

Understanding material properties and performance enabled by molecular simulations
and machine learning potentials

by

Hao Deng

B.S., Hebei University of Technology, 2017
M.S., Tianjin University, 2020

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Tim Taylor Department of Chemical Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2025

Abstract

Machine learning potentials (MLPs) are being rapidly adopted to describe complex potential energy surfaces and to solve emerging materials science and engineering problems. Trained against quantum mechanical datasets, MLPs are versatile mathematical surrogates to describe various chemical and physical phenomena originating from atomistic interactions. MLPs are notably superior over classical interatomic force fields owing to their high training flexibility, efficiency, and accuracy.

In this thesis, artificial neural network (ANN), sparse Gaussian process (SGP), and neural equivariant MLPs were developed and applied in the investigations of mechanical and thermodynamic behaviors of metals, main-group semiconductors, and perovskites. In Chapter 3, a high-dimensional NN potential (HDNNP) was developed for the condensed phase nickel. This HDNNP trained using the geometry and force data extracted directly from *ab initio* molecular dynamics can predict the melting point of face-centered cubic nickel within a few Kelvins of the true value. In Chapter 4, the thermodynamic stabilities of icosahedral boron allotropes, its phase diagram were predicted with a SGP MLP trained using an on-the-fly active learning scheme. In Chapter 5, a neural equivariant interatomic potential was employed to tackle challenges associated with the variations of elemental configurations in a high-performance air electrode perovskite (i.e., $\text{PrNi}_{1-\delta}\text{Co}_\delta\text{O}_3$) for protonic ceramic electrochemical cell (PCEC) applications. MLP-based phonon calculations suggest that the Ni/Co occupancy affects lattice thermal and chemical expansions differently, impacting a tradeoff between PCEC performance and stability. In Chapter 6, the neural equivariant MLP was used to study $\text{BaZr}_{1-\delta}\text{Y}_\delta\text{O}_3$ (BZY), a proton-conducting electrolyte used in PCEC assembly, paving the way to provide solutions to improve scalability of future PCEC manufacturing.

Understanding material properties and performance enabled by molecular simulations
and machine learning potentials

by

Hao Deng

B.S., Hebei University of Technology, 2017
M.S., Tianjin University, 2020

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Tim Taylor Department of Chemical Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2025

Approved by:

Major Professor
Bin Liu

Copyright

© Hao Deng 2025.

Abstract

Machine learning potentials (MLPs) are being rapidly adopted to describe complex potential energy surfaces and to solve emerging materials science and engineering problems. Trained against quantum mechanical datasets, MLPs are versatile mathematical surrogates to describe various chemical and physical phenomena originating from atomistic interactions. MLPs are notably superior over classical interatomic force fields owing to their high training flexibility, efficiency, and accuracy.

In this thesis, artificial neural network (ANN), sparse Gaussian process (SGP), and neural equivariant MLPs were developed and applied in the investigations of mechanical and thermodynamic behaviors of metals, main-group semiconductors, and perovskites. In Chapter 3, a high-dimensional NN potential (HDNNP) was developed for the condensed phase nickel. This HDNNP trained using the geometry and force data extracted directly from *ab initio* molecular dynamics can predict the melting point of face-centered cubic nickel within a few Kelvins of the true value. In Chapter 4, the thermodynamic stabilities of icosahedral boron allotropes, its phase diagram were predicted with a SGP MLP trained using an on-the-fly active learning scheme. In Chapter 5, a neural equivariant interatomic potential was employed to tackle challenges associated with the variations of elemental configurations in a high-performance air electrode perovskite (i.e., $\text{PrNi}_{1-\delta}\text{Co}_\delta\text{O}_3$) for protonic ceramic electrochemical cell (PCEC) applications. MLP-based phonon calculations suggest that the Ni/Co occupancy affects lattice thermal and chemical expansions differently, impacting a tradeoff between PCEC performance and stability. In Chapter 6, the neural equivariant MLP was used to study $\text{BaZr}_{1-\delta}\text{Y}_\delta\text{O}_3$ (BZY), a proton-conducting electrolyte used in PCEC assembly, paving the way to provide solutions to improve scalability of future PCEC manufacturing.

Table of Contents

List of Figures	x
List of Tables	xix
Acknowledgements	xxi
Chapter 1 - Introduction.....	1
1.1 Limitation of Traditional Methodologies	2
1.2 Advantages of Machine Learning Potential.....	3
1.3 Scope of the Dissertation	5
1.3.1 Phase Transition of Condensed Phase Nickel.....	6
1.3.2 Thermodynamic Stability of Icosahedral Boron.....	7
1.3.3 Lattice Expansion of Protonic Ceramic Electrochemical Cell Oxygen Electrode and Electrolyte	10
1.3.3.1 $\text{PrNi}_{1-\delta}\text{Co}_\delta\text{O}_3$ (PNC) Perovskite for PCEC Electrode	12
1.3.3.2 $\text{BaZr}_{1-\delta}\text{Y}_\delta\text{O}_3$ (BZY) Perovskite for PECE Electrolyte.....	13
1.4 Thesis Outline	14
Chapter 2 - Computational Methods.....	16
2.1 Density Functional Theory (DFT)	18
2.1.1 Kohn–Sham Formulation.....	18
2.1.2 Plane Wave Basis Set.....	20
2.1.3 Pseudopotential	20
2.1.4 Electron Exchange-Correlation and Self-Interaction Errors	21
2.2 Machine Learning Potential (MLP)	22
2.2.1 High-Dimensional Neural Network Potential (HDNNP)	23

2.2.2 Sparse Gaussian Process (SGP) Potential.....	27
2.2.3 Allegro – a Neural Equivariant Interatomic Potential	29
2.3 Molecular Dynamics (MD).....	32
2.4 Phonon Calculation.....	33
2.4.1 Vibrational Entropy Calculations	33
2.4.2 Quasi-Harmonic Approximations (QHA).....	34
Chapter 3 - Phase Transition Simulations of Condensed Phase Nickel.....	36
3.1 Introduction.....	36
3.2 Methods	39
3.2.1 Training Data Generation.....	39
3.2.2 NNP Training.....	41
3.2.3 Molecular Dynamics Simulations.....	41
3.3 Results and Discussion	43
3.3.1 Training Results	43
3.3.2 Structure Properties.....	44
3.3.3 Physical Properties	48
3.3.4 Melting and Crystallization of FCC Ni.....	50
3.4 Conclusion	52
Chapter 4 - Thermodynamic Stability Prediction of Icosahedral Boron Allotropes	54
4.1 Introduction.....	54
4.2 Methods	57
4.2.1 Training Set Generation by OTF MD	57
4.2.2 SGP Potential Offline Training.....	60

4.2.3 Molecular Dynamics Simulation	60
4.3 Results and Discussion	61
4.3.1 Thermodynamics and Structural Properties	61
4.3.2 Vibrational Properties	63
4.3.3 Phase Stability Predictions	65
4.4 Conclusion	72
Chapter 5 - Lattice Expansion Investigation of $\text{PrNi}_{1-\delta}\text{Co}_\delta\text{O}_3$ (PNC) Perovskite induced heating and hydration	73
5.1 Introduction	73
5.2 Methods	76
5.2.1 Density Functional Theory (DFT)	77
5.2.2 Dataset Generation	77
5.2.3 Machine Learning Potential (MLP) Training	79
5.2.4 MLP-based Phonon Calculation and Expansion Prediction	80
5.3 Results and Discussion	80
5.3.1 MLP Benchmark	80
5.3.2 Thermal Expansion	81
5.3.3 Chemical Expansion	85
5.3.4 Simultaneous Effect of Ni/Co Occupancy on Expansion and Performance	87
5.4 Conclusion	89
Chapter 6 - Revealing the Lattice expansion and Electrode Incompatibility in Y-doped BaZrO_3 Electrolytes	91
6.1 Introduction	91

6.2 Methods.....	95
6.2.1 Density Functional Theory (DFT)	95
6.2.2 Machine Learning Potential (MLP).....	95
6.2.3 Energy Minimization and Phonon Calculations	98
6.2.4 Material Preparation and Characterizations.....	99
6.3 Results and Discussion	100
6.3.1 Effects of Atomic Environments on Defect Formation and CE	100
6.3.2 Effects of Atomic Environments on TE.....	102
6.3.3 Ordering Screening and Lattice Expansion Investigation of BZY20	105
6.4 Conclusion	109
Chapter 7 - Conclusions and Outlook.....	110
References.....	115
Appendix A - A high-dimensional neural network potential for molecular dynamics simulations of condensed phase nickel and phase transitions.....	147
Appendix B - Predictions of Boron Phase Stability Using an Efficient Bayesian Machine Learning Interatomic Potential.....	151
Appendix C - Atomistic Simulations of Thermal and Chemical Expansions of $\text{PrNi}_x\text{Co}_{1-x}\text{O}_{3-\delta}$ Accelerated by Machine Learning Potentials	164
Appendix D - Uncovering the Atomic Origins of Lattice Expansion and Electrode Incompatibility in Y-doped BaZrO_3 Electrolytes.....	189
Appendix E - Reprint Permissions.....	203

List of Figures

Figure 1.1 Numbers of publications about structurally complex materials in the past decade, , data collected from Web of Science ⁶	1
Figure 1.2 The workflow of (a) physics-based potentials and (b) machine learning potentials, adopted from Ref [⁴⁶]......	5
Figure 1.3 Illustration of the unit cell of FCC nickel.....	6
Figure 1.4 Primitive cells of (A) 12-atom α -B in a rhombohedral cell, (B) 105-atom β -B in a rhombohedral cell, and (C) 28-atom γ -B in an orthorhombic cell. The basic building blocks are displayed in different colors: corner icosahedra (green), edge icosahedra (red), face icosahedra (cyan), interstitial B2 pair (blue), and triple-fused icosahedra (yellow).....	8
Figure 1.5 Phase diagram of elemental boron. Red ⁶¹ , blue ⁶² , and green ⁶³ curves are from different experimental studies, adopted from Ref. [⁵⁹]......	9
Figure 1.6 Schematic of PCECs, a reversible energy storage device, operated in the electrolysis (top) and fuel cell (bottom) modes ⁶⁴	11
Figure 1.7 Schematic diagram of the ideal crystal structure of (A) $\text{PrNi}_{1-\delta}\text{Co}_8\text{O}_3$ (PNC) and (B) $\text{BaZr}_{1-\delta}\text{Y}_\delta\text{O}_3$ (BZY).....	13
Figure 2.1 Illustration of MLP training process, including (1) training set (reference database) generation, (2) choosing the descriptor and representation, and (3) choosing the MLP model and training (regression), adopted from Ref. [⁷⁷]......	16
Figure 2.2 Illustration of the chemical environment. The central atom is marked as blue, and all the atoms within the dashed sphere defined by the cutoff radius are the neighbors of the central atom, adopted from Ref. [²⁶].....	17
Figure 2.3 The self-consistent scheme to solve the Kohn–Sham Eqn., adopted from Ref. [⁷⁹] ...	19

Figure 2.4 (A) Schematic structure of a feed-forward neural network defining a two-dimensional input vector $\mathbf{G} = (G_1, G_2)$ that describes the atomic configuration and the potential energy E . (B) Structure of an artificial neuron i in layer j . Adapted from Ref. [26].	24
Figure 2.5 (A) Radial symmetry functions. The values from dark blue to green are $\eta = 0.0, 0.03, 0.08, 0.16, 0.3$, and $0.5 \text{ \AA}^{-2}$ with $R_s = 0 \text{ \AA}$. (B) Angular Terms with $\lambda = 1$ are shown in blue to black, terms with $\lambda = -1$ are shown in orange to brown. The employed exponents are $\zeta = 1, 2, 4$, and 16 (from light to dark tones).	26
Figure 2.6 Active learning (OTF MD) workflow based on the inherent uncertainty quantification mechanism of the SGP model. Adapted from Ref. [40].	28
Figure 3.1 An FCC 108-atom Ni supercell. The solid black lines represent the cell boundaries.	39
Figure 3.2 A two-dimensional potential energy surface diagram illustrates the T-p parameter space sampling for FCC Ni. The potential energy values (eV/atom) are shown in different colors.	40
Figure 3.3 Initial structures for solid phase (A), liquid phase (B), and phase transition (C) simulations.	42
Figure 3.4 RMSE of energy (A) and forces (B) and comparisons between the training (black) and testing (red) sets at different epochs of training.	44
Figure 3.5 Lattice constants ($\text{\AA}$) predicted using the NNP for 0–800 K. Simulation results are represented by black squares. The error bars are shown in red.	45
Figure 3.6 Thermal expansion coefficients between 200-800 K obtained using NNP and from literature data: red dots (EI+Ph ¹⁴³), blue triangles (EAM ¹⁰⁸), green triangles (Expt. 1: Touloukian ¹⁴⁴), and purple diamonds (Expt. 2: Touloukian ¹⁴²).	46

