability field of new hydrous phase, δ-AlOOH, with implications for water transport into the deep mantle. *Geophysical Research Letters* **28**, 3991–3993 (2001).

[29] He, Y. *et al.* Superionic hydrogen in Earth's deep interior (2018). 1810.08766.

[30] Tsuchiya, J. First principles prediction of a new high-pressure phase of dense hydrous magnesium silicates in the lower mantle. *Geophysical Research Letters* **40**, 4570–4573 (2013).

[31] Tsuchiya, T., Tsuchiya, J., Dekura, H. & Ritterbex, S. Ab Initio Study on the Lower Mantle Minerals. *Annual Review of Earth and Planetary Sciences* **48**, 99–119 (2020).

[32] Sawai, M., Katayama, I., Hamada, A., Maeda, M. & Nakashima, S. Dehydration kinetics of antigorite using in situ high-temperature infrared microspectroscopy. *Physics and Chemistry of Minerals* **40**, 319–330 (2013).

[33] Ingrin, J., Hercule, S. & Charton, T. Diffusion of hydrogen in diopside: Results of dehydration experiments. *Journal of Geophysical Research: Solid Earth* **100**, 15489–15499 (1995).

[34] Cavazzoni, C. *et al.* Superionic and Metallic States of Water and Ammonia at Giant Planet Conditions. *Science* **283**, 44–46 (1999).

[35] Katoh, E., Yamawaki, H., Fujihisa, H., Sakashita, M. & Aoki, K. Protonic Diffusion in High-Pressure Ice VII. *Science* **295**, 1264–1266 (2002).

[36] Goldman, N., Fried, L. E., Kuo, I.-F. W. & Mundy, C. J. Bonding in the Superionic Phase of Water. *Physical Review Letters* **94**, 217801 (2005).

[37] Wilson, H. F., Wong, M. L. & Militzer, B. Superionic to Superionic Phase Change in Water: Consequences for the Interiors of Uranus and Neptune. *Physical Review Letters* **110**, 151102

(2013).

[38] Sun, J., Clark, B. K., Torquato, S. & Car, R. The phase diagram of high-pressure superionic ice. *Nature Communications* **6**, 8156 (2015).

[39] Millot, M. *et al.* Nanosecond X-ray diffraction of shock-compressed superionic water ice. *Nature* **569**, 251–255 (2019).

[40] Piet, H. *et al.* Dehydration of δ-AlOOH in Earth's Deep Lower Mantle. *Minerals* **10**, 384 (2020).

[41] Behler, J. & Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Physical Review Letters* **98**, 146401 (2007).

[42] Bartók, A. P., Payne, M. C., Kondor, R. & Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Physical Review Letters* **104**, 136403 (2010).

[43] Thompson, A., Swiler, L., Trott, C., Foiles, S. & Tucker, G. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *Journal of Computational Physics* **285**, 316–330 (2015).

[44] Shapeev, A. V. Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials. *Multiscale Modeling & Simulation* **14**, 1153–1173 (2016).

[45] Artrith, N., Urban, A. & Ceder, G. Efficient and accurate machine-learning interpolation of atomic energies in compositions with many species. *Physical Review B* **96**, 014112 (2017).

[46] Zhang, L. *et al.* End-to-end symmetry preserving inter-atomic potential energy model for finite and extended systems. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, 4441–4451 (Curran Associates Inc., Red Hook, NY, USA, 2018).

[47] Zuo, Y. *et al.* Performance and Cost Assessment of Machine Learning Interatomic Potentials. *The Journal of Physical Chemistry A* **124**, 731–745 (2020).

[48] Smith, J. S., Nebgen, B., Lubbers, N., Isayev, O. & Roitberg, A. E. Less is more: Sampling chemical space with active learning. *The Journal of Chemical Physics* **148**, 241733 (2018).

