Supplementary Information (SI) for Physical Chemistry Chemical Physics.

Supporting information for "Data-Driven Molecular Dynamics Simulation of Water Isotope Separation using a Catalytically Active Ultrathin Membrane "

Jinu Jeong[a], Chenxing Liang[b], and Narayana R. Aluru*[bc]

[a.] *Department of Mechanical Science and Engineering, The University of Illinois at Urbana–Champaign, Urbana, Illinois, 61801 United States.*
[b.] *Walker Department of Mechanical Engineering, The University of Texas at Austin, Austin, Texas, 78712 United States.*
[c.] *Oden Institute for Computational Engineering & Sciences, The University of Texas at Austin, Austin, Texas, 78712 United States. Email: aluru@utexas.edu*

S.1. AIMD system for stability tests

We observed that D2O, deuterium-dominant hydronium isotopes, and H^+ ions dominate net inflow (Figure 3.A) and that Mo-active site is more frequently occupied by OH than OD (Figure 3.C). Inspired by that the molecules in the net inflow are related to OH/OD bonds: H_2O, H_3O^+, and H_2DO^+ have more OH bonds than OD bonds; D_2O, HD_2O^+, and D_3O^+ have more OD bonds than OH bonds; HDO has the same number of OH and OD bonds; H^+/D^+ ions are the products of OH/OD bond breaks. We perform AIMD simulations of the systems shown in Figure S.1 to explain the observations (Figure 3.A and C).

S.2. Subsystems for data generation

Since our target system (Figure S2.D) is composed of 552 water molecules, a large number for AIMD simulation, the computational cost is burdensome even for training data generation. For example, a cluster computer environment with four nodes (48 cores per node and Intel Xeon Platinum 8160 2.1GHz) can run 0.63 ps of AIMD simulation in 24 hours. To efficiently generate training data, we consider three subsystems shown in Figures S2.A-C, which are designed to enrich water-water interaction, water-wall interaction, and water-pore interaction, respectively.

S.3. Physics validations

To explicitly validate that our neural network potential learns isotopic differences and AIMD-level physics, we compute several physical quantities from DeePMD simulations and compare them with AIMD results. Here, we consider RDF, bond length distribution, bond angle distribution, mean-squared deviations (MSD), velocity autocorrelation function (VACF), and power spectra and show that our neural network potential can accurately reproduce those physical properties for H_2O and D_2O, which are shown in Figures S3.A-F and Figures S3.G-L, respectively.

S.4. Computational cost of AIMD and DeePMD compared

We conduct a comprehensive benchmark of the simulation speed between AIMD and DeePMD across a range of systems. The specifics of each system, including the duration of the simulations and the computing resources utilized, are detailed in Table S.1. While DeePMD is still more expensive than classical MD simulation, our benchmarks reveal that DeePMD offers a reduction in computational cost compared to AIMD, allowing for much longer simulation times and larger system sizes without compromising accuracy.

S.5. Benchmark comparison of critical quantities against higher accuracy methods

We perform additional benchmark simulation with more advanced simulation techniques. The revPBE functional is a revised version of the PBE functional, designed to correct some of the deficiencies of the original PBE, particularly for surface science applications and systems where adsorption processes are important. The SCAN functional, based on the SCAN meta-GGA framework, provides improved accuracy for a wide range of systems, including those with intermediate range van der Waals interactions. The cohesion energy obtained with PBE-D3 is 5.09026 eV, while the values obtained with revPBE and SCAN are 4.77152 eV and 5.59619 eV, respectively. Given that the SCAN functional is considered the most accurate due to its meta-GGA essence, we compared the differences between the PBE-D3 and SCAN functionals. The difference is within 9%, indicating that PBE-D3 provides a reasonably accurate approximation, even outperforming the revPBE functional for our system. These benchmark comparisons demonstrate that the PBE-D3 functional is reliable for our study and provides results that are in agreement with those from more accurate methods.