Figure 3.7 RDFs as a function of temperature predicted using the reported NNP. The peak and valley locations predicted for each temperature using the EAMFS potentials ¹¹⁴ are included for comparison.	47
Figure 3.8 (A) Mean square displacements ($\text{\AA}^2$) of Ni atoms in the liquid phase at various temperatures. (B) Calculated diffusivity ($10^{-9} \text{ m}^2/\text{s}$) in this work (black squares) compared with the literature: red dots (EAMFS ¹¹⁴), blue triangles (EAM ¹⁴⁷), green triangles (Expt. 1: Meyer ¹⁴⁶), and purple diamonds (Expt. 2: Meyer ¹⁴⁵).	49
Figure 3.9 The number of atoms along axial direction and snapshots of the simulated solid-liquid two-phase systems for (A) initial configuration, (B) 1700 K, (C) 1726 K, and (D) 1750 K.	52
Figure 4.1 (A) Workflow for SGP potential training. Boron crystal structures are visualized using the VESTA software ¹⁶⁵ . Black solid lines indicate cell boundaries; (B) Composition of the training set for offline training; (C) DFT- versus SGP-predicted energies; (D) DFT- versus SGP-predicted forces.	59
Figure 4.2 (A) Equation of state for α -B, β -B, and γ -B. Dashed lines (DFT) and solid lines (SGP) were fitted to the 4 th -order Birch-Murnaghan formula. (B) Enthalpy differences (ΔH) for β -B (black) and γ -B (red) versus α -B. (SGP: solid lines; DFT: dashed lines) (C) Radial distribution functions of solid-state α -B, β -B, and γ -B. RDF analyses were performed using OVITO ¹³⁹ . (SGP: black; DFT: red) (D) Volume expansions of α -B (black) and β -B (red). SGP: solid lines with circles; AIMD: dashed lines with squares; experimental volume expansion of α -B ¹⁶⁷ and β -B ¹⁶⁸ : dotted lines with triangles.	62
Figure 4.3 Phonon dispersion and phonon density of states (PDOS) of α -B, β -B, and γ -B calculated from DFT (blue) and MD simulations (red).	64

Figure 4.4 Phase diagram predicted by SGP-based MD simulation and phonon calculations considering (A) defect-free β -phase (β -B ₁₀₅) and (B) β -phase with one B13 defect (β -B ₁₀₄). Squares, dots, diamonds, and triangles represent thermodynamically preferable conditions of α -B, β -B ₁₀₅ , β -B ₁₀₄ , and γ -B, respectively. Red curves and black dashed lines indicate the respective phase boundaries and possible locations of the α - β - γ triple point. (C) Schematic representation of defect-free β -B (B ₁₀₅). Thermodynamic differences (ΔX , $X = H, -TS, ZPE, G$) between B ₁₀₅ and B ₁₀₄ at (D) $T = 500\text{K}$ and (E) $P = 2\text{ GPa}$.	66
Figure 4.5 Thermodynamic differences between β -B and α -B (A) $T = 900\text{ K}$ and (B) $P = 2\text{ GPa}$; between γ -B and α -B (C) $T = 900\text{ K}$ and (D) $P = 2\text{ GPa}$; between β -B and γ -B (E) $T = 900\text{ K}$ and (F) $P = 2\text{ GPa}$.	70
Figure 5.1 The workflow of PNC MLP training and detailed dataset distribution of (B) PNC37, (C) PNC55, and (D) PNC73.	78
Figure 5.2 Thermal expansion (TE) predicted by MLP-based phonon calculations on (A) pristine, (B) OHO $\cdot$ -defective, and (C) VO $\cdot$ -defective PNC.	82
Figure 5.3 (A) The oxygen vacancy formation energies ($\Delta E_{\text{VO}\cdot}$) of PNC37 (grey), PNC55 (red), PNC73 (blue). (B) Thermal expansions from experiments (hollow symbols) and MLP phonon calculations (solid symbols) for $C_{\text{VO}\cdot, \text{PNC37}} = 4.17\%$, $C_{\text{VO}\cdot, \text{PNC55}} = 4.76\%$, and $C_{\text{VO}\cdot, \text{PNC73}} = 8.33\%$. The experimental results were adapted from Ref. ⁷⁰ .	83
Figure 5.4 Chemical expansions (ζ) of PNC37 (black), PNC55 (red), and PNC73 (blue) from MLP phonon calculations for (A) VO $\cdot$ formation ($\zeta_{\text{VO}\cdot}$), (B) OHO $\cdot$ formation ($\zeta_{\text{OHO}\cdot}$), and (C) hydration (ζ_{hyd}).	85
Figure 5.5 Contour plots for lattice expansions of (A) PNC37, (B) PNC55, and (C) PNC73, varying with temperature (x-axis, 330-1000 K) and hydration (y-axis, 0-100%).	87

Figure 6.1 Top view (A-D) and cross-sectional SEM images of PNC73-BZY20 cells treated under (A, E) room temperature (RT) and 0% steam, (B, F) RT and 50% steam, (C, G) 600°C and 0% steam, and (D, H) 600°C and 50% steam. The black region is PNC73, and the gray region with large radius is BZY20. (I) Illustration of expansion mismatch between BZY20 and PNC73 under operando condition. The contribution from chemical expansion is higher than thermal expansion.	93
Figure 6.2 Illustration of four $2\times 2\times 2$ 40-atom supercells including one 1-Y substitution (A, 1Y-sub) and three 2-Y substitutions (B, 2Y-sub-1, C, 2Y-sub-2, and D, 2Y-sub-3) configurations. (E) Schematic illustration of components involved in D_{def-Y} determination as described in Eqn. 6.1. Atoms existing on periodic boundaries are omitted. Distinguishable lattice oxygen sites in supercell are indicated and indexed in the sequence of increasing D_{def-Y}	97
Figure 6.3 Training and benchmarking of BZY-MLP. (A) Training set composition, parity plots and mean absolute error (MAE) of (B) energy and (C) force, darker color represents higher data density region, comparison with DFT on predictions of (D) defect formation energy (ΔE), (E) equation of state (EOS), and (F) BaZrO ₃ phonon density of states (PDOS).	98
Figure 6.4 Relationships between the averaged defect-yttrium distance (D_{def-Y}) and defect formation energy (ΔE_i) or chemical expansion (ϵ_i) in $2\times 2\times 2$ supercells.	101
Figure 6.5 Linear expansion of pristine and V _O -defective (A) 1Y-sub, (B) 2Y-sub-1, (C) 2Y-sub-2, (D) 2Y-sub-3 $2\times 2\times 2$ BZY supercells, and (E) corresponding thermal expansion coefficients (TECs) compared with literature ⁷⁶ which are shown as black dash baselines.	103
Figure 6.6 (A) Three-dimensional colormap surface illustrating energies of 5Y-sub $3\times 3\times 3$ BZY supercells. (B) Thermal expansion coefficients (TECs) of BZY20 ($3\times 3\times 3$ 5Y-sub-1 supercell) predicted by MLP and experiments from this work and literature ⁷⁶ (C) Lattice expansion of	

BZY20 affected by temperature and hydration, calculated based on 5Y-sub-1 supercell. (D)	
Thermal expansion and (E) chemical expansion of 5Y-sub-1 supercell. (F) Lattice expansion mismatch between PNC73 and BZY20.	106
Figure 7.1 Fe(111) morphology: room temperature vs. in operando condition. Top view of final structures obtained from 20-ns molecular dynamics simulations. Bright regions indicate adatoms, while dark regions the formation of vacancies on the surface. Adapted from Ref. [233].	113
Figure 7.2 The trace of a proton in a perovskite showing the two principle features of proton transport: rotational diffusion and proton transfer (data have been obtained by a AIMD simulation of a proton in cubic BaCeO ₃). Adapted from Ref. [198].	114
Appendix Figure A.1 Comparison between energies and forces predicted by NNP and AIMD on training set: (A) energy, (B) x-component force, (C) y-component force, (D) z-component force.	149
Appendix Figure A.2 Comparison between energies and forces predicted by NNP and AIMD on testing set: (A) energy, (B) x-component force, (C) y-component force, (D) z-component force.	150
Appendix Figure B.1 System temperature during OTF MD of β -B with a scaling factor of 1.05. The steps associated with each DFT call are marked with red dots.	155
Appendix Figure B.2 Force mean absolute errors (MAEs) as a function of n_{max} and l_{max} . Cutoff radius = 5.5 Å.	156
Appendix Figure B.3 Force mean absolute errors (MAEs) as a function of the cutoff radii. $n_{\text{max}} = 14$, $l_{\text{max}} = 6$	157

Appendix Figure B.4 Atomic energy mean absolute error (MAE) of every frame during offline training. Structures included in the sparse set are marked with red dots.	158
Appendix Figure B.5 Force mean absolute errors (MAEs) of every frame during offline training. Structures included in the sparse set are marked with red dots.	159
Appendix Figure B.6 Enthalpy differences (ΔH) versus pressure (0~150 GPa) for β -B (black) and γ -B (red) versus α -B. (SGP: solid lines; DFT: dashed lines).....	160
Appendix Figure B.7 Free energy (F), entropy (S), heat capacity (C_v), and enthalpy (E) of (A) α -B, (B) β -B ₁₀₅ , (C) γ -B, and (D) β -B ₁₀₄ based on the SGP potential simulations and DFT using the using the Phonopy code. DFT results are shown in hollow circles, and SGP results are shown in solid lines.....	161
Appendix Figure B.8 (A) Equation of states, (B) atomic enthalpies (eV/atom), (C) radial distribution function $g(r)$, and (D) phonon density of states (PDOS) for β -B ₁₀₄ based on DFT and SGP potential predictions.....	162
Appendix Figure B.9 Principal component analysis (PCA) projection using SOAP descriptors. For each allotrope, 2500 configurations extracted from MD simulations every 0.2 ps under 500 K and 0 Pa were plotted. For building SOAP descriptors, cutoff radius, maximum radial and angular lables are 5.5 Å, 12, and 6, respectively.	163
Appendix Figure C.1 Schematic illustrations of defective PNC structures for MLP-based phonon calculations. Color scheme: Pr (yellow), Ni (green), Co (dark blue), O (red), VO $\cdot$ (cyan), and H (pink). Defects are highlighted with black dashed circles.	176
Appendix Figure C.2 Illustrations of PNCs with different Ni/Co ordering for calculations in Figure S9. Color scheme: Pr (yellow), Ni (green), Co (dark blue), and O (red).....	177
Appendix Figure C.3 Relative energies of PNCs with different Ni/Co ordering.	178

Appendix Figure C.4 Equations of state (EOS) of pristine and defective PNC unit cells, PNC37 (top), PNC55 (middle), and PNC73 (bottom), based on MLP (solid red squares) and DFT (hollow black dots).	179
Appendix Figure C.5 Phonon dispersion and phonon density of states (PDOS) of PNC55 calculated from DFT (blue) and MLP (red) calculations.....	180
Appendix Figure C.6 Equation of state (EOS) of pristine and defective (A) PNC37, (B) PNC55, and (C) PNC73, adjusted with the same volume scaling factor.	181
Appendix Figure C.7 Pressures (P_{total}) and their normal components (P_{xx} , P_{yy} , and P_{zz}) of the PNCs with and without defects under expansion (dL/d_0) of 5×10^{-3} to 20×10^{-3}	182
Appendix Figure C.8 Thermal expansions predicted by MLP (solid symbols, based on unit cell with one $VO^{\bullet}$) and comparisons with experiments ⁷⁰ (hollow symbols).	183
Appendix Figure C.9 (A) $OHO^{\bullet}$ formation energy ($\Delta E_{OHO^{\bullet}}$), and chemical expansion (ζ) induced by (B) $VO^{\bullet}$, (C) $OHO^{\bullet}$, and (D) hydration, for PNC37 (grey), PNC55 (red), and PNC73 (blue), from DFT calculations.....	184
Appendix Figure C.10 Parity plots for the predicted energies and forces based on the testing set (MAE: mean absolute error). Red color indicates the highest density.	185
Appendix Figure C.11 Free energies of (A) pristine PNC37, (B) pristine PNC55, and (C) pristine PNC73, as a function of cell volume at 300-1000K (with a temperature interval of 50K). The red lines indicate the locations of the energy minima.....	186
Appendix Figure C.12 Free energies of PNC consisting of a $VO^{\bullet}$, (A) PNC37, (B) PNC55, and (C) PNC73, as a function of the cell volume at 300-1000K (with a temperature interval of 50K). The red lines indicate the locations of the energy minima.	187

Appendix Figure C.13 Free energies of PNC consisting of a $\text{OHO}\cdot$, (A) PNC37, (B) PNC55, and (C) PNC73, as a function of the cell volume at 300-1000K (with a temperature interval of 50K). The red lines indicate the locations of the energy minima.	188
Appendix Figure D.1 Cross-sectional SEM images with amplify magnitude of 2000 of PNC73-BZY20 cells treated under various conditions, corresponding to Figure 1E to 1H.	196
Appendix Figure D.2 Comparison of the equation of state (EOS) of 2x2x2 BZO and BZY supercells between DFT and MLP.	197
Appendix Figure D.3 SEM images of BZY samples.	198
Appendix Figure D.4 Schematic illustration of determining descriptors including averaged V_O - V_O distance ($D_{\text{V}_\text{O}-\text{V}_\text{O}}$), averaged Y-Y distance ($D_{\text{Y}-\text{Y}}$), and averaged Y coordination (Coord_Y) number. as described in Eqn. 6.2-6.4.	199
Appendix Figure D.5 Three-dimensional colormap surface illustrating energies of 3Y-sub 3x3x3 BZY supercells.	200
Appendix Figure D.6 Scheme of (A) 5Y-sub-1 and (B) 5Y-sub-2 3x3x3 supercells.	201
Appendix Figure D.7 (A) TEC comparison between 5Y-sub-1 and 5Y-sub-2, and (B) Lattice expansion difference between 5Y-sub-1 and 5Y-sub-2.	202