[49] Zhang, L., Lin, D.-Y., Wang, H., Car, R. & E, W. Active learning of uniformly accurate interatomic potentials for materials simulation. *Physical Review Materials* **3**, 023804 (2019).

[50] Zhang, C. *et al.* Deep machine learning potential for atomistic simulation of Fe-Si-O systems under Earth's outer core conditions. *Physical Review Materials* **6**, 063802 (2022).

[51] Wan, T., Luo, C., Sun, Y. & Wentzcovitch, R. M. Thermoelastic properties of bridgmanite using Deep Potential Molecular Dynamics (2023). 2307.07127.

[52] Sun, J. *et al.* Accurate first-principles structures and energies of diversely bonded systems from an efficient density functional. *Nature Chemistry* **8**, 831–836 (2016).

[53] Zhang, L., Wang, H., Car, R. & E, W. Phase Diagram of a Deep Potential Water Model. *Physical Review Letters* **126**, 236001 (2021).

[54] Zhang, C. *et al.* Modeling Liquid Water by Climbing up Jacob's Ladder in Density Functional Theory Facilitated by Using Deep Neural Network Potentials. *The Journal of Physical Chemistry B* **125**, 11444–11456 (2021).

[55] Umemoto, K. *et al.* Nature of the Volume Isotope Effect in Ice. *Physical Review Letters* **115**, 173005 (2015).

[56] Eberle, M. A., Grasset, O. & Sotin, C. A numerical study of the interaction between the mantle wedge, subducting slab, and overriding plate. *Physics of the Earth and Planetary Interiors* **134**, 191–202 (2002).

[57] Ohtani, E. Hydrous minerals and the storage of water in the deep mantle. *Chemical Geology* **418**, 6–15 (2015).

[58] Litasov, K. D. & Ohtani, E. Phase relations in hydrous MORB at 18–28GPa: Implications for heterogeneity of the lower mantle. *Physics of the Earth and Planetary Interiors* **150**, 239–263 (2005).

[59] Behler, J. Four Generations of High-Dimensional Neural Network Potentials. *Chemical Reviews* **121**, 10037–10072 (2021).

[60] Bronstein, Y., Depondt, P. & Finocchi, F. Thermal and nuclear quantum effects in the hydrogen bond dynamical symmetrization phase transition of δ-AlOOH. *European Journal of Mineralogy* **29**, 385–395 (2017).

[61] Wentzcovitch, R. M., Wu, Z. & Carrier, P. First Principles Quasiharmonic Thermoelasticity of Mantle Minerals. *Reviews in Mineralogy and Geochemistry* **71**, 99–128 (2010).

[62] Suzuki, A. Compressibility of the high-pressure polymorph of AlOOH to 17 GPa. *Mineralogical Magazine* **73**, 479–485 (2009).

[63] Kuribayashi, T., Sano-Furukawa, A. & Nagase, T. Observation of pressure-induced phase transition of δ-AlOOH by using single-crystal synchrotron X-ray diffraction method. *Physics and Chemistry of Minerals* **41**, 303–312 (2014).

[64] Kang, D. *et al.* Hydrogen-Bond Symmetrization of δ-AlOOH. *Chinese Physics Letters* **34**, 108301 (2017).

[65] Pillai, S. B., Jha, P. K., Padmalal, A., Maurya, D. M. & Chamyal, L. S. First principles study of hydrogen bond symmetrization in δ-AlOOH. *Journal of Applied Physics* **123**, 115901 (2018).

[66] Mashino, I., Murakami, M. & Ohtani, E. Sound velocities of δ-AlOOH up to core-mantle boundary pressures with implications for the seismic anomalies in the deep mantle. *Journal of Geophysical Research: Solid Earth* **121**, 595–609 (2016).

[67] Simonova, D. *et al.* Structural Study of δ-AlOOH Up to 29 GPa. *Minerals* **10**, 1055 (2020).

