Supplementary Information (SI) for Digital Discovery.

Supplementary information

Enhancing High-Dimensional Neural Network Potentials Accuracy in OLED Systems via Element-Relabeling

Authors: Yonghwan Yun[a], Dongmin Park[a], Junyoung Choi[a], Dong Shin Choi[b], Yousung Jung[*,ac]

[a]Department of Chemical and Biological Engineering, and Institute of Chemical Processes, Seoul National University, 1 Gwa-nak-ro, Gwanak-gu, Seoul 08826, Republic of Korea.

[b]Virtual Design Engineering Division, LG Display, 245 LG-ro, Wollongong-myeon, Paju-si, Gyeonggi-do 10845, Republic of Korea

[c]Institute of Engineering Research, Seoul National University, 1 Gwa-nak-ro, Gwanak-gu, Seoul 08826, Republic of Korea.

CONTENTS

1. Computation Details

For reproducibility, the exact software versions used in this work are as follows: LAMMPS (2 Aug 2023) for classical and replica-exchange molecular dynamics, Packmol (v20.14.2) for generation of initial molecular configurations, CP2K (v9.1) for ab initio molecular dynamics and reference calculations, and n2p2 (v2.2.0) for HDNNP training and evaluation.

1.1 Replica-exchange molecular dynamics simulation

Replica exchange molecular dynamics (REMD) simulations were performed using the LAMMPS simulation package (2 Aug 2023) [1]. The all-atom optimized potential for liquid simulations (OPLS-AA) force field [2] was utilized to sample the energy landscape of the organic molecular systems. All force field parameters were taken from OPLS-AA, with the exception of atomic partial charges. Partial atomic charges were assigned following the protocol described by Lin *et al.* [3] Long-range electrostatics were computed using the particle–particle particle–mesh (PPPM) method [4]. Van der Waals interactions were truncated at a cutoff distance of 8.8 Å. A time step of 0.5 fs was employed for integration of the equations of motion. Temperature and pressure were maintained using the Nosé–Hoover thermostat and barostat, respectively.

Initial configurations comprising ten bathocuproine (BCP) molecules were generated using Packmol (v20.14.2) [5]. Following energy minimization of the initial structure, constant-pressure and constant-temperature (NPT) simulations were performed for 10 ns at 300 K and 1 atm for structural relaxation. NVT simulations at densities of 1.1, 1.2, and 1.3 g/cm^3 were conducted to generate initial configurations for REMD.

The REMD simulations employed 40 replicas distributed over temperatures ranging from 300 K to 2000 K. The replica temperatures were not assigned using simple linear or logarithmic spacing. Instead, the temperature ladder was generated using the REMD temperature generator based on the method of Patriksson and van der Spoel [6,7], with the target of achieving an average replica exchange acceptance rate of approximately 30%. Exchange attempts were conducted at 0.1 ps. For each 120 ns REMD simulation at densities of 1.1, 1.2, and 1.3 g/cm^3, the final 100 ns of trajectory were used for sampling. At a density of 1.2 g/cm^3, selected temperature points for representative replicas included 300, 483, 720, 1058, and 1468 K. Replicas at 300 K were additionally included for the systems at 1.1 and 1.3 g/cm^3. From each selected replica, 500 snapshots were extracted, resulting in 3,500 total configurations across all densities and temperatures.

1.2 *ab initio* Molecular Dynamics Simulation

To overcome the intrinsic limitations of the dataset generated via classical REMD, *ab initio* molecular dynamics (AIMD) simulations were performed. AIMD simulations were carried out using the CP2K software package (v9.1) [8]. The Perdew–Burke–Ernzerhof (PBE) exchange–correlation functional [9] was used in combination with the Grimme D3 dispersion correction [10]. Kohn–Sham orbitals were represented using the Gaussian and plane waves (GPW) approach [11] implemented in the Quickstep module of CP2K [12], which utilizes atom-centered Gaussian basis sets for wavefunctions and plane waves for the electron density. Atomic species were modeled using DZVP basis sets [12] in conjunction with Goedecker–Teter–Hutter (GTH) pseudopotentials [13]. A plane wave cutoff energy of 300 Ry was applied for the auxiliary density representation.

Initial configurations for the AIMD simulations were obtained from equilibrated REMD trajectories. AIMD simulations were conducted under the same temperature and density conditions as their corresponding REMD counterparts. A 1 ps NVT simulation was carried out using the Born–Oppenheimer molecular dynamics framework. Temperature was controlled via a canonical sampling velocity rescaling (CSVR) thermostat with a coupling constant of $\tau = 0.1$ ps [14]. The equations of motion were integrated with a timestep of 0.5 fs.