List of Tables

Table 2.1 Comparisons of computational and training range among HDNNP, SGP, and Allegro.	23
Table 3.1 Summary of interatomic potentials for Ni in MD simulations.	37
Table 3.2 Diffusivity (10^{-9} m ² /s) of Ni atoms in the liquid phase.....	50
Appendix Table A.1 Selected temperatures and corresponding scaling factors.....	148
Appendix Table B.1 Parameters used in OTF and offline training	153
Appendix Table B.2 4 th -order Birch-Murnaghan (BM) Eqn. of state (EOS) ¹⁶⁶ coefficients calculated from DFT and MLP results. And the mean absolute error (MAE) and mean absolute difference (MAD) of MLP predictions.....	154
Appendix Table C.1 Lattice parameters (Å) of optimized pristine and defective PNC unit cells.	164
Appendix Table C.2 Bader charges (Q_B , e) of atoms in pristine and VO $\cdot$ -defective PNC37 and the charge change (ΔQ_B) due to the defect formation.....	165
Appendix Table C.3 Bader charge (Q_B , e) of atoms in pristine and OHO $\cdot$ -defective PNC37 and the charge change (ΔQ_B) due to the defect formation.....	166
Appendix Table C.4 Bader charge (Q_B , e) of atoms in pristine and VO $\cdot$ -defective PNC55 and the charge change (ΔQ_B) due to the defect formation.....	167
Appendix Table C.5 Bader charge (Q_B , e) of atoms in pristine and OHO $\cdot$ -defective PNC55 and the charge change (ΔQ_B) due to the defect formation.....	168

Appendix Table C.6 Bader charge (Q_B , e) of atoms in pristine and VO $\cdot$ -defective PNC73 and the charge change (ΔQ_B) due to the defect formation.....	169
Appendix Table C.7 Bader charge (Q_B , e) of atoms in pristine and OHO $\cdot$ -defective PNC73 and the charge change (ΔQ_B) due to the vacancy formation.	170
Appendix Table C.8 Parameters of two-dimensional polynomial fitting of the relationship between expansion (dL/d $_0$) and co-effect of temperature (T) and hydration (H), as shown in Figure 5. The R 2 values of each fitting are shown in the last row.....	171
Appendix Table C.9 Comparisons between GGA-PBE and r 2 SCAN+rVV10 functionals on energies (in eV) of PNC73 and corresponding VO $\cdot$ formation.....	172
Appendix Table C.10 Parameters used in OTF-MD for dataset generation with FLARE	173
Appendix Table C.11 Parameters used in final offline training using Allegro	174
Appendix Table C.12 Wall time (in second) comparison between DFT and MLP on single-point energy calculation and geometry optimization on PNC73 unit cell.	175
Appendix Table D.1 OTF-MD parameters for dataset generation based on 2x2x2 supercell....	192
Appendix Table D.2 Parameters used in final offline training using Allegro	193
Appendix Table D.3 Lattice constants of pristine and V $_O$ -defective 2 $\times$ 2 $\times$ 2 supercells, and formation energy (ΔE_i) or chemical expansion (ε_i) due to the defect i (V $_O$ or OH) formation at distinct sites of 1Y-sub and 2Y-sub BZY, as shown in Figure 6.4.	194
Appendix Table D.4 Linear fitting parameters of relationships between averaged oxygen-yttrium distance (D_{O-Y}) and defect formation energy (ΔE) or chemical expansion (ε) in 2 $\times$ 2 $\times$ 2 supercells, as shown in Figure 6.4.	195

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisor, Prof. Bin Liu, for his unwavering support and invaluable guidance throughout my Ph.D. journey. He not only taught me how to conduct rigorous research but also inspired me to grow as a thoughtful and independent researcher. His mentorship extended beyond academics, offering insightful advice on career development and life planning. The time I spent under his supervision has been truly transformative and will remain invaluable throughout my life.

I am also deeply grateful to my committee members—Prof. James Edgar, Prof. Jeffrey Comer, and Prof. Davood Pourkargar—for their time, feedback, and continued support during my doctoral studies. Special thanks to Prof. Chuancheng Duan for the stimulating discussions and the opportunity to collaborate on his project.

I would like to acknowledge Prof. Liu's former and current group members—Dr. Jiayi Xu, Dr. Narges Manavi, Dakota Braaten, Yao Du, Zixian Wang, and Nabanita Mondal—for their helpful discussions, creative ideas, and collaborative spirit.

My heartfelt appreciation also goes to Prof. Christine Aikens and her group members—Dr. Zhen Liu, Dr. Yuchen Wang, Dr. Olivia Hull, Dr. Shana Havenridge, Sulalith Samarasinghe, Dmitri Leenatali, Gayathri Habarakadage, and Paul Javed—for the engaging joint group meetings and enlightening scientific conversations.

The final two years of my Ph.D. research were conducted at Idaho National Laboratory under the supervision of Dr. Dong Ding. I am truly thankful for Dr. Ding's mentorship, encouragement, and generous support during this critical phase of my research. I am equally appreciative of his current and former group members—Dr. Meng Li, Dr. Wei Wu, Dr. Lucun Wang, Dr. Qian Zhang, Dr. Wenjuan Bian, Dr. Zeyu Zhao, Dr. Yuqing Meng, Dr. Wanhua Wang,

Dr. Haixia Li, Dr. Fan Liu, Dr. Samuel Koomson, Dr. Pengxi Zhu, Dr. Weilin Zhang, Quanwen Sun, Yuchen Zhang, Matthew Naughton, and Linan Zhu—for their support both in research and daily life.

I am also thankful for the friends I met during this journey—Bolun Xu and his wife Jia Tan, Fengqing Di, Armin Ezzati, Xiaoting Xu, Yizhi Wang, Sunming Qin, and Yu Lu. Thank you for sharing countless memorable moments and experiences with me.

Finally, I would like to express my deepest gratitude to my parents and family. Your understanding, love, and unwavering support have been the foundation that made this journey possible. I could not have done it without you.

Chapter 1 - Introduction

The demand for highly functional materials grows rapidly to meet the advancement of technologies in energy conversions. In the past decade, the structural and compositional complexity of synthesized materials has increased significantly, driven by the pursuit of higher performance. As shown in Figure 1.1, prominent materials include metal-organic frameworks (MOFs)¹, 2D materials², multi-dopant perovskites³, high-entropy alloys⁴, and liquid electrolytes⁵. Improving materials' intrinsic properties (e.g., hardness, expansion, diffusivity, stability) and performance (e.g., catalysis, absorption) is a perpetual goal of worldwide material researchers. Tuning properties of these materials relies on a deep understanding of the mechanisms governing their observable macroscopic behaviors.

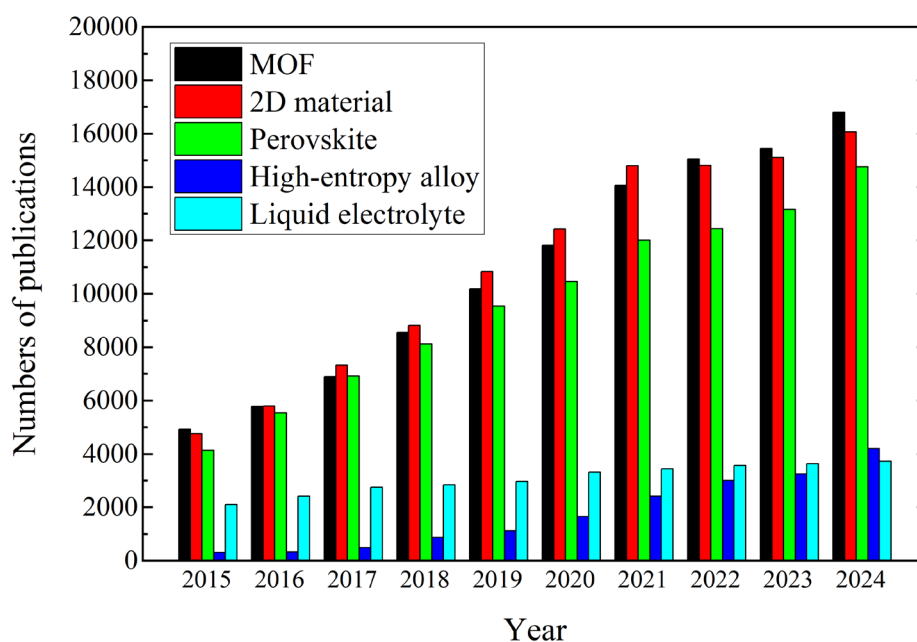


Figure 1.1 Numbers of publications about structurally complex materials in the past decade, , data collected from Web of Science⁶.

Computational methods are expected to play an indispensable role in material research by providing atomically resolved insights due to restricted space and time scale, temperature, atmosphere in commonly used experimental techniques. Since the first atomic-scale simulation was reported in 1957⁷, various methodologies have been developed and applied in fundamental research. Mature computational techniques include density functional theory (DFT)^{8,9}, Hartree-Fock (HF)¹⁰, molecular dynamics (MD)¹¹, and Monte Carlo (MC)¹². However, as the materials of interest become more complex, limitations on simulation scale and accuracy become evident. Even though computational resources are also growing rapidly, modelers are still struggling to balance the efficiency and fidelity of molecular simulations. Thus, it is vital for modern material research to identify new modeling paradigm to strengthen atomistic simulations with high predictive power at reduced computational costs. The machine learning potentials (MLPs) emerge as a promising option.

1.1 Limitation of Traditional Methodologies

For condensed phase material research, the two most popular simulation methodologies are DFT^{8,9} and MD. Empowered by continuous improvement of computing hardware, DFT has become perhaps one of the most widely used modeling tools in the past twenty years. Founded on quantum mechanical principles¹³, DFT calculations provide rich information in molecular configurations, ground-state energies, electronic band structures, densities of states (DOS), and valence charges, and magnetisms. Applications of DFT methods have surged in chemistry¹⁴, condense matter physics¹⁵, biology¹⁶, catalysis¹⁷, material science¹⁸, and artificial intelligence¹⁹.

In classical MD simulations, individual atomic trajectories are dictated by empirical interatomic potentials according to the Newtonian law of motion²⁰. MD simulations have long been used to probe systems and phenomena over the dimension and time scales beyond the range of typical quantum mechanical methods. It works best for studying atomic-scale dynamics, thermodynamics, and transport properties over time. Nevertheless, DFT and classic MD both have intrinsic limitations.

DFT and *ab initio* molecular dynamics (AIMD) suffer from high computational expenses, which lead to poor scaling when monitoring the behaviors of large model systems over a long timescale. Moreover, even routine structural or compositional complexity will likely pose computational challenges to efficiently investigate materials' thermodynamic stabilities (especially phonon calculations), phase transitions, and the migrations of atomic species or charge carriers.

Since the first MD study based on empirical potential was published in 1977,²¹ a plethora of potentials emerged, such as Lennard-Jones potential,²² embedded atom method (EAM),²³ and reactive force field (ReaxFF),²⁴ which are all based on physical rules to describe interatomic relationships. However, the acquisition of these potentials can be highly strenuous due to the delicate training of force field parameters bearing specific physical meanings. Often, the predictability deteriorates notably for complex, multi-component systems.

1.2 Advantages of Machine Learning Potential

In recent years, there has been a surging interest in machine learning potentials (MLPs) as the mathematical surrogates to describe complex atomic interactions owing to their training

flexibility, high computational efficiency, and accuracy when implemented in molecular simulations. Traditional interatomic potentials are typically constructed based on physical principles. As illustrated in Figure 1.2A, these physics-based models calculate the energy of each atom by considering the positions of neighboring atoms within a predefined cutoff radius, using fixed analytical expressions with preset parameters. The system's total energy is obtained by summing over the individual atomic energies. In contrast, MLPs offer a more flexible approach. As shown in Figure 1.2B, an atom's local environment (defined by the atoms within a cutoff sphere) is encoded into a set of structural descriptors. These descriptors are then mapped to atomic energies through a regression model. Summing over the atomic contributions yields the total energy of the system, effectively defining a point on the potential energy surface (PES), which represents the energy–structure relationship. Mathematical nature enables MLP to break the limitations of physical rules. In principle, flexible models like artificial neural networks (ANNs) can reach arbitrarily high accuracy. By training against datasets extracted from *ab initio* calculations, MLP can be delicately fitted to build the potential energy of the system; then, PES can be established.

We are witnessing a surge in the interest in MLPs because high-quality computational data sets (generated from DFT calculations) have become more accessible. Thus far, the relatively mature, more established MLPs include neural network potential (NNP)²⁵⁻²⁸, spectral neighbor analysis potential (SNAP)²⁹⁻³², Gaussian approximation potential (GAP)³³⁻³⁵, and moment tensor potential (MTP)³⁶⁻³⁸, sparse Gaussian process (SGP)³⁹⁻⁴², and neural equivariant interatomic potential (NequIP)^{43,44}. With the help of MLP, much more complex simulations can be performed, including larger scale MD simulations and phonon calculations. Notable examples include the

modeling of inter-dimensional transformation of 2D stanene ⁴⁰, and heterogeneous H₂/Pt(111) system (over 1 billion atoms) involving chemisorption of H₂ molecules⁴⁵.