[68] Benoit, M., Marx, D. & Parrinello, M. Quantum effects on phase transitions in high-pressure ice. *Computational Materials Science* **10**, 88–93 (1998).

[69] Benoit, M., Marx, D. & Parrinello, M. Tunnelling and zero-point motion in high-pressure ice. *Nature* **392**, 258–261 (1998).

[70] Wang, H. *et al.* Quantum structural fluxion in superconducting lanthanum polyhydride. *Nature Communications* **14**, 1674 (2023).

[71] Huang, D., Badro, J., Brodholt, J. & Li, Y. Ab Initio Molecular Dynamics Investigation of Molten Fe–Si–O in Earth's Core. *Geophysical Research Letters* **46**, 6397–6405 (2019).

[72] Tuckerman, M. E. *Statistical Mechanics: Theory and Molecular Simulation* (Oxford University Press, Oxford ; New York, 2010).

[73] Umemoto, K., Kawamura, K., Hirose, K. & Wentzcovitch, R. M. Post-stishovite transition in hydrous aluminous SiO2. *Physics of the Earth and Planetary Interiors* **255**, 18–26 (2016).

[74] Brown, J. M. & Shankland, T. J. Thermodynamic parameters in the Earth as determined from seismic profiles. *Geophysical Journal of the Royal Astronomical Society* **66**, 579–596 (1981).

[75] Grasselli, F. Investigating finite-size effects in molecular dynamics simulations of ion diffusion, heat transport, and thermal motion in superionic materials. *The Journal of Chemical Physics* **156**, 134705 (2022).

[76] Potashnikov, S., Boyarchenkov, A., Nekrasov, K. & Kupryazhkin, A.Ya. High-precision molecular dynamics simulation of UO2–PuO2: Anion self-diffusion in UO2. *Journal of Nuclear Materials* **433**, 215–226 (2013).

[77] Klochko, L., Baschnagel, J., Wittmer, J. P. & Semenov, A. N. General relations to obtain the time-dependent heat capacity from isothermal simulations. *The Journal of Chemical Physics* **154**, 164501 (2021).

[78] Tuckerman, M. E., Marx, D., Klein, M. L. & Parrinello, M. Efficient and general algorithms for path integral Car–Parrinello

molecular dynamics. *The Journal of Chemical Physics* **104**, 5579–5588 (1996).

[79] Marx, D. & Parrinello, M. Ab initio path integral molecular dynamics: Basic ideas. *The Journal of Chemical Physics* **104**, 4077–4082 (1996).

[80] Ly, K. K. & Ceperley, D. M. Stability and distortion of fcc LaH 10 with path-integral molecular dynamics. *Physical Review B* **106**, 054106 (2022).

[81] Wang, R. & Yoshino, T. Electrical conductivity of diaspore, δ-AlOOH and ϵ-FeOOH. *American Mineralogist* **106**, 774–781 (2021).

[82] Terasaki, H. *et al.* Stability of Fe–Ni hydride after the reaction between Fe–Ni alloy and hydrous phase (δ-AlOOH) up to 1.2Mbar: Possibility of H contribution to the core density deficit. *Physics of the Earth and Planetary Interiors* **194–195**, 18–24 (2012).

[83] Wang, H., Zhang, L., Han, J. & E, W. DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Computer Physics Communications* **228**, 178–184 (2018).

[84] Zeng, J. *et al.* DeePMD-kit v2: A software package for deep potential models. *The Journal of Chemical Physics* **159**, 054801 (2023).

[85] Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization (2017). 1412.6980.

[86] Zhang, Y. *et al.* DP-GEN: A concurrent learning platform for the generation of reliable deep learning based potential energy models. *Computer Physics Communications* **253**, 107206 (2020).

[87] Thompson, A. P. *et al.* LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications* **271**, 108171 (2022).

[88] Kresse, G. & Furthmüller, J. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Physical Review B* **54**, 11169–11186 (1996).