Both initial and final configurations from the AIMD trajectories were incorporated into the training dataset. To enhance the accuracy and convergence of the electronic structure calculations, both initial and final AIMD geometries were recalculated using a higher cutoff energy of 650 Ry and the MOLOPT-DZVP-SR-GTH basis set, replacing the original DZVP set. Because the initial and final configurations from each 1 ps AIMD trajectory were both included in the reference dataset, we examined their configurational distinctness by comparing potential-energy distributions and representative dihedral-angle distributions (Figs. S3–S4). These analyses show that the 0 ps and 1 ps structures are not trivially identical snapshots, but represent meaningfully evolved configurations under finite-temperature dynamics. We note, however, that such decorrelation analysis alone does not fully exclude the possibility of train/test correlation if paired endpoint structures are randomly split across subsets; this possibility was therefore examined explicitly using a trajectory-wise split, as described below.

1.3 Symmetry Functions Selection

Atom-centered symmetry functions (ACSFs) were employed to encode the local chemical environments of individual atoms, enabling the conversion of structural configurations into input representations for atomic neural networks [15,16]. Careful selection of ACSFs is critical for capturing key structural features [17,18] and ensuring that distinct configurations on the potential energy surface (PES) are mapped to unique ACSF descriptors.

Following the procedure described in Ref. [19], an initial general set of ACSFs was constructed with a cutoff radius of 12 Bohr (6.35 Å). Radial symmetry functions were defined for each elemental pair using a fixed Gaussian width of 0.308 Bohr uniformly distributed within the cutoff distance. Angular symmetry functions were assigned to each elemental triplet using a fixed Gaussian width of 0.012 Bohr, with λ values of ±1 and ζ values of 1, 2, 4, 8, 16, and 32 (see Tables S1 and S2 for details).

The initial descriptor set comprised 30 radial and 108 angular functions per element. All ACSF outputs were normalized, and features with low variance (range $< 10^{-2}$) were removed using the n2p2 pruning protocol. A refined ACSF subset was selected based on the chemical diversity present in the training data to optimally represent the OLED system. Feature selection reduced the initial 138 descriptors per element to 113 for H, 111 for C, and 79 for N.

Upon introducing the pseudo-element X, the ACSF pool was expanded to include the additional pair and triplet channels involving X. Importantly, the symmetry functions for X were not re-optimized specifically for the pseudo-element. Instead, the same predefined ACSF parameter pool used for the standard elements (H, C, and N) was extended to the enlarged set of elemental combinations. Accordingly, the same radial and angular parameter grids (i.e., η, R_s, ζ, and λ), as well as the same cutoff radius, were retained, while the interaction channels were expanded to include X–H, X–C, X–N, and X–X terms. This procedure increased the initial ACSF pool from 138 to 232 functions per element.

To maintain computational efficiency while retaining environment-specific information, the expanded descriptor set was again pruned using the same n2p2 protocol. Specifically, descriptors with negligible variation in the training dataset (maximum–minimum $< 10^{-2}$) were discarded. After pruning, 174, 162, 116, and 167 symmetry functions were retained for H, C, N, and X, respectively. The resulting performance improvement therefore does not arise from introducing a new ACSF functional form for X, but from assigning a separate element label—

and hence a separate descriptor-to-neural-network mapping—to a distinct high-error environment class while preserving the same underlying ACSF construction principle as in the baseline model.

1.4 Neural Network Training Set Ups

HDNNP tailored for organic light-emitting diode (OLED) systems was developed using the selected ACSFs and curated training dataset. The HDNNP architecture was based on the Behler–Parrinello framework [15,16], as implemented in the *n2p2* software package (v2.2.0) [20,21]. The network consisted of two hidden layers, each comprising 25 neurons. Hyperbolic tangent (tanh) activation functions were applied to the hidden layers, and a linear activation function was used in the output layer. Network weights were optimized using a recursive Kalman filter algorithm [22].

During training, 100% of the energy data and 1 % of the force components per configuration were utilized. Force components were assigned weights 10-fold larger relative to energy weights. Unless otherwise stated, the reference dataset was divided at the configuration level into 90% training and 10% testing using the internal n2p2 random split determined by the random_seed value specified in input.nn for each model. Because both the initial and final configurations from each 1 ps AIMD trajectory were included in the dataset, we additionally evaluated a stricter trajectory-wise split in which the paired endpoint configurations from each AIMD trajectory were always assigned to the same subset. This additional analysis was performed to explicitly assess the possibility of train/test leakage arising from short-trajectory correlations. The results of this comparison are shown in Fig. S7.