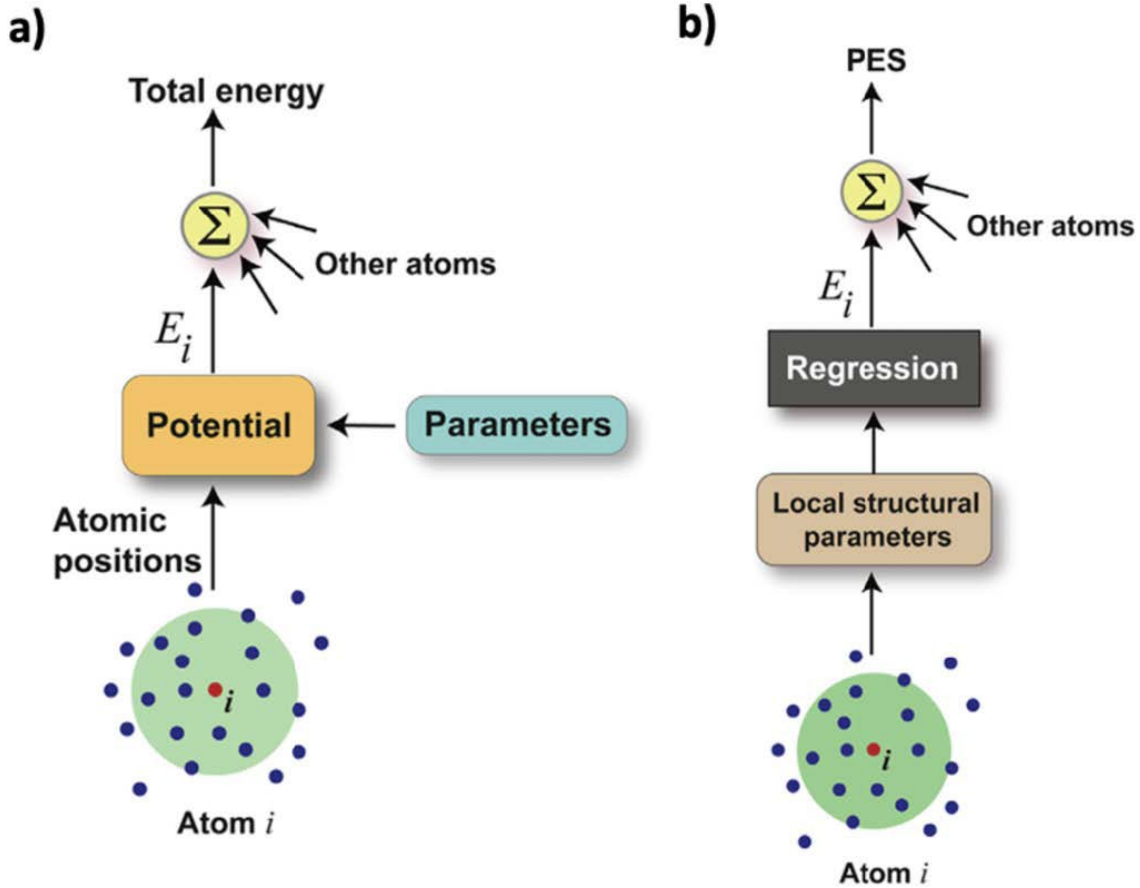


Figure 1.2 The workflow of (a) physics-based potentials and (b) machine learning potentials, adopted from Ref [46].

1.3 Scope of the Dissertation

As the capability of MLPs grows, there are convincing evidence suggesting that MLPs are applicable in simulations of complex material systems for energy applications ⁴⁷⁻⁴⁹. This thesis will discuss several projects involving three different MLP methodologies on these systems: condensed

phase nickel, icosahedral boron, and perovskites for proton-conducting electrochemical cells. The three systems cover transition metal, main group semiconductors, and perovskites, and their research interests will be discussed in the following sections. This thesis also demonstrates that MLPs are powerful and versatile tools in investigating functional materials' physical and chemical phenomena for energy applications.

1.3.1 Phase Transition of Condensed Phase Nickel

The phase transition of Ni has been subjected to extensive investigations ⁵⁰⁻⁵². For this reason, abundant data have been accumulated over the years in literature, providing an ideal opportunity to benchmark the quality of interatomic force fields and other simulation tools. Nickel exhibits a face-centered cubic (FCC) structure under room temperature and atmosphere pressure, as shown in Figure 1.3. Moreover, nickel has been regarded as a versatile metal with influential impacts on steel manufacturing ⁵³, catalysis ^{54, 55}, and various chemicals and energy conversions ^{56, 57}.

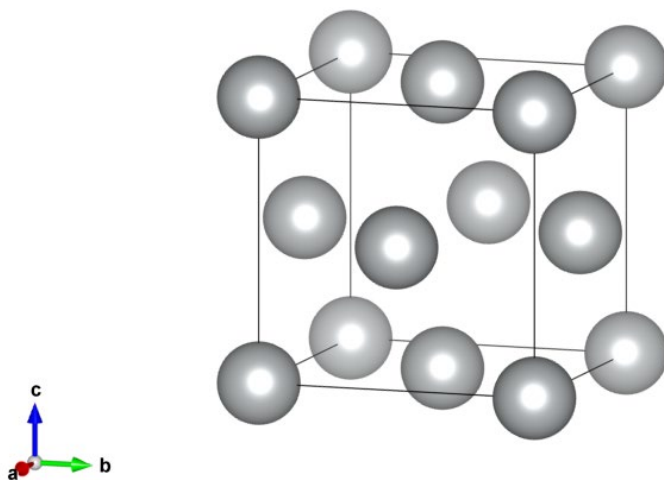


Figure 1.3 Illustration of the unit cell of FCC nickel.

In Chapter 3, a high-dimensional neural network potential (HDNNP) is reported for MD simulations of condensed-phase Ni and Ni systems with liquid-solid phase coexistence. The reference data set was generated to train this HDNNP by sampling the potential energy surface using *ab initio* MD simulations over a broad temperature-pressure domain. Excellent agreements were achieved between bulk FCC nickel thermal expansion simulations and experimental data. This HDNNP also yields accurate structures and diffusivities in the liquid state. The phase transition between the liquid and solid phases was simulated using the two-phase interface method. The predicted melting point temperature is within a few kelvins of the literature value. The general methodology could be applied to describe crystals with much more complex phase behaviors.

1.3.2 Thermodynamic Stability of Icosahedral Boron

Elemental boron is characterized as low weight (2.3g/cm^3), high melting point (2200°C), and high degree of hardness (9.3 on Mohs' scale). Stable pure boron crystals contain characteristic icosahedral B₁₂ motifs. Among boron's allotropes⁵⁸, α -B (rhombohedral, B₁₂), β -B (rhombohedral, B₁₀₅), and γ -B (orthorhombic, B₂₈) are the three principal members, which all consist of B₁₂ icosahedra as the fundamental building block (Figure 1.4).

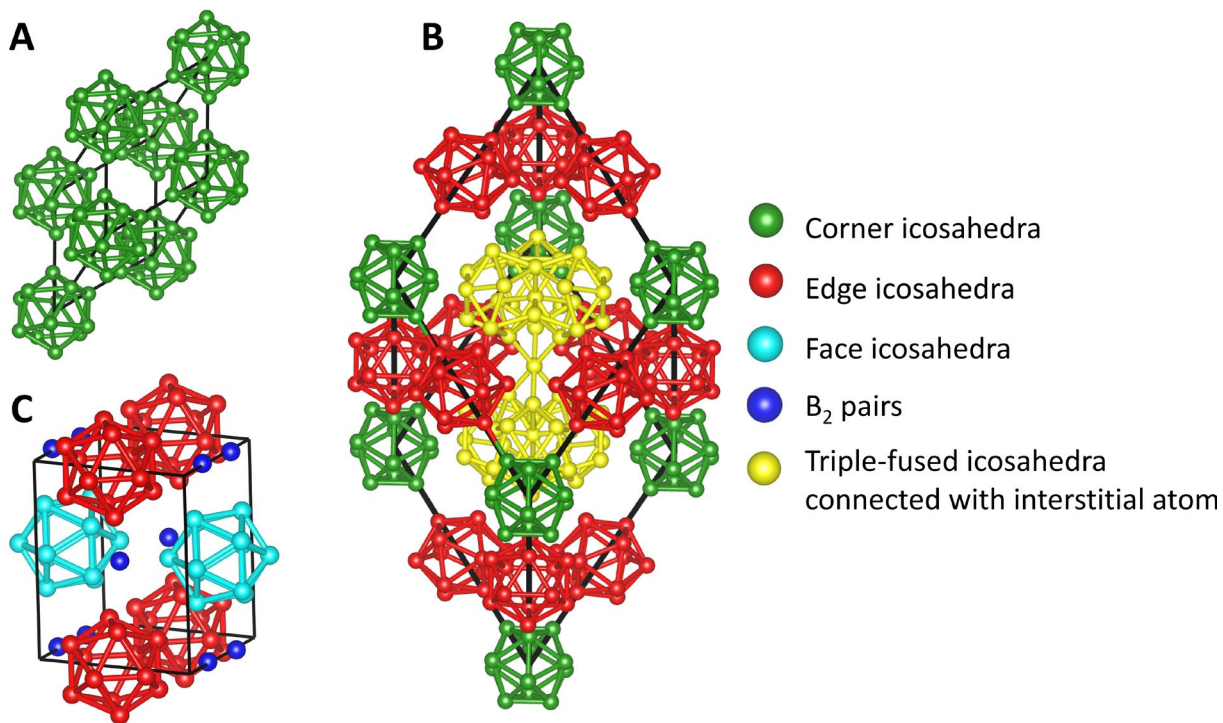


Figure 1.4 Primitive cells of (A) 12-atom α -B in a rhombohedral cell, (B) 105-atom β -B in a rhombohedral cell, and (C) 28-atom γ -B in an orthorhombic cell. The basic building blocks are displayed in different colors: corner icosahedra (green), edge icosahedra (red), face icosahedra (cyan), interstitial B₂ pair (blue), and triple-fused icosahedra (yellow).

The thermodynamic stability of elemental boron allotropes remains debated, as shown in Figure 1.5. While the γ -B phase dominates at high pressures, the relative phase stability between α -B and β -B at moderate conditions remains debated⁵⁹. Moreover, the experimental determination of phase stability between α -B and β -B is challenged by a substantial phase transition kinetic barrier⁶⁰.

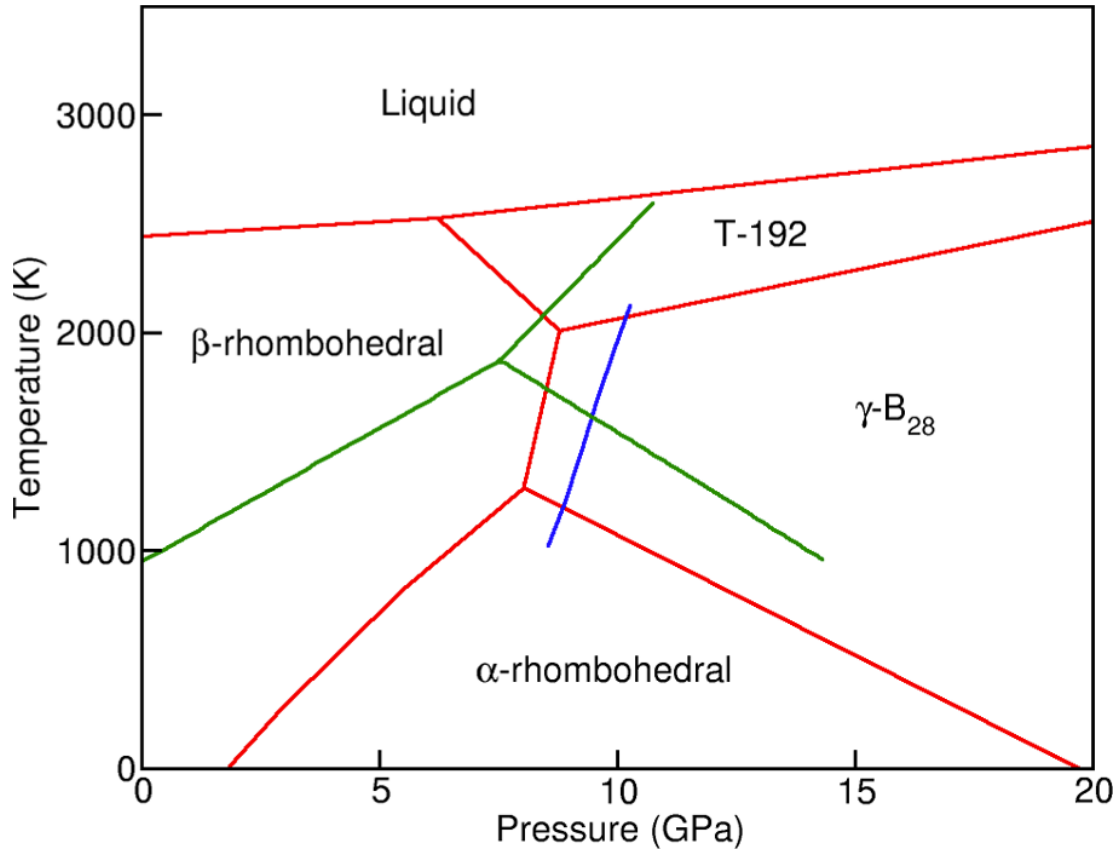


Figure 1.5 Phase diagram of elemental boron. Red ⁶¹, blue ⁶², and green ⁶³ curves are from different experimental studies, adopted from Ref. [⁵⁹].

The B12 motif presents a unique challenge for an MLP to adapt to various icosahedral boron allotropes. Hence, the SGP approach was selected, which is based on on-the-fly active learning and improves the efficiency of training set generation. Thermodynamic phase stability of three elemental boron allotropes, i.e., α -B, β -B, and γ -B, was investigated using a Bayesian interatomic potential trained via a sparse Gaussian process (SGP). SGP potentials trained with datasets from on-the-fly active learning achieve quantum mechanical level accuracy when employed in molecular dynamics simulations to predict wide-ranging thermodynamic, structural, and vibrational properties. The simulated phase diagram (500~1400 K and 0~16 GPa) is in good

agreement with experimental measurements. The SGP-based MD simulations also successfully predicted that the B13 defect is critical in stabilizing β -B below 700 K. At higher temperatures, the entropy becomes the dominant factor, making β -B the more stable phase over α -B. This study demonstrates that SGP potentials based on a training set consisting of defect-free-only systems could make correct predictions of defect-related phenomena in solid-state crystals, paving the path to investigate crystal phase stability and transitions.

1.3.3 Lattice Expansion of Protonic Ceramic Electrochemical Cell Oxygen Electrode and Electrolyte

Protonic ceramic electrochemical cells (PCECs) are a next-generation energy conversion device. As shown in Figure 1.6 ⁶⁴, PCECs are assembled into a planar sandwich structure, consisting of an oxygen electrode (also called steam electrode or cathode), electrolyte, and a hydrogen electrode (also called fuel electrode or anode). The protonic electrolyte in PCEC conducts protons, which can be performed under reversible modes and enable PCEC to be functional under both electrolysis and fuel cell modes. In electrolysis mode, hydrogen evolution reaction (HER) and oxygen evolution reaction (OER) occur at the anode and cathode, respectively. When switching to the fuel cell mode, reactions become hydrogen oxidation reaction (HOR) and oxygen reduction reaction (ORR).