[89] Sun, J., Ruzsinszky, A. & Perdew, J. P. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Physical Review Letters* **115**, 036402 (2015).

[90] Hoover, W. G. & Holian, B. L. Kinetic moments method for the canonical ensemble distribution. *Physics Letters A* **211**, 253–257 (1996).

[91] Zhang, L., Han, J., Wang, H., Car, R. & E, W. Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics. *Physical Review Letters* **120**, 143001 (2018).

[92] Köster, J. & Rahmann, S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* **28**, 2520–2522 (2012).

[93] Mölder, F. *et al.* Sustainable data analysis with Snakemake. *F1000Research* **10**, 33 (2021).

[94] Towns, J. *et al.* XSEDE: Accelerating Scientific Discovery. *Computing in Science & Engineering* **16**, 62–74 (2014).

Supplementary Information

High throughput sampling of phase space with deep learning potentials: δ-AlOOH at geophysical conditions

Chenxing Luo [iD][1], Yang Sun [iD][*,1,2], Renata M. Wentzcovitch [iD][†,1,3,4]

[1]*Department of Applied Physics and Applied Mathematics, Columbia University, 500 West 120 Street, New York, 10027, New York, USA*
[2]*Department of Physics and Astronomy, Iowa State University, Ames, Iowa 50011, USA*
[3]*Department of Earth and Environmental Sciences, Columbia University, 1200 Amsterdam Avenue, New York, 10027, New York, USA*
[4]*Lamont–Doherty Earth Observatory, Columbia University, 61 Route 9 West, Palisades, 10964, New York, USA*

This PDF file includes:

- Tables SI and SII

- Figures S1 to S8

[*]yangsun@iastate.edu
[†]rmw2150@columbia.edu

TABLE SI – Details for RMSE of force and potential energies for validation.

T	V	Energy RMSE/N (eV/atom)	Force RMSE (eV/Å)
600.0	54.0	0.001873	0.035725
600.0	44.0	0.002231	0.030642
600.0	40.0	0.002494	0.031809
600.0	48.0	0.002179	0.032381
1200.0	44.0	0.002140	0.047643
1200.0	48.0	0.002202	0.049318
1200.0	54.0	0.001965	0.052093
1200.0	40.0	0.002487	0.052524
1800.0	48.0	0.002060	0.067983
1800.0	44.0	0.001998	0.064704
1800.0	54.0	0.002045	0.070243
1800.0	40.0	0.002315	0.073160
2400.0	48.0	0.002096	0.084835
2400.0	44.0	0.002090	0.084830
2400.0	40.0	0.002492	0.091925
2400.0	54.0	0.001497	0.101425
3000.0	48.0	0.001981	0.122193
3000.0	54.0	0.001948	0.124198
3000.0	44.0	0.001831	0.106954
3000.0	40.0	0.002675	0.119190

TABLE SII – Details of the DP-GEN iterations for generating the training dataset.