Representative CP2K input files used for the AIMD and endpoint recalculation steps, together with the deposited input.nn files, explicit split artifacts, and trained model files, are provided in the Zenodo repository [https://doi.org/10.5281/zenodo.19078769].

1.5 Validation of Train/Test Split Integrity

To directly assess whether the inclusion of both initial and final configurations from each 1 ps AIMD trajectory could introduce train/test leakage, we performed an additional trajectory-wise split. In this protocol, the paired endpoint structures from each AIMD trajectory were always assigned to the same subset, thereby preventing correlated trajectory endpoints from being divided between the training and test sets.

Figure S7 compares the force and energy RMSE values obtained from repeated random splits and from the trajectory-wise split. The trajectory-wise split yields errors comparable to those obtained from random splits across different seeds, indicating that the predictive performance reported in the main text is not dependent on inadvertently separating paired AIMD endpoint structures across the training and test sets.

1.6 Comparison of Atom-Selection Criteria for Relabeling in the DWCNT

To clarify the atom-selection criterion used for the nanotube example, we directly compared force-error-based and structure-based relabeling in the DWCNT system. In the force-error-based model, the carbon atoms with the highest force RMSE values were reassigned to the pseudo-element X, following the same procedure used for the BCP system. This reduced the carbon and hydrogen force RMSE values from 121.18 and 56.39 meV/Å in the baseline model to 97.54 and 45.35 meV/Å, respectively.

We then examined a structure-based criterion in which the carbon atoms in the outer shell were relabeled. As shown in Fig. S8, this strategy further reduced the carbon and hydrogen force RMSE values to 84.69 and 36.98 meV/Å, respectively. These results confirm that the force-error-based criterion is an effective starting point, while also showing that, for the highly symmetric DWCNT system, the physically defined inner/outer-shell partition provides a cleaner separation of distinct local environments and leads to greater improvement.

1.7 Molecular Dynamics using High-dimensional Neural Network Potential

A committee of four neural networks was constructed to perform molecular dynamics (MD) simulations. For each network in the ensemble, the full reference dataset was randomly subsampled, excluding 10% of the data in each case. Neural network ensembles generally provide improved predictive accuracy over individual models, owing to their ability to reduce overfitting and capture model uncertainty [23,24,25].

MD simulations using the four-network ensemble were carried out with the *CP2K* software package [8]. During simulation, atomic forces and energies were obtained by averaging the predictions of the four neural networks at each timestep. To ensure convergence of the radial distribution function (RDF) and potential of mean force (PMF), simulations were run for 60 ps using the same protocol adopted in the AIMD simulations.

1.8 Glass Transition Temperature Prediction

The element-relabeled HDNNP was applied in MD simulations to estimate the glass transition temperature (Tg). MD simulations were performed using the *LAMMPS* engine interfaced with the *n2p2* package [20]. The protocol proposed by Lin et al. [3] was followed for Tg prediction.

Prior to the application of the HDNNP, classical MD simulations were performed to generate equilibrated structures at 800 K. Simulation parameters and force field settings were consistent with those employed in the REMD simulations. An initial configuration of 128 BCP molecules was generated in a periodic simulation box using *Packmol* [5]. The system underwent staged thermal processing, including heating from 0 K to 300 K, equilibration at 300 K until density convergence, further heating to 800 K at 0.5 K ps^{-1}, and final equilibration at 800 K for 10 ns. Following classical equilibration at 800 K, the system was further equilibrated for 2.5 ns using the element-relabeled HDNNP. Both the classically and HDNNP-equilibrated systems were then cooled linearly from 800 K to 0 K at a rate of 50 K ns^{-1}. During the cooling process, temperature-dependent density profiles were recorded and used to estimate Tg via two-phase linear fitting (Figure 9).

Table S1. Parameter sets for radial symmetry functions used within a cutoff radius of 12 Bohr.

#	η / bohr^{-2}	r_s / bohr
1	0.307795	0.264562
2	0.307795	1.539104
3	0.307795	2.813645
4	0.307795	4.088187
5	0.307795	5.362729
6	0.307795	6.637271
7	0.307795	7.911813
8	0.307795	9.186355
9	0.307795	10.46090
10	0.307795	11.73544

Table S2. Parameter sets for angular symmetry functions used within a cutoff radius of 12 Bohr.

#	η / bohr^{-2}	λ	ζ
1	0.011906	1	1
2	0.011906	-1	1
3	0.011906	1	2
4	0.011906	-1	2
5	0.011906	1	4
6	0.011906	-1	4
7	0.011906	1	8
8	0.011906	-1	8
9	0.011906	1	16
10	0.011906	-1	16
11	0.011906	1	32
12	0.011906	-1	32

Table S3. Summary of descriptor growth and pruning before and after introduction of the pseudo-element X.