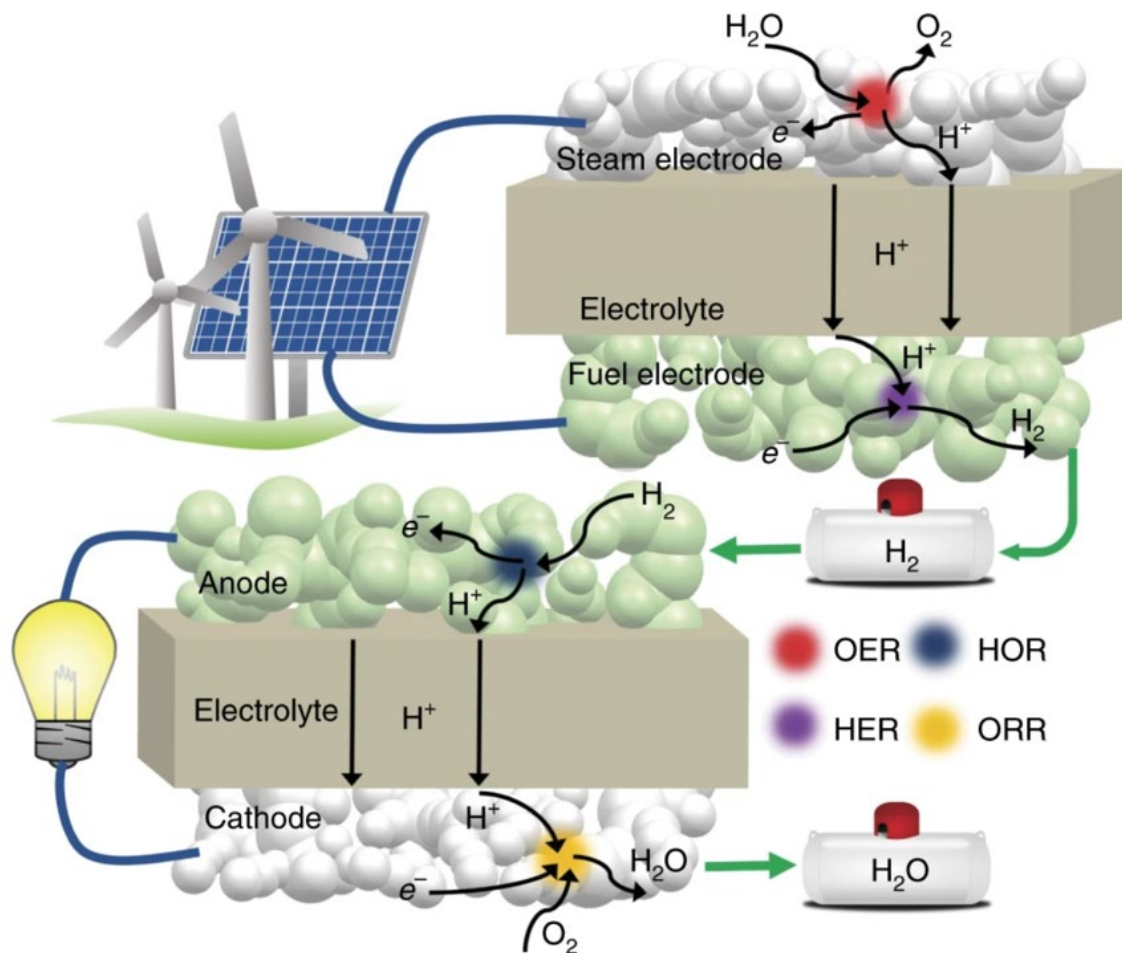


Figure 1.6 Schematic of PCECs, a reversible energy storage device, operated in the electrolysis (top) and fuel cell (bottom) modes ⁶⁴.

Electrodes and solid-state electrolytes in PCECs experience significant lattice expansions when exposed to high steam concentrations at elevated temperatures, especially under the electrolysis mode ^{65, 66}. The expansion mismatch between electrode and electrolyte results in interfacial delamination and causes rapid deterioration in PCEC performance ^{67 68}. Accurate and fast predictions of volume expansions are pivotal to material selection and cell manufacturing with sustained electrode-electrolyte interfacial stability.

Under *operando* conditions, lattice expansions of PCEC components are simultaneously contributed by thermal expansion (TE) and hydration-induced chemical expansion (CE). Experimental measurements of decoupled TE and CE under *operando* conditions can be challenging, because TE and CE are often coupled. Computationally, phonon calculations based on quasiharmonic approximations (QHA) offer pathways to measure and decouple TE and CE. In this thesis, phonon calculations based on neural equivariant MLP were conducted to elucidate the volume expansions of the proton-conducting $\text{PrNi}_x\text{Co}_{1-x}\text{O}_{3-\delta}$ (PNC) and $\text{BaZr}_x\text{Y}_{1-x}\text{O}_{3-\delta}$ (BZY) lattices, which are potential candidates of oxygen electrode and electrolyte, respectively. The lattice expansion was manifested under a combined influence of temperature, oxygen vacancies ($V_{\text{O}}^{\bullet\bullet}$) and proton uptake (OH_{O}^+) in the bulk at varying B-site occupancies.

1.3.3.1 $\text{PrNi}_{1-\delta}\text{Co}_{\delta}\text{O}_3$ (PNC) Perovskite for PCEC Electrode

$\text{PrNi}_{1-\delta}\text{Co}_{\delta}\text{O}_3$ (PNC) is a triple conducting oxide and exhibits promising performance as an oxygen electrode material^{69, 70}. As shown in Figure 1.7a, PNC displays an ABO_3 orthorhombic perovskite structure, where Pr occupies the A site, and Ni, Co occupy the B site.

In Chapter 5, it is shown that the Ni/Co occupancy contributes to thermal and chemical expansions of PNC differently, where thermal expansions are related to Co occupancy. In contrast, chemical expansions are more closely associated with the Ni occupancy. Both $V_{\text{O}}^{\bullet\bullet}$ and OH_{O}^+ lead to higher thermal expansion than the pristine PNC. The temperature increase will produce a negative impact on hydration-induced chemical expansions. For the combined thermal and chemical expansions, it was found that, at a higher Ni occupancy, the strategies that boost the PCEC's electrochemical performance may harm the electrode-electrolyte interfacial stability

caused by severe chemical expansions. Mitigating chemical expansions of the Ni-abundant PNC will benefit the interfacial stability.

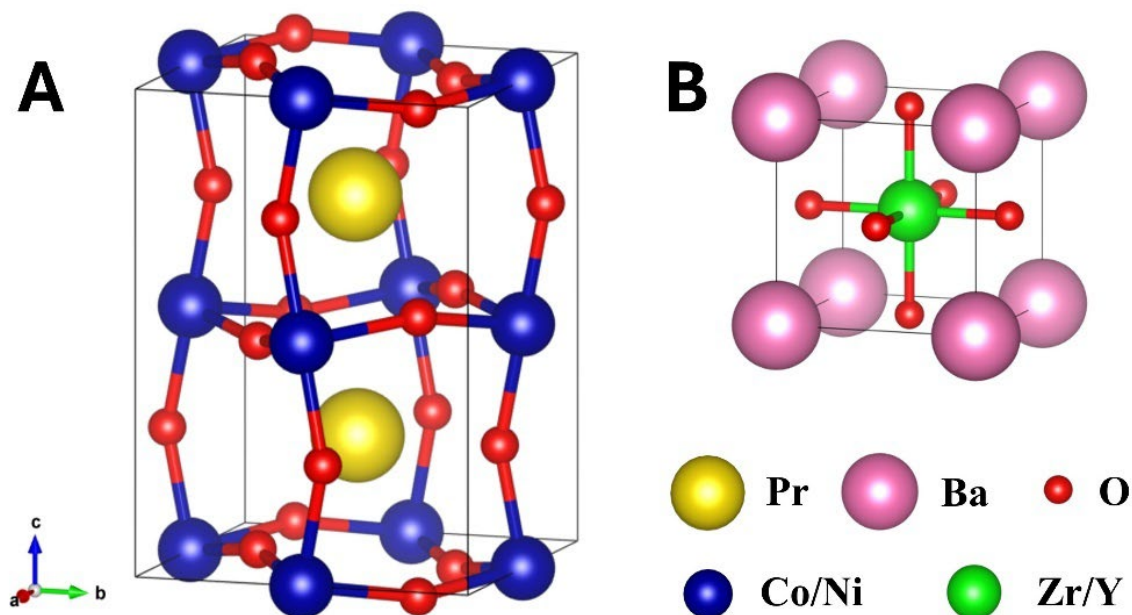


Figure 1.7 Schematic diagram of the ideal crystal structure of (A) $\text{PrNi}_{1-\delta}\text{Co}_\delta\text{O}_3$ (PNC) and (B) $\text{BaZr}_{1-\delta}\text{Y}_\delta\text{O}_3$ (BZY).

1.3.3.2 $\text{BaZr}_{1-\delta}\text{Y}_\delta\text{O}_3$ (BZY) Perovskite for PECE Electrolyte

Among proton-conducting electrolyte materials, doped- BaZrO_3 and doped- BaCeO_3 have been extensively studied. Doped- BaZrO_3 oxides have better mechanical and chemical stability than doped- BaCeO_3 oxide in terms of stability⁷¹. Y-doped BaZrO_3 (BZY) perovskite is one of the earliest identified compositions with promising performance⁷²⁻⁷⁵. As shown in Figure 1.7b, BZY exists as cubic perovskite structure, where Ba and Zr/Y occupy A and B sites, respectively. The replacement of Zr^{4+} with trivalent Y^{3+} on B-site leads to V_O formation in the lattice, thus water

molecules can be more easily incorporated into the lattice, leading to enhanced hydration and higher protonic conductivity.

The lattice expansion of popular electrolytes is typically less than electrodes ⁶⁵. The differences between PCEC electrolyte and electrode on hydration-induced chemical expansion also attract an increasing attention, since potential risk of delamination ⁷⁶. To comprehensively quantify the lattice expansion mismatch and identify its origin, we conducted molecular simulations based on DFT and MLP to identify the popular atomic environments in BZY20 and their contribution to lattice expansion. We found that $V_{\text{O}}^{\bullet}$ and $\text{OH}_{\text{O}}^{\bullet}$ tend to form closely with Y atoms, and the resulted CE is more significant than defects far from Y atoms. The main contributor of CE is filling $V_{\text{O}}^{\bullet}$. When the concentration of Y dopant is high ($\sim 25\%$), local atomic environment around $V_{\text{O}}^{\bullet}$ plays a more significant role on TE. By screening the thermodynamically stable configurations, it is found that $V_{\text{O}}^{\bullet}$ and dopant tend cluster. Expansion of BZY20 is more sensitive to temperature increase instead of hydration. Therefore, when assembled with hydration-sensitive electrode PNC73, the interface will be destabilized mainly due to the water molecule incorporation.

1.4 Thesis Outline

This thesis is organized into seven chapters. Chapter 1 introduces the background and motivation of the research included in the thesis. Chapter 2 provides the basic methodologies employed in the thesis research, including density functional theory (DFT), three machine learning potential frameworks (HDNNP, SGP, and Allegro), molecular dynamics (MD) simulations, and phonon calculations. Chapter 3 employed the high-dimensional neural network potential (HDNNP) to predict the properties and phase transition of condensed phase nickel. Chapter 4

predicted the thermodynamic phase diagram of icosahedral boron allotropes using a sparse Gaussian process (SGP) potential. Chapters 5 and 6 investigated the lattice expansion behaviors by performing large-scale phonon calculations enabled by the Allegro framework. The conclusions and outlook of the research presented are summarized in Chapter 7.

Chapter 2 - Computational Methods

To realize MLP-based molecular simulations, multiple techniques are needed throughout the process. Starting with training set generation, a well-behaved MLP will be trained, followed by MD simulations and phonon calculations.