Iteration	Temperature	Criteria for selection (eV/Å)	# of configurations collected
Initial	300	N/A	118
1	300, 600, 900	$[0.20, 0.30]$	878
2	300, 600, 900	$[0.20, 0.30]$	547
3	300, 600, 900, 1200, 1500	$[0.25, 0.35]$	554
4	300, 600, 900, 1200, 1500, 1800, 2100	$[0.25, 0.35]$	390
5	300, 600, 900, 1200, 1500, 1800, 2100, 2700	$[0.25, 0.35]$	1000
Total			3487
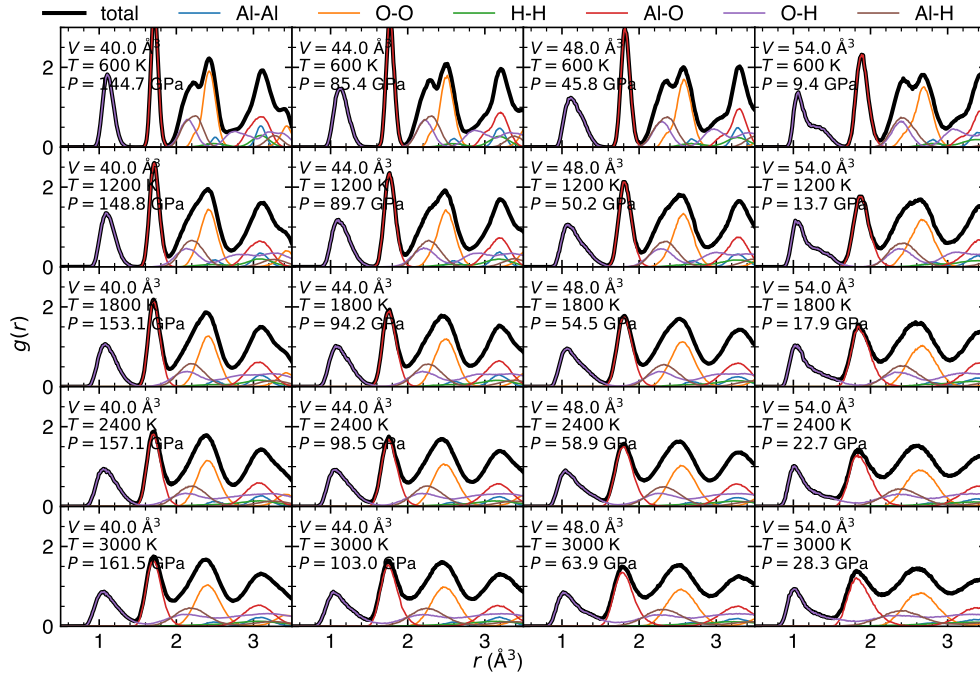