Category	**Elemental Composition**	**Total SFs before Pruning (A)**	**Total SFs after Pruning (B)**	**Total Reduction Rate (1-B/A)**
Baseline	H, C, N	414	303	**26.8%**
Proposed	H, C, N, X	928	619	**33.3%**

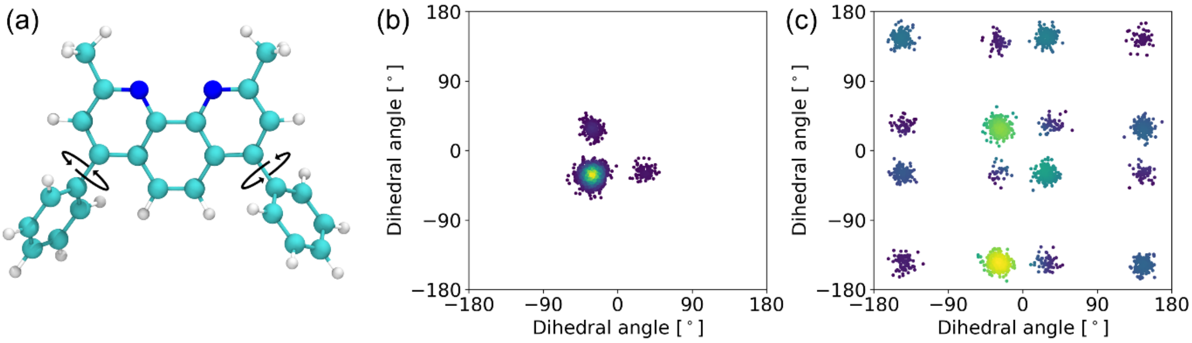

Figure S1. (a) Schematic illustration of two representative dihedral angles in a BCP molecule. (b) and (c) show the corresponding dihedral angle distributions obtained from regular MD and REMD simulations at 300 K, respectively.

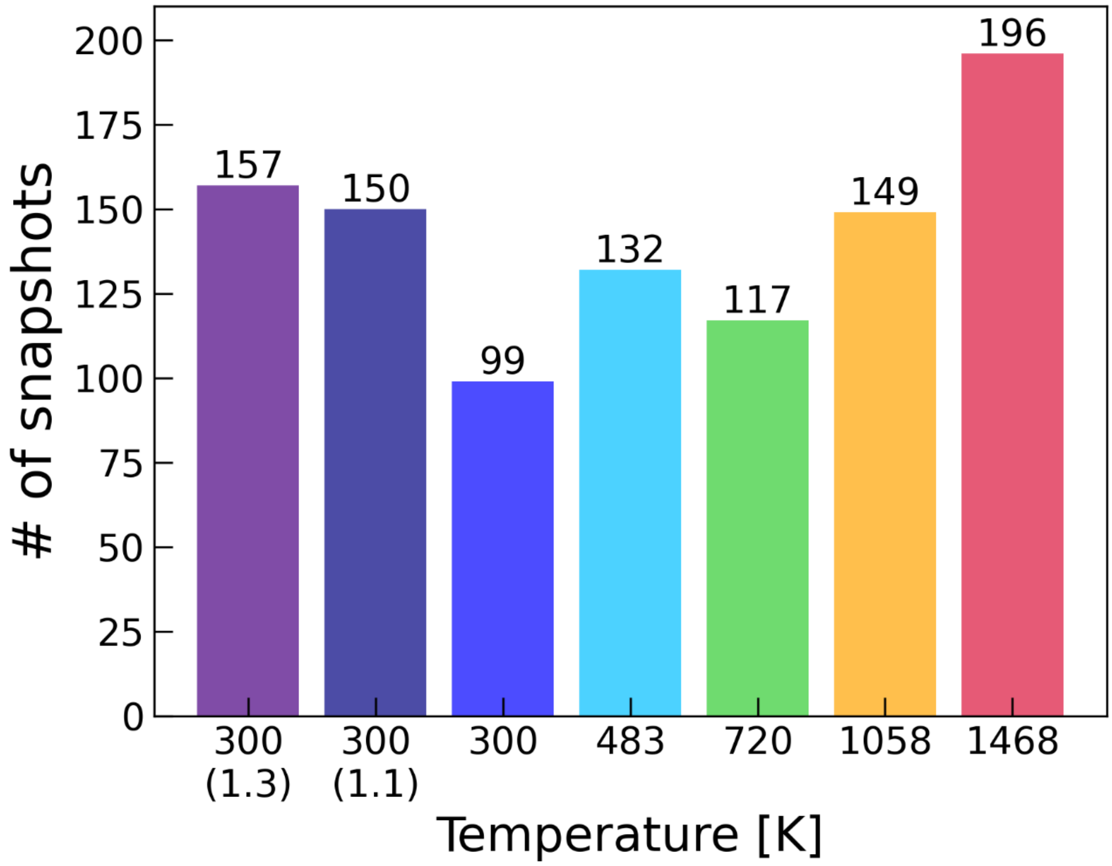

Figure S2. Number of snapshots selected from 3,500 configurations across all replicas using the Farthest Point Sampling (FPS) algorithm.