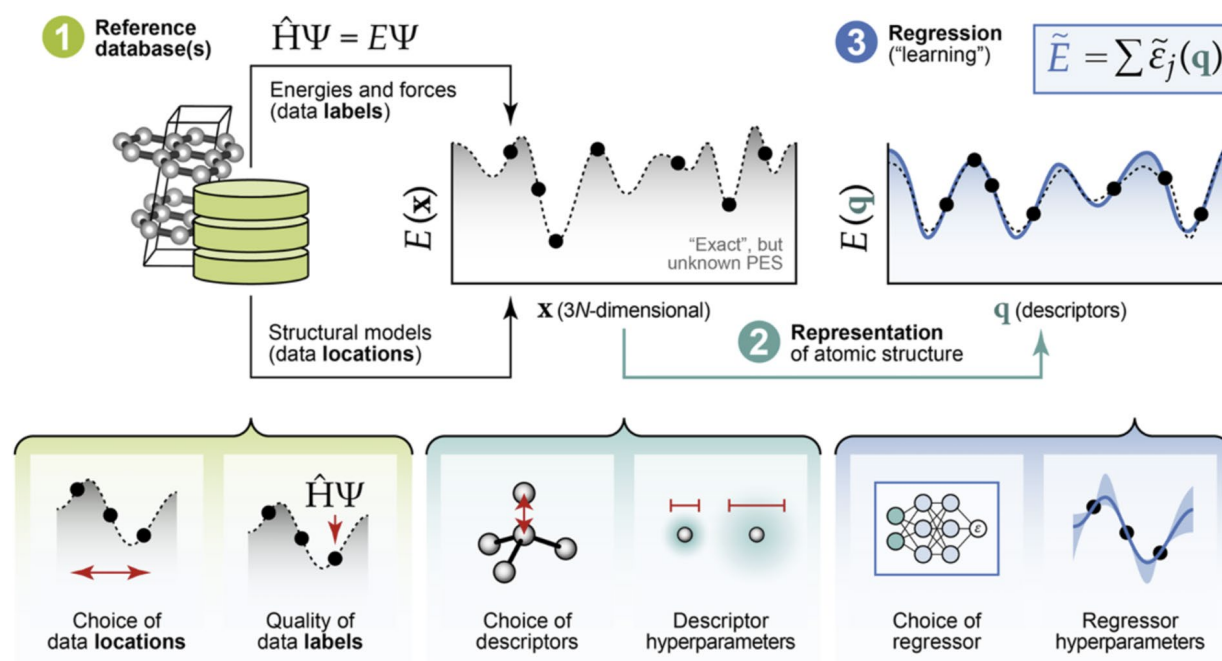


Figure 2.1 Illustration of MLP training process, including (1) training set (reference database) generation, (2) choosing the descriptor and representation, and (3) choosing the MLP model and training (regression), adopted from Ref. [77]

To initiate training an MLP, as shown in Step 1 in Figure 2.1⁷⁷, a training set (reference database) must be acquired from first-principles calculations. The structural data extracted from the first-principles calculations will be decomposed as a set of chemical environments containing a central atom and its neighbors within a predetermined cutoff radius (Figure 2.2). However, data from the first-principles calculations are based on cartesian coordinates, which are not transferable

with respect to rotation and translation, as well as the order of the neighboring atoms. Therefore, to make the data compatible with MLP training, the cartesian coordinates must be represented by another vector called a descriptor (Step 2 in Figure 2.1), invariant to rotation, translation, and ordering. After being represented by the descriptor, the training set will be ready for the chosen MLP model to conduct the training process (Step 3 in Figure 2.1). Once the MLP is validated and tested, production simulations can be performed to simulate the process of interest.

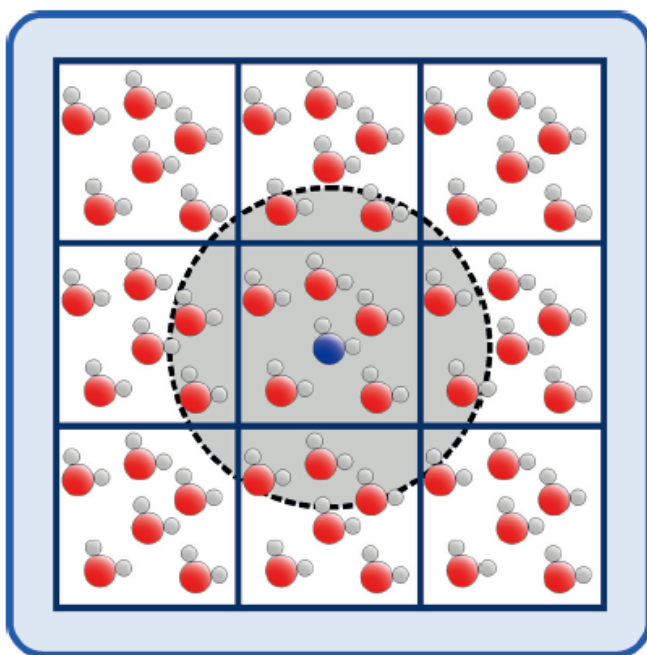


Figure 2.2 Illustration of the chemical environment. The central atom is marked as blue, and all the atoms within the dashed sphere defined by the cutoff radius are the neighbors of the central atom, adopted from Ref. [26]

This chapter outlines the fundamental methodologies used in this MLP-based research, encompassing density functional theory (DFT), three machine learning potential frameworks (HDNNP, SGP, and Allegro), along with molecular dynamics (MD) simulations and phonon

calculations. DFT serves as the training set generator throughout this thesis. HDNNP and SGP were employed as MLP models in Chapters 3 and 4 as the MLP model, respectively, and Allegro was used for Chapters 5 and 6. MD simulations and phonon calculations are the production simulation techniques used in the thesis.

2.1 Density Functional Theory (DFT)

As one of the most mature and efficient first-principles calculations, DFT was widely employed as the training set generator for MLP due to its high reliability. This subchapter will introduce the basic framework of DFT based on plane waves.

2.1.1 Kohn–Sham Formulation

In first-principles methods, the wave function of the N -particle system $\Psi(r_1, \dots, r_N)$ can be represented as the product of single-particle wave functions $\psi_i(r_i)$ of particle i :

$$\Psi(r_1, \dots, r_N) = \psi_1(r_1) \cdots \psi_N(r_N) \quad (2.1)$$

Based on Kohn-Sham's formulation⁷⁸, the single-particle wave function can be acquired by the Schrödinger equation in the form of Eqn. 2.2, which is the foundation of the density functional theory calculation.

$$\left[-\frac{\hbar}{2m} \nabla_i^2 + v_{eff}(r) \right] \psi_i(r) = \varepsilon_i \psi_i(r) \quad (2.2)$$

The Hamiltonian in the square bracket contains the electronic kinetic term and the effective potential functional $v_{eff}(r)$, which can be written as Eqn. 2.3:

$$v_{eff}(r) = V_{en}(r) + \int \frac{\rho(r')}{|r - r'|} dr' + V_{xc}[\rho(r)] \quad (2.3)$$

In the form above, the $v_{eff}(r)$ is the functional of the electron density $\rho(r)$, which can be calculated by:

$$\rho(r) = \sum_i |\psi_i(r_i)|^2 \quad (2.4)$$

Eqn. 2.2 to 2.4 form a circular problem that must be solved using a self-consistent algorithm shown in Figure 2.3.

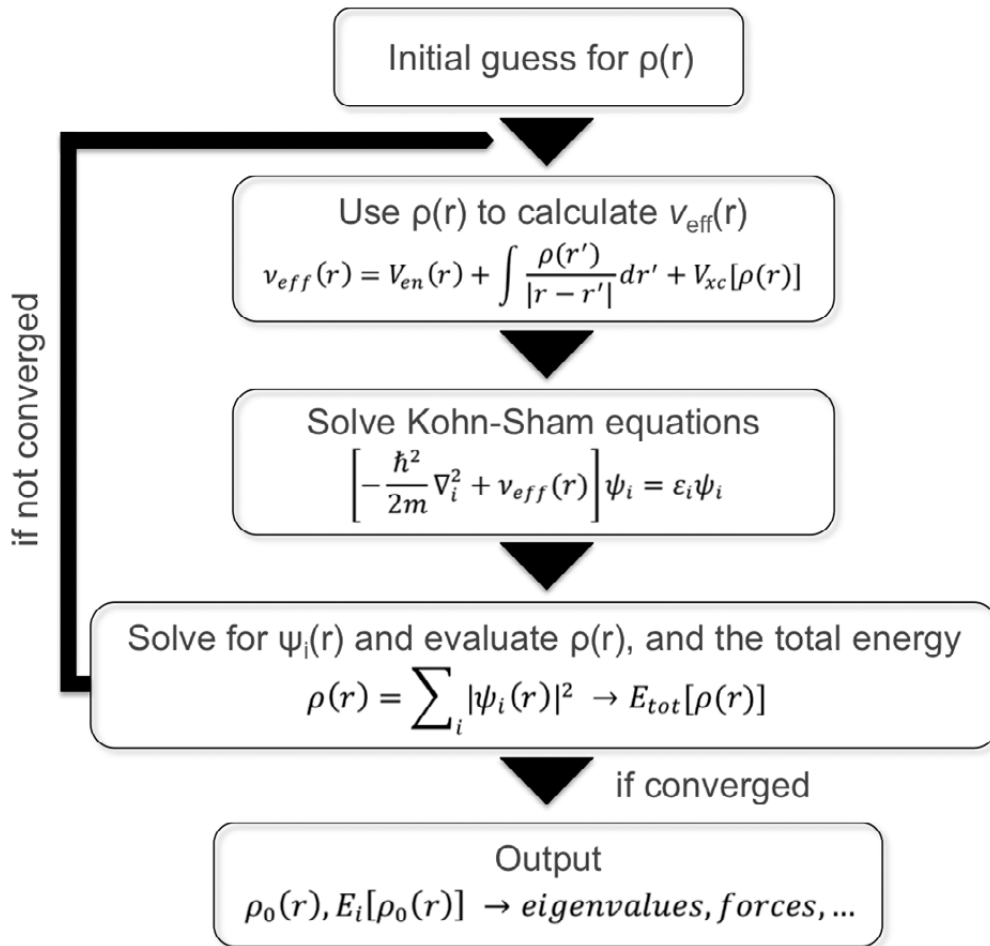


Figure 2.3 The self-consistent scheme to solve the Kohn-Sham Eqn., adopted from Ref. [79]

2.1.2 Plane Wave Basis Set

For the solid phase material with periodic arrangement, Bloch's theorem⁸⁰ can be employed, which uses periodic functions to represent the wavefunctions. The wave function of an electron occupying a state of k (wave vector confined to the first Brillouin Zone) can be expressed as Eqn. 2.5:

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{\vec{k}}(\vec{r}) \quad (2.5)$$

$u_{\vec{k}}(\vec{r})$ is a periodic function in space with the same periodicity as the supercell. Due to periodicity, $u_{\vec{k}}(\vec{r})$ can be expanded in terms of a set of plane waves:

$$u_{\vec{k}}(\vec{r}) = \sum_{\vec{G}} c_{\vec{G}} e^{i\vec{G}\cdot\vec{r}} \quad (2.6)$$

$\vec{G}$ is the lattice vector in the reciprocal space. Combining Eqn. 2.5 and 2.6, the plane wave function can be written in:

$$\psi_{ik}(\vec{r}) = \sum_{\vec{G}} c_{i\vec{k}}(\vec{G}) e^{i(\vec{k}+\vec{G})\cdot\vec{r}} \quad (2.7)$$

where $c_{i\vec{k}}$ is the expansion coefficient of vector $\vec{G}$. The series in Eqn. 2.7 can be truncated using the predefined cut-off energy (E_{cut}):

$$E_{cut} = \frac{\hbar^2}{2m} \vec{G}_{cut}^2 \quad (2.8)$$

$$|\vec{k} + \vec{G}| < \vec{G}_{cut} \quad (2.9)$$

2.1.3 Pseudopotential

Due to the rapid oscillations of electron wavefunctions near atomic nuclei, all-electron calculations using the full Coulombic potential are often computationally expensive. However,

since chemical bonding and material properties are primarily governed by the weakly bound valence electrons, using pseudopotentials (PPs) can provide an effective solution. By treating core electrons as frozen and replacing their complex behavior with smoother mathematical forms, PPs significantly reduce the number of plane waves required in calculations. Among the widely used methods, the projector augmented-wave (PAW) ⁸¹ approach is advantageous for its ability to closely reproduce all-electron wavefunctions, especially beyond a certain cutoff radius where accurate representation of valence electron behavior is critical. This method balances computational efficiency and accuracy and was adopted for all density functional theory (DFT) calculations in this thesis.

2.1.4 Electron Exchange-Correlation and Self-Interaction Errors

In Kohn–Sham density functional theory (DFT), the exchange-correlation energy $V_{xc}[\rho(r)]$ accounts for the complex many-body effects of electron–electron interactions. Specifically, it captures both the kinetic energy difference between an interacting electron system and its non-interacting counterpart, as well as the corrections for self-interactions in the Coulomb potential, since the exact form of $V_{xc}[\rho(r)]$ is unknown, various approximations have been developed over the years, which built Jacob’s ladder ^{82, 83} of exchange-correlation functionals.

The base of this ladder is the Local Density Approximation (LDA) ⁸⁴, which assumes a uniform electron gas. LDA often performs reasonably well for predicting molecular geometries and vibrational frequencies, but tends to underestimate binding energies and overestimate electron localization. A more accurate and widely used level of approximation is the Generalized Gradient Approximation (GGA), which incorporates both the local electron density and its gradient. Well-known GGA functionals include Perdew-Burke-Ernzerhof (PBE) ⁸⁵, Perdew-Wang (PW91) ⁸⁶,

revised PBE (RPBE)⁸⁷, and PBEsol⁸⁸. GGA functionals achieve a good balance between accuracy and computational cost and are especially well-suited for systems such as transition metals, surfaces, and interfaces.

To further improve predictive power, higher level functionals were developed and applied. Meta-GGA functionals such as strongly constrained and appropriately normed (SCAN)⁸⁹, and its regularized version r²SCAN⁹⁰, showed improved accuracy on predictions of formation and decomposition enthalpies, equilibrium volumes, and fundamental band gaps⁹¹. In addition, the hybrid functionals, such as B3LYP^{92, 93}, HSE06⁹⁴ and PBE0⁹⁵, mix a portion of exact Hartree-Fock exchange with DFT exchange-correlation, enhancing performance for band gap calculations and electronic structure predictions.

Despite the high accuracy, the high computational cost of higher level functionals makes them less practical for large systems or dataset generation. Therefore, in this thesis, the GGA-PBE functional will be primarily employed due to its balance of accuracy and computational efficiency.

The PBE functional suffers from self-interaction errors (SIEs) when systems involve localized d- or f-electrons. Such phenomena must be addressed in the DFT calculations of PNC perovskite containing Ni and Co atoms. To mitigate these effects, Hubbard U correction is commonly introduced within the GGA framework, resulting in the GGA+U methods^{96, 97}. Empirical tuning and benchmarking of the U parameter are often necessary, typically based on experimental values. In this thesis, we use the U parameters for PNC acquired from the literature.