FIGURE S1 – $g(r)$ and partial $g(r)$ at various V, T.
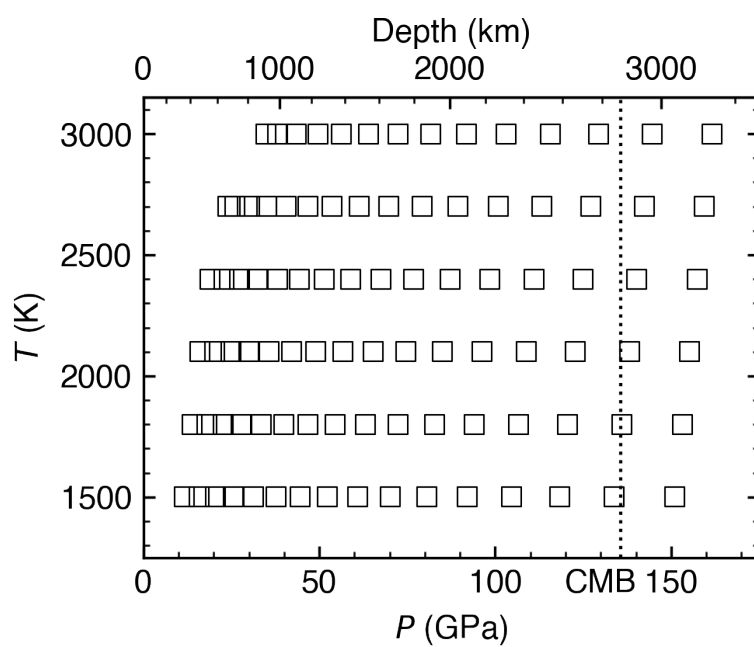
FIGURE S2 – The (P, T)'s for systematic DPMD simulations.
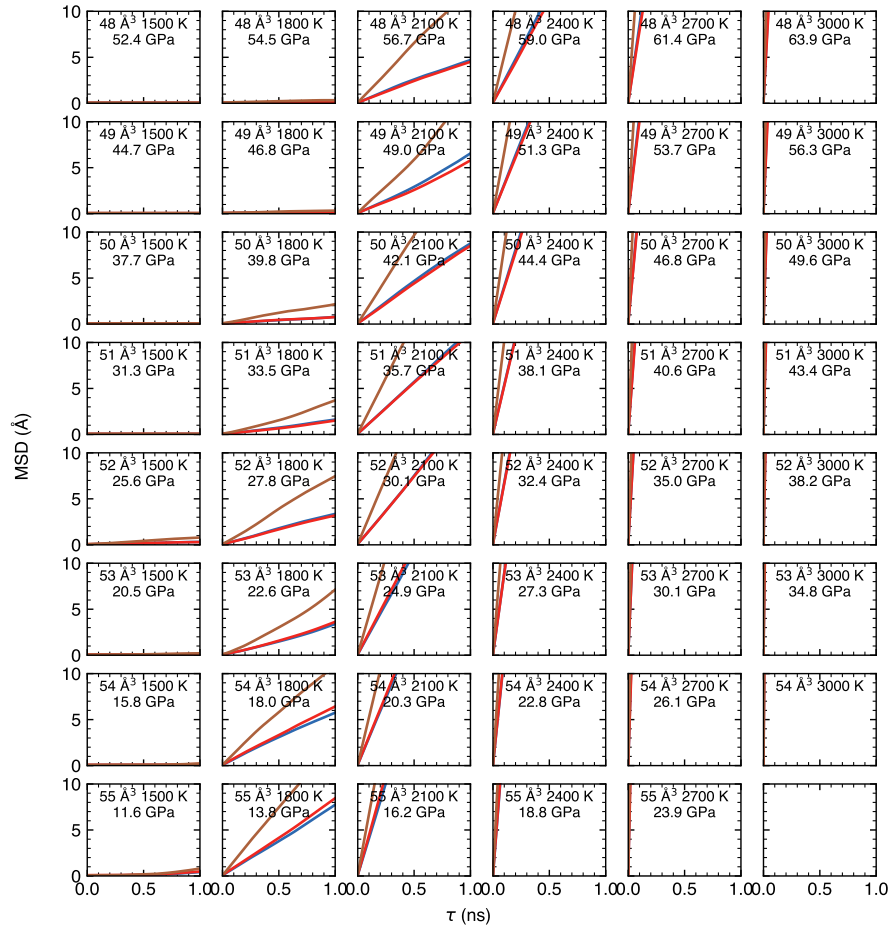
FIGURE S3 – (a) The proton diffusivity at various V, Ts at $V < 48$ Å^3
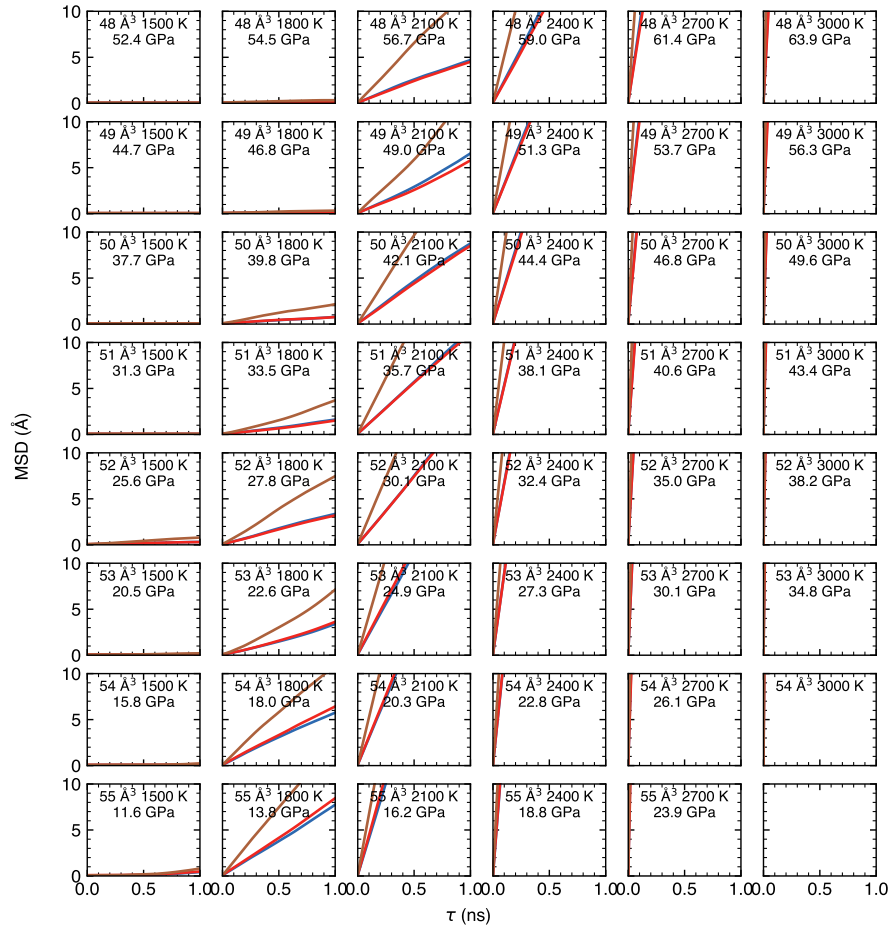
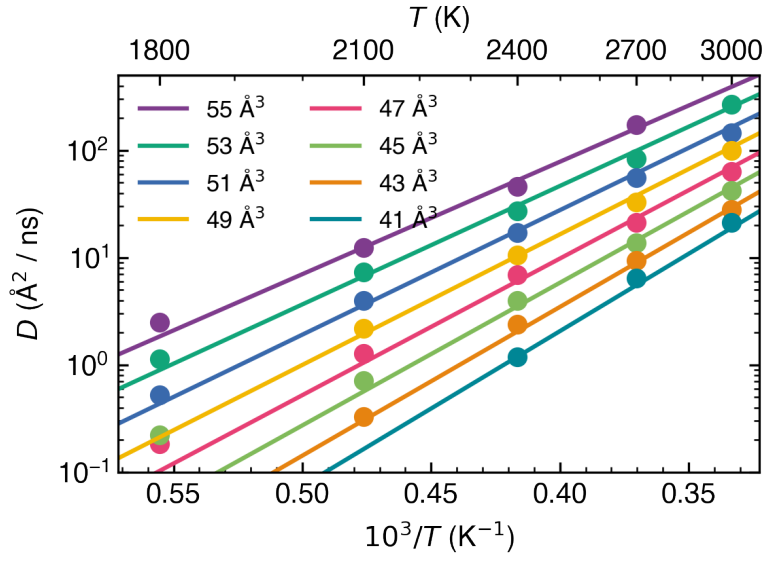**FIGURE S3** – (b) The proton diffusivity at various V, Ts at $V \geq 48$ Å^3