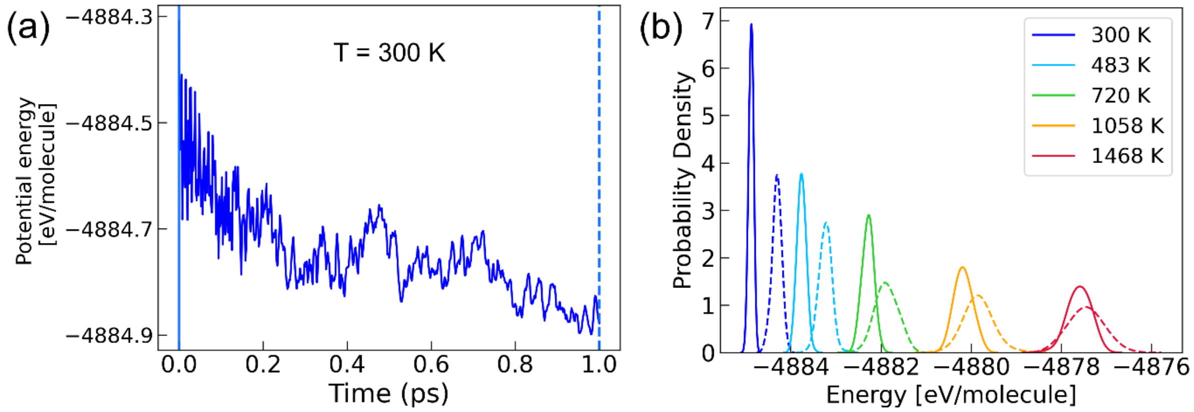

Figure S3. (a) Average potential energy over 1 ps of AIMD at 300 K, with initial structures obtained from REMD. (b) Comparison of initial and final energy distributions at various temperatures, shown as dotted and solid lines, respectively.

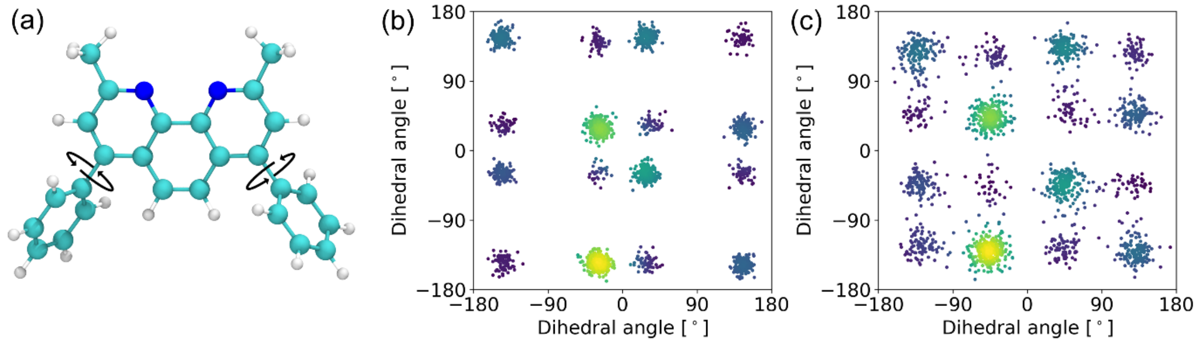

Figure S4. (a) Schematic illustration of two representative dihedral angles in a BCP molecule. Comparison of dihedral angle distributions at 300 K: (b) from REMD and (c) from AIMD after 1 ps, initiated from the REMD configurations.