2.2 Machine Learning Potential (MLP)

In this thesis, three MLP models, i.e., high-dimensional neural network potential (HDNNP), sparse Gaussian process (SGP), and Allegro, were employed to obtain the structural

and thermodynamic properties of several condensed-phase systems based on their respective characteristics. Table 2.1 compares the computational and training range of these MLP models considered in this work. HDNNP is one of the most straightforward MLP models that can be trained using various data sources. Typically, an extensive training set is needed. As the complexity increased, as for icosahedral boron and the $\text{PrNi}_{1-\delta}\text{Co}_\delta\text{O}_3$ perovskite mentioned in Chapter 4 and 5, respectively, two more advanced MLP models, i.e., SGP and Allegro, which enable active learning and equivariant operation, respectively, were employed. More importantly, as demonstrated in the following discussion, SGP and Allegro are indeed more efficient and suitable MLPs.

Table 2.1 Comparisons of computational and training range among HDNNP, SGP, and Allegro.

MLP framework	Number of frames in training sets	MD speed of multi-element system	MD speed related to training set size	Active learning
HDNNP	1000s~10,000s	Low	No	No
SGP	~100s	Low	Yes	Yes
Allegro	100s~10000s	High	No	No

2.2.1 High-Dimensional Neural Network Potential (HDNNP)

Behler and Parrinello ⁹⁸ first devised the high-dimensional neural network potential (HDNNP) framework in 2007. The HDNNP family started from early neural network potentials

designed for small-scale and simple systems, like small molecules in the gas phase and molecular interactions with frozen metal surfaces. The development of HDNNP was accomplished by using the summation of atomic energy contributions and symmetry functions as descriptors^{99, 100}, simplifying the training process.

The neural network model is a biology-inspired mathematical formulation. As shown in Figure 2.4A, the neural network consists of many artificial neurons (nodes) arranged in layers, where the input and output layers provide structural descriptions (G) and potential energies (E), respectively. The number of hidden layers and nodes inside each layer defines the specific functional form of the neural network. The more hidden layers and nodes per layer, the more flexible the functional form.

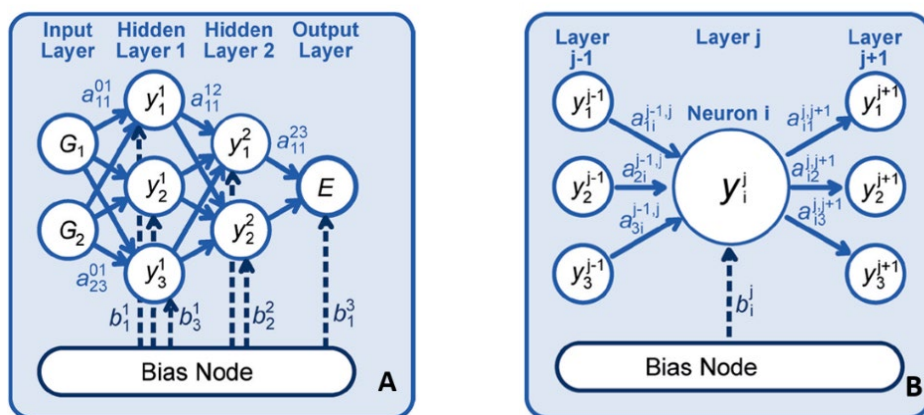


Figure 2.4 (A) Schematic structure of a feed-forward neural network defining a two-dimensional input vector $\mathbf{G} = (G_1, G_2)$ that describes the atomic configuration and the potential energy E . (B) Structure of an artificial neuron i in layer j . Adapted from Ref. [26].

As shown in Figure 2.4B, the value of a generic node i in layer j (y_i^j) can be determined by Eqn. 2.10:

$$y_i^j = f_i^j(x_i^j) = f_i^j\left(b_i^j + \sum_{k=1}^{N_{j-1}} a_{k,i}^{j-1,j} \cdot y_k^{j-1}\right) \quad (2.10)$$

where $a_{k,i}^{j-1,j}$ is the weight and b_i^j is the bias. f_i^j represents the activation function, typically a continuous nonlinear function. The full analytical form of the feed-forward neural network (Figure 2.4A) can be expressed as Eqn. 2.11:

$$E = f_1^3\left(b_l^3 + \sum_{l=1}^{N_{hid,2}} a_{kl}^{23} \cdot f_k^3\left(b_k^2 + \sum_{k=1}^{N_{hid,1}} a_{jk}^{12} \cdot f_j^1\left(b_j^1 + \sum_{i=1}^{N_G} a_{ij}^{01} \cdot G_i\right)\right)\right) \quad (2.11)$$

The output of the neural network (E) is the decomposed energy of a single atom. The total energy of the system is obtained by summing over all atoms in the system (Eqn. 2.12):

$$E_s = \sum_{\mu=1}^{N_{atoms}} E_{\mu} \quad (2.12)$$

For multi-component systems, this Eqn. becomes a sum of all the atomic energy contributions of all the elements (Eqn. 2.4):

$$E_s = \sum_{v=1}^{N_{elem}} \sum_{\mu=1}^{N_{atoms,v}} E_{\mu}^v \quad (2.13)$$

In Behler-Parrinello's HDNNP formulation, the structural descriptor is represented by symmetry functions, which must satisfy invariance concerning rotation, translation, and the order

of neighboring atoms. The radial and angular symmetry functions used to describe the local chemical environment are expressed by Eqn.s 2.14 and 2.15, respectively.

$$G_i^2 = \sum_j e^{-\eta(R_{ij}-R_s)^2} \cdot f_c(R_{ij}) \quad (2.14)$$

$$G_i^4 = 2^{1-\zeta} \sum_j \sum_{k \neq j} (1 + \lambda \cdot \cos \theta_{ijk})^\zeta f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk}) \quad (2.15)$$

where R_s is the radial shift, θ_{ijk} is the angle centered at atom i with neighbors j and k . Three constants, i.e., η , λ , and ζ are predetermined factors. As shown in Figure 2.5⁹⁹, by defining η , λ , and ζ , we obtain a set of different functions to extract the fingerprint of the atomic position.

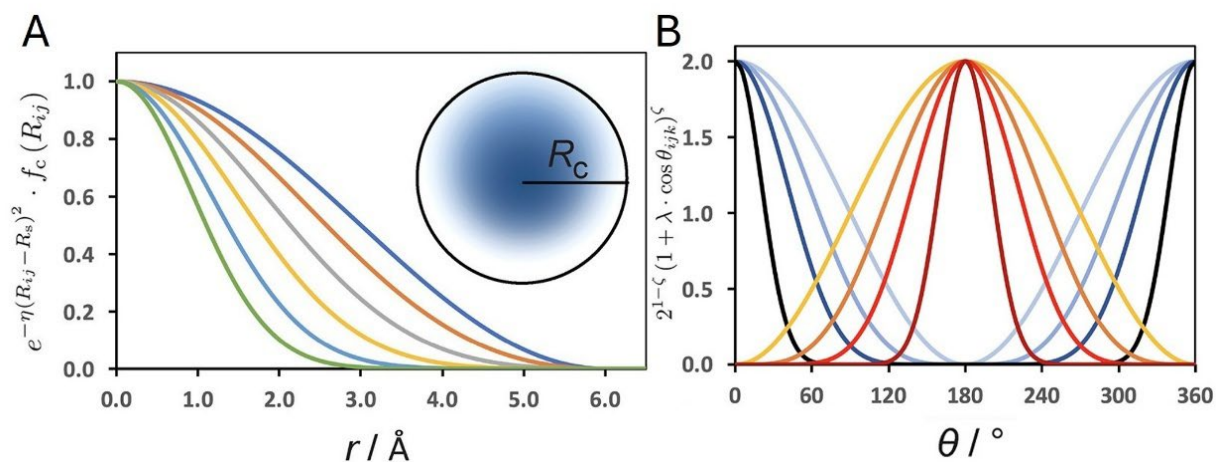


Figure 2.5 (A) Radial symmetry functions. The values from dark blue to green are $\eta = 0.0$, 0.03, 0.08, 0.16, 0.3, and 0.5 $\text{\AA}^{-2}$ with $R_s = 0 \text{\AA}$. (B) Angular Terms with $\lambda = 1$ are shown in blue to black, terms with $\lambda = -1$ are shown in orange to brown. The employed exponents are $\zeta = 1, 2, 4$, and 16 (from light to dark tones).⁹⁹

f_c is the cutoff function (Eqn. 2.16), which ensures that the value and slope of symmetry functions decay to zero at the cutoff radius R_c .

$$f_c(R_{ij}) = \begin{cases} 0.5 \cdot \left[\cos\left(\frac{\pi R_{ij}}{R_c}\right) + 1 \right], & R_{ij} \leq R_c \\ 0, & R_{ij} > R_c \end{cases} \quad (2.16)$$

2.2.2 Sparse Gaussian Process (SGP) Potential

The accuracy and stability of the trained HDNNP are sensitive to the contents included in the training set. One effective remedy is constructing training sets by introducing active learning techniques. The SGP potential supports active learning with a built-in mechanism to estimate simulation uncertainties.

SGP predicts atomic forces by comparing the environment of interest with a dataset of atomic environments with known forces. To quantify the similarity between the target environment and those in the training set, the normalized dot product kernel function ($k(\mathbf{d}_1, \mathbf{d}_2)$) is determined by Eqn. 2.17,

$$k(\mathbf{d}_1, \mathbf{d}_2) = \sigma^2 \left(\frac{\mathbf{d}_1 \cdot \mathbf{d}_2}{d_1 d_2} \right)^2 \quad (2.17)$$

where $\mathbf{d}_1$ and $\mathbf{d}_2$ are atomic cluster expansion (ACE) descriptors¹⁰¹ for environments 1 and 2, and σ is the signal variance. With the kernel function, the local energy (ε) predictions of a given environment i can be achieved by Eqn. 2.18:

$$\varepsilon_i = \sum_t^{t \in N_s} k(\mathbf{d}_i, \mathbf{d}_t) \alpha_t \quad (2.18)$$

where S represents a sparse set, N_S is the number of environments in the sparse set, $\mathbf{d}_i$ and $\mathbf{d}_t$ are ACE descriptor vectors for the respective target and sparse set environment. α_i is the element in the training coefficient vector (α).

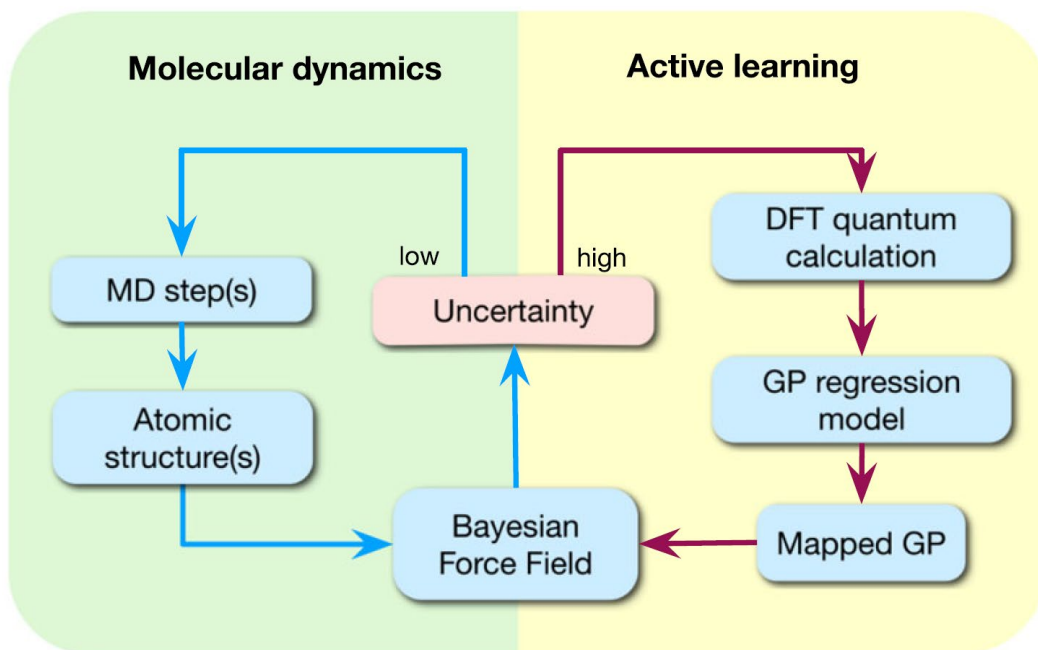


Figure 2.6 Active learning (OTF MD) workflow based on the inherent uncertainty quantification mechanism of the SGP model. Adapted from Ref. [40].

On-the-fly (OTF) MD simulation, which drives active learning, has been successfully integrated with SGP model development. As shown in Figure 2.5, during the OTF MD, once the configuration is updated by a new MD step based on the current potential, the algorithm will estimate the uncertainty of the new structure based on an inherent quantification mechanism (Eqn. 10) to determine whether the predetermined uncertainty threshold is exceeded. If the uncertainty exceeds the threshold, a DFT calculation will be performed based on the current configuration.

This new configuration, along with its energy, force, and stress information, will be added to the training set to update the current SGP potential.

To determine if the DFT call is necessary for an updated structure during OTF training, the total uncertainty (σ_t^2) of the new structure will be determined using Eqn. 2.19:

$$\sigma_{t,i}^2 = \sigma_{i\alpha}^2 + \sigma_n^2 \quad (2.19)$$

where σ_n is the observation noise (a trainable parameter), and $\sigma_{i\alpha}^2$ is the epistemic uncertainty expressed by Eqn. 2.20:

$$\sigma_{i\alpha}^2 = k(\mathbf{d}_i, \mathbf{d}_i) - \bar{k}_{i\alpha}^T (K + \sigma_n^2 I)^{-1} \bar{k}_{i\alpha} \quad (2.20)$$

where $\bar{k}_{i\alpha}$ is the force kernel between the target and sparse set environments. K is the covariance matrix of the training set. Unlike HDNNP, SGP works closely with the training set when predicting energy and forces, which is the leading cause of the OTF MD becoming less efficient when the OTF MD trajectory becomes longer.