FIGURE S4 – D_H vs. T. Solid lines represent the interpolation $\log D_H(V, T) = aV/T + b/T + c$ at different Vs.
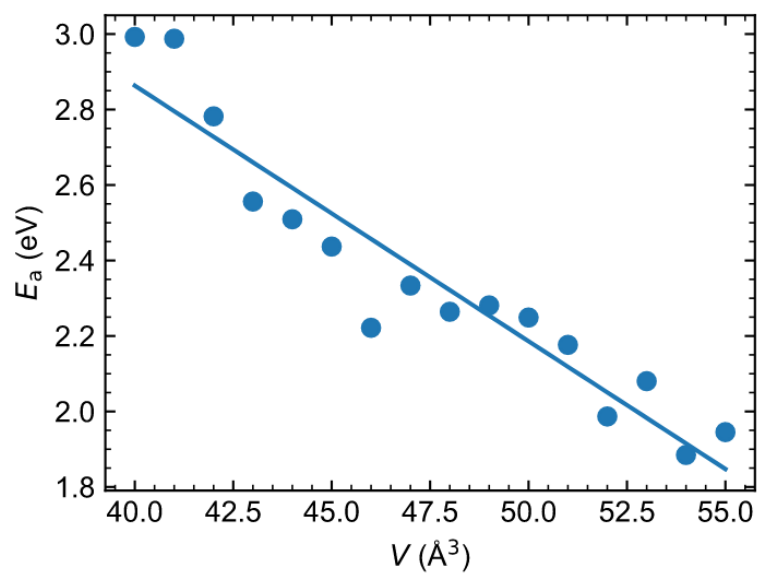
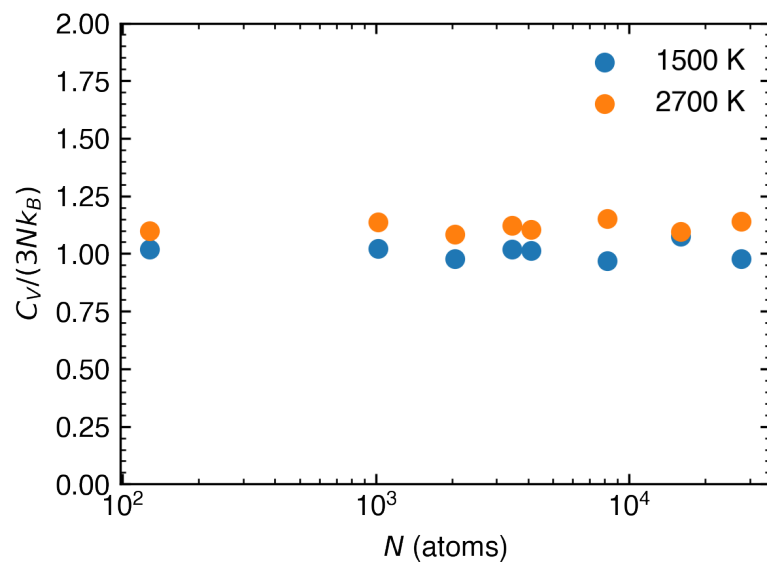**FIGURE S5** – The diffusion activation energy E_a vs. unit-cell volume V.