S.6. DFT analysis of OH adsorption at Mo atoms in the nanopore

We investigated the interaction between one OH group and the nanopore by obtaining charge transfer density maps (Figure S.4.) Here, the red isosurface represents electron accumulation, while the blue isosurface represents electron depletion. The results indicate that the oxygen atom interacts strongly with the Mo atom in the nanopore, with charge transfer from the Mo atom to the oxygen atom. This significant hybridization between the O 2p orbitals and the Mo d orbitals suggests the formation of a strong covalent bond, resulting in the stable attachment of the oxygen atom to the Mo atom throughout the simulation. Additionally, it is observed that the hydrogen atom loses some electrons to the oxygen atom, though this interaction is not as strong as the charge transfer between the Mo and O atoms. This weaker interaction between the O atom and the H/D atom makes the exchange between H and D atoms feasible. We also quantified the interaction between the Mo atom and the OH group by calculating the cohesion energy. Using the PBE-D3 functional, the cohesion energy was calculated to be 5.09026 eV, indicating a strong interaction between the Mo atom and the OH group.

Figures

Figure S.1.

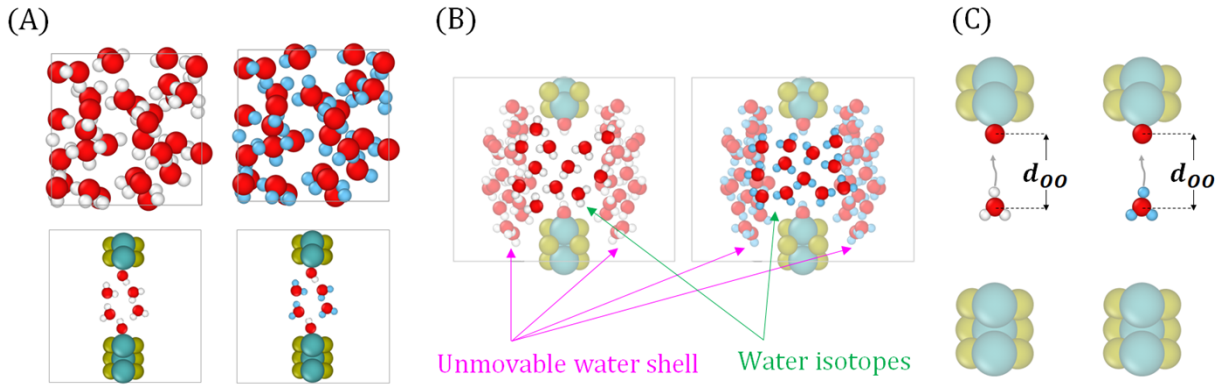

Figure S.1. AIMD systems for OH/OD bond stability test. (A) Bulk H_2O (top left), bulk D_2O (top right), confined H_2O (bottom left), and confined D_2O (bottom right). To quantify the relative instability of each water isotope in confinement, potential energy of each system is measured and compared. (B) Pore-confined H_2O and D_2O molecules surrounded by fixed water molecules. Bond length and the force acting on H/D atoms are computed and their standard deviation, a quantification of their fluctuation, are calculated to compare stability difference of OH and OD bond in confinement. (C) A single H_3O^+ and D_3O^+ in a MoS_2 nanopore, where a single oxygen atom is attached to a Mo atom. H^+/D^+ transfer time considering various O to O distances are measured.