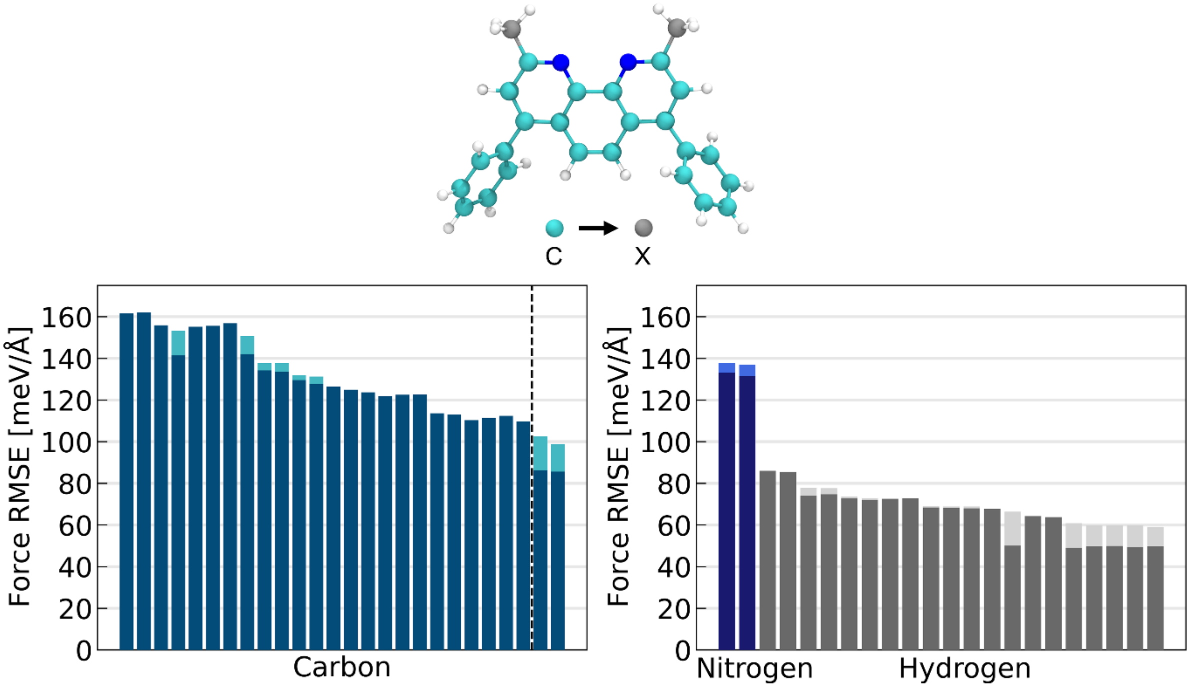

Figure S5. Individual atomic force RMSE variations before and after carbon relabeling with virtual element 'X'. Initial RMSE values, arranged by magnitude, are depicted using bar colors corresponding to the standard atomic colors for carbon, nitrogen, and hydrogen, while dark-colored bars represent the RMSE values after relabeling. In the left panel, a dashed line separates the atoms; those to the right of the line correspond to sites where carbon was relabeled as 'X'.

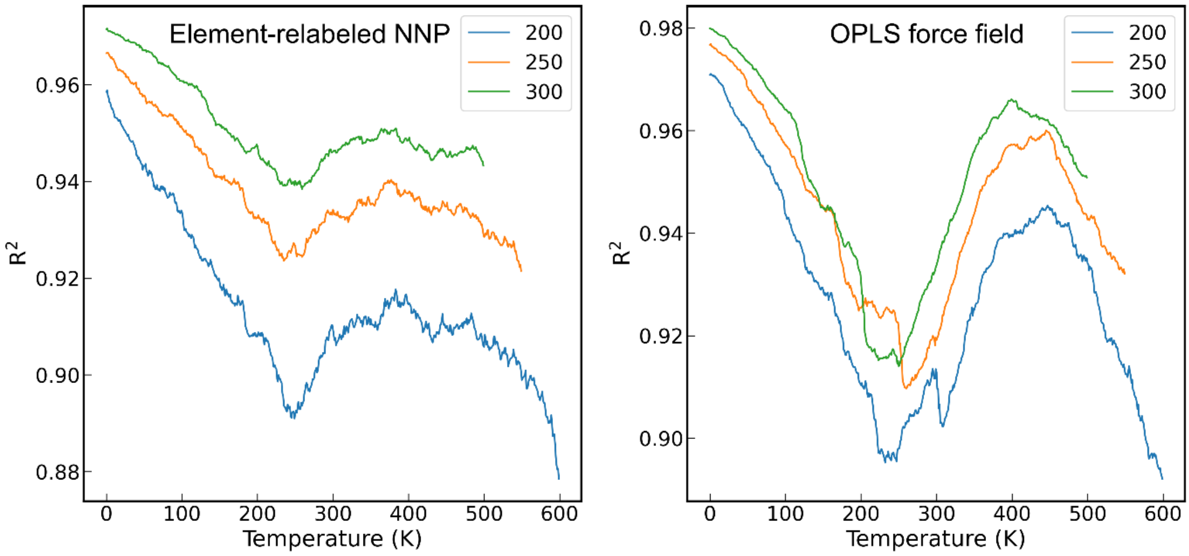

Figure S6. R^2 values for linear regressions of density–temperature data (Fig. 9), computed using sliding windows of 200 K (blue), 250 K (orange), and 300 K (green). Left and right panels represent results from the element-relabeled NNP and the OPLS force field, respectively.