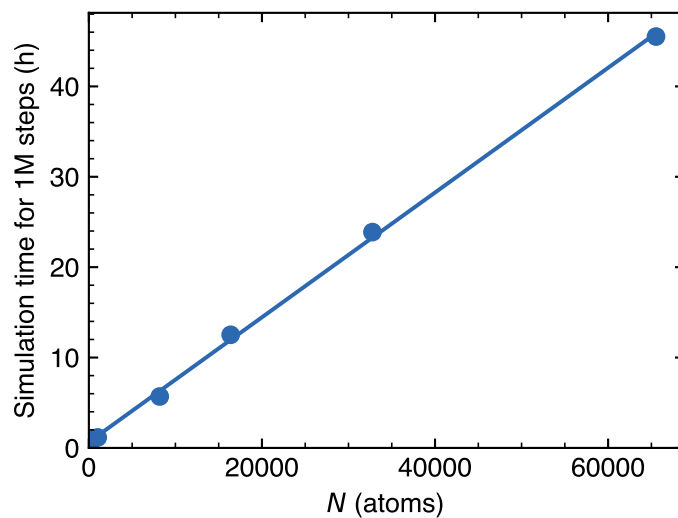**FIGURE S6** – Size effect on C_V at $\sim$110 GPa.

FIGURE S7 – Machine hours for DPMD simulation with 10^6 MD timesteps vs. simulation cell size (128, 256, 512, 1024, 8192, 16,384, 32,768, and 65,536 atoms) on an NVIDIA A100-SXM4-40GB GPU.