Figure S.2

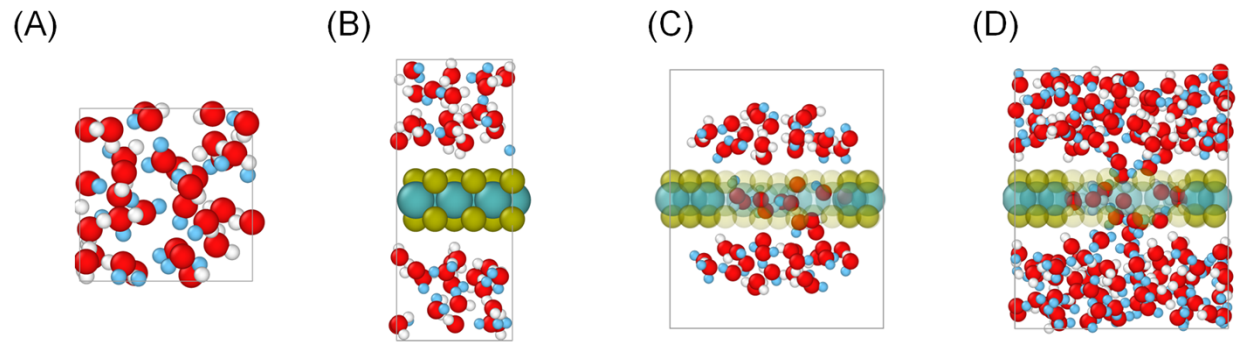

Figure S.2. Configurations of the subsystems and the target system. (A) A bulk water system composed of eight H_2O, sixteen HDO, and eight D_2O, considered for water-water interaction enrichment. (B) A system composed of a single-layer MoS_2 wall (without pore) surrounded by eleven H_2O, twenty-two HDO, and eleven D_2O, considered for water-wall interaction enrichment. (C) A system composed of a single-layer MoS_2 nanopore with a fixed water shell that confines two H_2O, five HDO, and two D_2O, considered for water-pore interaction enrichment. (D) The target system.

Figure S.3

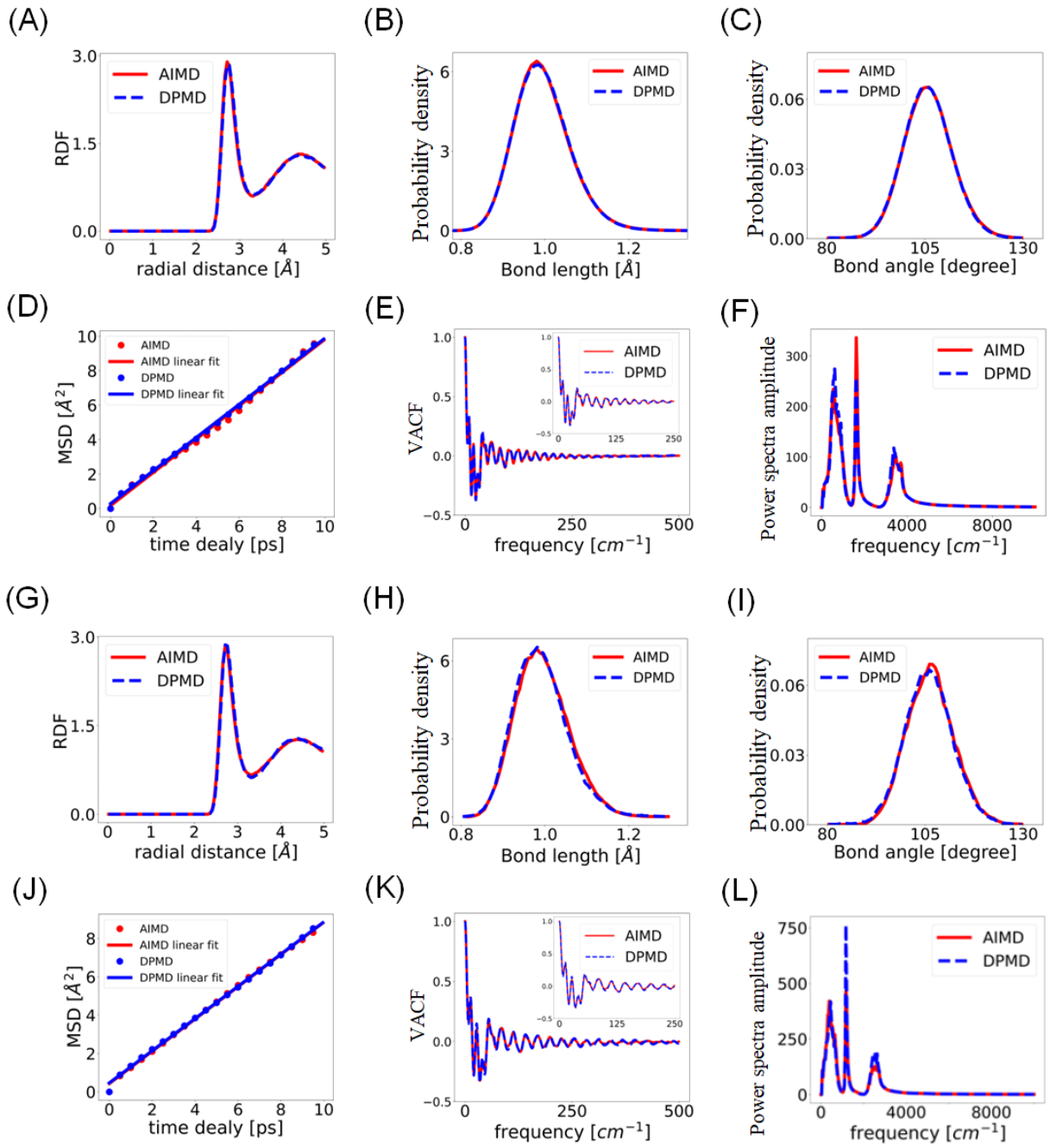

Figure S.3. Validation of DPMD simulations. The physical properties of DeePMD simulations are computed and compared with AIMD results. (A-F) RDF, bond length distribution, bond angle distribution, MSD, VACF, and power spectra of bulk H_2O. (G-L) The same properties of bulk D_2O.