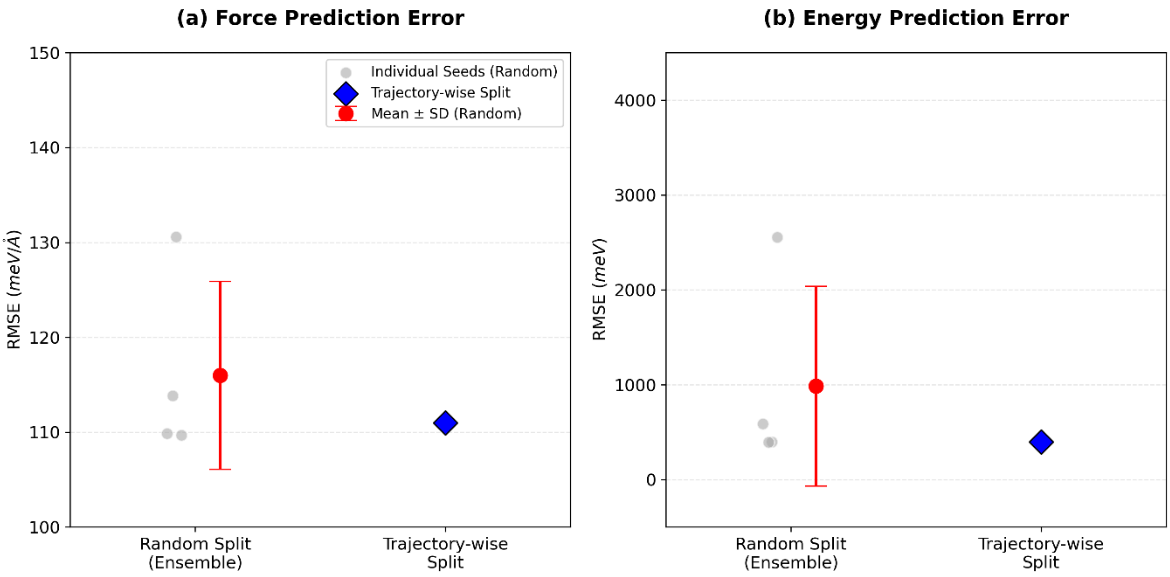

Figure S7. Comparison of model errors obtained using repeated random splits and a trajectory-wise split designed to prevent paired 0 ps/1 ps AIMD endpoint structures from being assigned to different subsets. Gray circles denote results from individual random seeds, red circles and error bars indicate the mean ± standard deviation of the random-split results, and blue diamonds denote the trajectory-wise split result. (a) Force RMSE. (b) Energy RMSE.