Figure S.4.

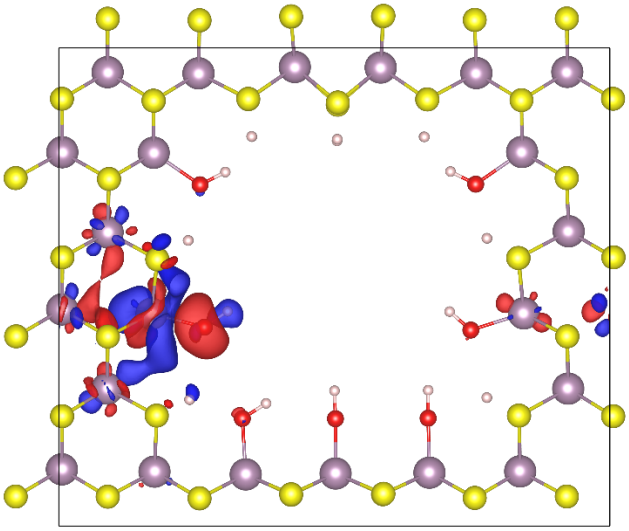

Figure S.4. Charge density map of the OH group interacting with the nanopore, where red indicates electron accumulation and blue represents electron depletion, illustrating strong covalent bonding between the O atom and the Mo atom.

Tables

Table S.1.

	Simulation time	**Wall clock time**	**# node**	**# core per node**	**CPU model and clock**
Bulk H2O (AIMD)	12.08 ps	48 hrs	4	48	Intel Xeon Platinum 8160 2.1GHz
Bulk D2O (AIMD)	13.97 ps	48 hrs	4	48	Intel Xeon Platinum 8160 2.1GHz
Bulk Mixture (AIMD)	13.51 ps	48 hrs	4	48	Intel Xeon Platinum 8160 2.1GHz
H2O with MoS2 nanopore (AIMD)	1.15 ps	48 hrs	4	48	Intel Xeon Platinum 8160 2.1GHz
D2O with MoS2 nanopore (AIMD)	1.26 ps	48 hrs	4	48	Intel Xeon Platinum 8160 2.1GHz
Mixture water with MoS2 nanopore (AIMD)	1.14 ps	48 hrs	4	48	Intel Xeon Platinum 8160 2.1GHz
Bulk H2O (DPMD)	220.86 ps	1 hr	1	16	Intel E5-2660 2.2 GHz with NVIDIA V100 GPU
Bulk D2O (DPMD)	219.51 ps	1 hr	1	16	Intel E5-2660 2.2 GHz with NVIDIA V100 GPU

Bulk Mixture (DPMD)	190.48 ps	1 hr	1	16	Intel E5-2660 2.2 GHz with NVIDIA V100 GPU
H2O with MoS2 nanopore (DPMD)	40.81 ps	1 hr	1	16	Intel E5-2660 2.2 GHz with NVIDIA V100 GPU
D2O with MoS2 nanopore (DPMD)	40.58 ps	1 hr	1	16	Intel E5-2660 2.2 GHz with NVIDIA V100 GPU
Mixture water with MoS2 nanopore (DPMD)	40.89 ps	1 hr	1	16	Intel E5-2660 2.2 GHz with NVIDIA V100 GPU

Table S.1. Simulation Speed Benchmarking of AIMD and DeePMD. The benchmarking results of AIMD and DeePMD across various systems are compared. The table includes the simulation time in picoseconds, wall clock time in hours, the number of nodes, the number of cores per node, and the CPU model (and GPU accelerator) and clock speed used for each simulation. The results highlight the differences in computational cost and efficiency between AIMD and DeePMD, showcasing the ability of DeePMD allowing for longer simulations.