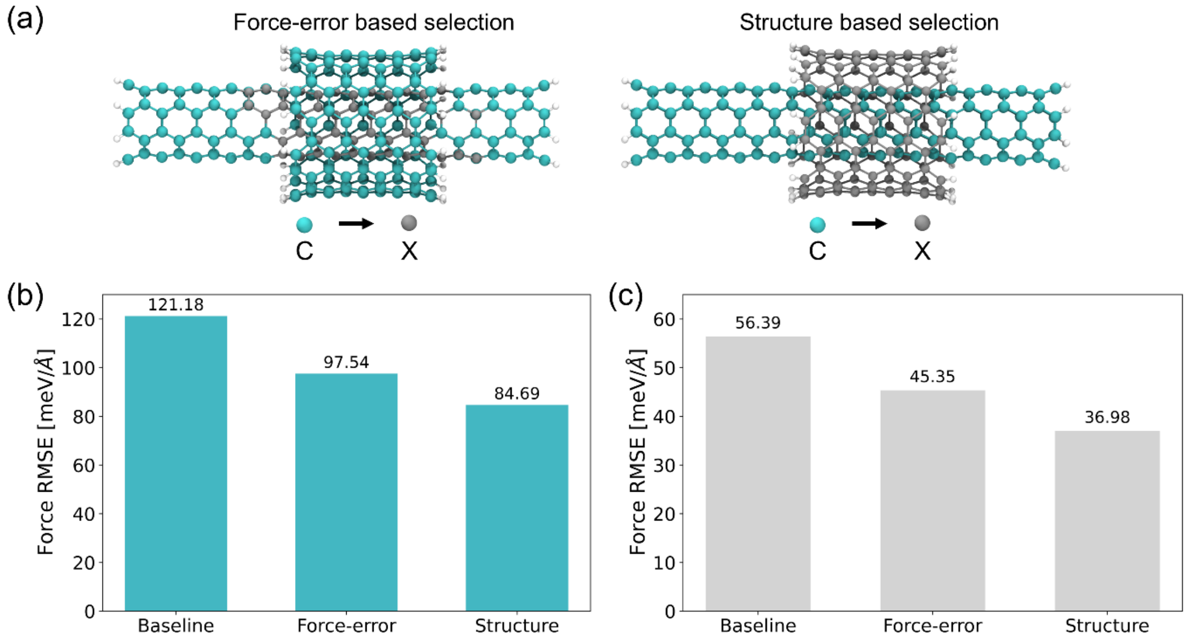

Figure S8. Comparison of atom-selection criteria for relabeling in the DWCNT system. (a) Schematic illustration of the two relabeling strategies: force-error-based selection and structure-based selection (outer-shell relabeling). Relabeled carbon atoms assigned to the pseudo-element X are shown in gray. (b) Carbon force RMSE and (c) hydrogen force RMSE for the baseline, force-error-based, and structure-based models. Both selection criteria improve the prediction accuracy relative to the baseline model, while the structure-based criterion yields the largest reduction in force RMSE.

References

1. A. P. Thompson, H. Aktulga, R. Berger, D. Bolintineanu, W. Brown, P. Crozier, P. Veld, A. Kohlmeyer, S. Moore, T. Nguyen, R. Shan, M. Stevens, J. Tranchida, C. Trott and S. Plimpton, Computer Physics Communications, 2021, **271**, 108171.
2. W. Jorgensen, D. S. Maxwell and J. Tirado-Rives, Journal of the American Chemical Society, 1996, **118** (45), 11225-11236.
3. K. Lin, L. Paterson, F. May, and D. Andrienko, npj Computational Materials, 2021, **7** (1), 179.
4. R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles* (McGraw-Hill, New York, 1981).
5. L. Martínez, R. Andrade, E. G. Birgin and J. M. Martínez, Journal of Computational Chemistry, 2009, **30**, 2157–2164.
6. A. Patriksson and D. Van Der Spoel, Phys. Chem. Chem. Phys., 2008, **10**, 2073.
7. https://virtualchemistry.org/remd-temperature-generator
8. T. D. Kühne, M. Iannuzzi, M. D. Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A. S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack, and J. Hutter, The Journal of Chemical Physics, 2020, **152**, 194103.
9. J. P. Perdew, K. Burke and M. Ernzerhof, Physical Review Letters, 1996, **77**, 3865–3868.
10. S. Grimme, J. Antony, S. Ehrlich and H. Krieg, The Journal of Chemical Physics, 2010, **132**, 154104.
11. G. Lippert, J. Hutter and M. Parrinello, Molecular Physics, 1997, **92**, 477–488.
12. J. VandeVondele and J. Hutter, The Journal of Chemical Physics, 2007, **127**, 114105.
13. S. Goedecker and M. Teter, Physical Review B, 1996, **54**, 1703–1710.
14. G. Bussi, D. Donadio and M. Parrinello, The Journal of Chemical Physics, 2007, **126**, 014101.
15. J. Behler, The Journal of Chemical Physics 2011, **134** (7), 074106.
16. J. Behler, Chemical Reviews, 2021, **121** (16), 10037-10072.
17. G. Imbalzano, A, Abekkum D, Giofre, S. Klees, J. Behler and M. Ceriotti, The Journal of Chemical Physics, 2018, **148**, 241730.
18. F. Musil, F. A. Grisafi, A. Bartok, C. Ortner, G. Csyani and M. Ceriotti, Chemical Reviews, 2021, **121**, 9759–9815.

19. C. Schran, F. Thiemann, P. Rowe, E. A. Müller, O. Marsalek and A. Michaelides, Proceedings of the National Academy of Sciences, 2021, **118** (38), e2110077118.

20. Singraber, A., Behler, J. & Dellago, C. Library-based LAMMPS implementation of high-dimensional neural network potentials. *J. Chem. Theory Comput.* **15**, 1827–1840 (2019).

21. A. Singraber, T. Morawietz, J. Behler, and C. Dellago Journal of Chemical Theory and Computation, 2019, **15** (5), 3075-3092.

22. T. B. Blank and S. Brown, Journal of Chemometrics, 1994, **8**, 391–407.

23. S. Hashem and B. Schmeiser, IEEE Transactions on Neural Networks. 1995, **6**, 792–794.

24. C. Schran, K. Brezina and O. Marsalek, The Journal of Chemical Physics, 2020, **153**, 104105.

25. A. Krogh and J. Vedelsby, Advances in Neural Information Processing Systems, 1995, **7**, 231–238.