2.2.3 Allegro – a Neural Equivariant Interatomic Potential

Allegro is a deep learning interatomic potential built on E(3)-equivariant convolutions over geometric tensors. The advantage of Allegro is that it can be trained over more information-rich and faithful representations of atomic environments. In most MLP models (e.g., symmetry function discussed in Chapter 2.1), the environment descriptors are purely scalar-based because of the requirement of being invariant to translation, permutation, and rotation. In contrast, Allegro employs the E(3)-equivariant convolutions for interactions of geometric tensors, which utilizes equivariant operation explicitly.

Formally, a function $f: X \rightarrow Y$ is equivariant with respect to a group G that acts on X and Y if:

$$D_Y[g]f(x) = f(D_X[g]x) \quad \forall g \in G, \forall x \in X \quad (2.21)$$

where $D_X[g]$ and $D_Y[g]$ in Eqn. 2.21 are the representations of the group element g in the vector spaces X and Y , respectively. In this case, Allegro uses relative position vectors directly instead of just the distance scalars, making the descriptor equivariant to rotation and enabling the usage of angular information^{43, 44}. The Allegro trained on liquid phase water can accurately predict the phase structures and phase equilibria, including equilibrium vapor-liquid density and vibrational density of states of ice⁵³. Allegro has been used to train MLP on other systems with high compositional and structural complexity, including solid acid electrolytes CsH_2PO_4 and CsHSO_4 ¹⁰² and eutectic solvent of ethylene glycol and choline chloride¹⁰³.

Allegro atomistic structures by assigning each atom to a node within a graph. The edges are created between pairs of atoms that lie within a cutoff radius. The nodes in deeper layers provide relationships between neighboring atoms. Allegro employs the E(3)-equivariant convolutions for interactions of geometric tensors, which utilize equivariant operation explicitly, making the descriptor equivariant to rotation and enabling the usage of angular information^{43, 44}.

The total energy of an N-atom system (E_{system}) is expressed as a sum over per-atom energies (E_i), which are, in turn, sums of pairwise energies (E_{ij}) between atom i and its neighbors j :

$$E_{\text{system}} = \sum_i^N \sigma_{Z_i} E_i + \mu_{Z_i} \quad (2.22)$$

$$E_i = \sum_{j \in N(i)} \sigma_{Z_i Z_j} E_{ij} \quad (2.23)$$

where σ_{Z_i} and μ_{Z_i} are trainable per-species scale and shift parameters, $\sigma_{Z_i Z_j}$ is a per-species-pair scaling factor. The forces ($\vec{F}_i$) are computed from the negative gradient of the total energy with respect to atomic positions, ensuring energy conservation:

$$\vec{F}_i = -\nabla_i E_{system} \quad (2.24)$$

The architecture of Allegro encodes information for ordered pairs of neighboring atoms using two distinct latent spaces: an invariant latent space composed of scalar features ($l = 0$) and an equivariant latent space that handles tensor features of arbitrary rank $l \geq 0$. These two spaces interact at every layer of the network. Ultimately, the pair energy E_{ij} is predicted by a multi-layer perceptron (MP) applied to the scalar features from the final layer.

The two-body invariant scalar latent features X_{ij}^L of the ordered pair of atoms ij at layer L is embedded by:

$$\mathbf{X}^{ij,L} = \text{MP}_{\text{two-body}} (1\text{Hot}(Z_i) \parallel 1\text{Hot}(Z_j) \parallel B(r_{ij})) \cdot u(r_{ij}) \quad (2.25)$$

where $\parallel$ denotes concatenation, $1\text{Hot}(\cdot)$ is a one-hot encoding of the center and neighbor atom species Z_i and Z_j , and $B(r_{ij})$ is the projection of the interatomic distance r_{ij} onto a radial Bessel basis.

$$B(r_{ij}) = (B_1(r_{ij}) \parallel \dots \parallel B_{N_{basis}}(r_{ij})) \quad (2.26)$$

The equivariant features $\mathbf{V}_{n,l,p}^{ij,L=0}$ are acquired by linear embedding of the l -th real spherical harmonic projection $Y_{l,p}^{ij}$ of the unit vector $\vec{r}_{ij}$ which has parity $p = (-1)^l$.

$$\mathbf{V}_{n,l,p}^{ij,L=0} = w_{n,l,p}^{ij,L=0} Y_{l,p}^{ij} \quad (2.27)$$

where the channel index is $n = 0, 1, \dots$, nequivariant - 1, and $w_{n,l,p}^{ij,L=0}$ is the scalar weights for each center-neighbor pair ij computed from the initial two-body scalar latent features:

$$w_{n,l,p}^{ij,L=0} = \text{MP}_{\text{embed}}^{L=0} (\mathbf{X}^{ij,L=0})_{n,l,p} \quad (2.28)$$

The equivariant features can be updated as a weighted sum of the tensor products of the current features with the geometry of the various other neighbor pairs ik in the local environment of i :

$$\mathbf{v}_{n,(l_1,p_1,l_2,p_2) \rightarrow (l_{out},p_{out})}^{ij,L} = \mathbf{v}_{n,l_1,p_1}^{ij,L-1} \otimes \left(\sum_{k \in N(i)} w_{n,l_2,p_2}^{ij,L} \vec{Y}_{l_2,p_2}^{ik} \right) \quad (2.29)$$

2.3 Molecular Dynamics (MD)

MD simulation is the production simulation method employed in Chapter 3 and 4, which is a powerful tool for relaxation and simulating the dynamical evolution of large-scale system. The high efficiency enables the MD to simulate a large scale up to micrometers ⁴⁵. The movement of the i th coordinate of an individual atom N is captured by solving the motion Eqn. of Newton's laws as Eqn. 2.30 ¹⁰⁴:

$$F_i = m_i a_i = m_i \frac{dv_i}{dt} \quad (2.30)$$

where F_i , a_i , v_i , and t are the force, acceleration, velocity and time, respectively. The force is related to the derivative of the total potential energy:

$$F_i = -\frac{\partial U}{\partial r_i} \quad (2.31)$$

where U is the potential energy. The U - r relationship can be acquired from the potential energy surface determined by interatomic potentials. The total energy E of the system can be given as the summation of potential energy U and kinetic energy K :

$$E = K + U \quad (2.32)$$

$$K = \frac{1}{2} \sum_{i=1}^{3N} m_i v_i^2 \quad (2.33)$$

One important operation during MD simulations is controlling the temperature and/or pressure to make the simulated system comparable to realistic condition. Such control is achieved by defining atoms in an ensemble. In this thesis, two ensembles were employed including

canonical ¹⁰⁵ and isothermal-isobaric ¹⁰⁶ ensembles. Canonical ensemble defines N , V , and T as constants (NVT ensemble), while isothermal-isobaric ensemble defines N , P , and T as constants (NPT ensemble). In both NVT and NPT ensemble, the Nose-Hoover thermostat tunes the kinetic energy K by:

$$dK/dt = k^2(K - K_0)^2 \quad (2.34)$$

where k is the constant with units of $1/t$, and K_0 is the target kinetic energy. And the barostat tune the pressure by change the volume where the relaxation rate is set by its inertia W :

$$W = (N + 1)k_B T_{target} P_{damp}^2 \quad (2.35)$$

where k_B is the number of atoms, k_B is the Boltzmann constant, T_{target} is the target temperature, and P_{damp} is stress damping parameter.

2.4 Phonon Calculation

Phonon is defined as the collective vibrations of atoms in the crystal. Phonon calculations are widely employed in research due to their strong predictive capabilities. They enable accurate estimation of thermal properties such as heat capacity, thermal expansion, and thermal conductivity ¹⁰⁷. In this thesis, phonon calculation was employed as the production tool of MLP to calculate vibrational entropy and thermal expansion.

2.4.1 Vibrational Entropy Calculations

To initiate the phonon calculations, supercells need to be generated with all possible atom displacements. Then the dynamical matrix $D_{\kappa\alpha,\kappa'\alpha'}(\mathbf{q})$ can be formulated as Eqn. 2.36:

$$D_{\kappa\alpha,\kappa'\alpha'}(\mathbf{q}) = \frac{1}{N\sqrt{m_\kappa m_{\kappa'}}} \sum_{ll'} \Phi_{l\kappa\alpha,l'\kappa'\alpha'} e^{i\mathbf{q}\cdot(\mathbf{R}_{l'\kappa'}^0 - \mathbf{R}_{lk}^0)} \quad (2.36)$$

where the l and κ label the lattice points and atoms in the primitive cell, and α denotes the Cartesian coordinates. m_κ is the atomic mass of atom κ . $\Phi_{l\kappa\alpha,l'\kappa'\alpha'}$ is the harmonic force constants, which can be calculated as the second derivative of the energy with respect to atomic positions. $\mathbf{q}$ is the wave vector. $(\mathbf{R}_{l'\kappa'}^0 - \mathbf{R}_{lk}^0)$ denotes the displacement. Then the phonon frequency $\omega_{\mathbf{q}v}$ can be obtained by solving the eigenvalue equation 2.37:

$$\sum D_{\kappa\alpha,\kappa'\alpha'}(\mathbf{q}) W_{\kappa'\alpha'}(\mathbf{q}v) = \omega_{\mathbf{q}v}^2 W_{\kappa\alpha}(\mathbf{q}v) \quad (2.37)$$

where $W_{\kappa\alpha}(\mathbf{q}v)$ is the eigen vector and v labels the phonon band index. The vibrational entropy $S^{vib}(T)$ can be given by:

$$S^{vib}(T) = \frac{1}{2T} \sum_{\mathbf{q}v} \hbar\omega_{\mathbf{q}v} \coth(\hbar\omega_{\mathbf{q}v}/2k_B T) - k_B \sum_{\mathbf{q}v} \ln[2\sinh(\hbar\omega_{\mathbf{q}v}/2k_B T)] \quad (2.38)$$

2.4.2 Quasi-Harmonic Approximations (QHA)

Based on quasi anharmonicity, the free energy is dependent on volume and temperature, which is given by Eqn. 2.39:

$$G(T, p) = U_{el}(V) - TS_{vib}(V) + pV \quad (2.39)$$

where U_{el} is the electronic ground-state energy calculated by density functional theory (DFT), and S_{vib} is the vibrational entropy from phonon calculations. In the QHA calculation, the free energy is sampled over unit cells with varied volumes and temperatures, and the energies are fitted to a set of equation of states (EOS), as shown in Figure C.11, C.12, and C.13. The volume under each

temperature can be identified by locating the minima of corresponding EOS curves, then the volume expansion can be calculated.

Chapter 3 - Phase Transition Simulations of Condensed Phase

Nickel

Chapter 3 is reproduced with permission from:

Deng H.; Comer J.; Liu, B. “A high-dimensional neural network potential for molecular dynamics simulations of condensed phase nickel and phase transitions”, *Molecular Simulations*. DOI: 10.1080/08927022.2022.2156561.

3.1 Introduction

For molecular simulations of the condensed-phase nickel (Ni) systems, empirical and semi-empirical interatomic potentials have been used in the literature, such as the embedded-atom-method (EAM)¹⁰⁸⁻¹¹⁰, the modified-embedded-atom-method (MEAM)¹¹¹⁻¹¹³, the embedded-atom-method Finnis-Sinclair (EAMFS) potential¹¹⁴, the effective medium theory (EMT)¹¹⁵, and the reactive force field (ReaxFF)^{116, 117}. These potentials yield reliable predictions of Ni's solid or liquid phase behaviors, a highly versatile material for catalysis^{54, 55} or material synthesis^{56, 57}. A list of potentials and their applications are summarized in Table 3.1. The accuracy and transferability of these potentials depend on the parameterization and training. Inevitably, the embedded physical parameters are trained against computational or experimental references, a tedious but critical step to obtaining reliable predictions.

Recently, MLPs emerged as an alternative class of potentials for MD simulations on Ni systems^{31, 118}. For instance, Li *et al.*³¹ applied the SNAP model to Ni-Mo binary alloys and suggested that it can be as accurate as the EAM potential. Also, Zuo *et al.*¹¹⁸ developed MLPs for

several solid-state metallic (i.e., Li, Mo, Cu, Ni) and semiconductor materials (i.e., Si, Ge). A systematic survey of the MLPs derived from NNP, SNAP, GAP, and MTP suggested that the MLPs trained in the NNP approach display a good balance between accuracy and computational cost. Nevertheless, MLPs trained to accurately describe the potential energy surface spanning the solid and liquid phases of condensed systems are rare.

Table 3.1 Summary of interatomic potentials for Ni in MD simulations.

Potential	Form	Applications
EAM	$E = \sum_i \left(F_i(\bar{\rho}_i) + \frac{1}{2} \sum_{j \neq i} \varphi_{ij}(R_{ij}) \right)$	• Thermal expansion, melting point (2011)¹⁰⁸ • Equilibrium properties, elastic constants, phonon-dispersions (1999)¹⁰⁹ • Activation energy of self-diffusion (1989)¹¹⁰
MEAM	Adding an angular dependency to EAM	• Elastic constants, structural energy differences, melting point (2015)¹¹¹ • Latent heat, melting expansion, liquid structure factor (2018)¹¹² • Bulk properties, point defect properties, thermal properties (2003)¹¹³
EAMFS	Same as EAM, except that ρ is a functional specific to elements	• Radial distribution function, mean square displacement (2018)¹¹⁴
EMT	$E = \sum_i E_{c,i}(n_i) + \Delta E_{AS} + \Delta E_{1el}$	• Energetics of metal atoms deposited on metal surfaces (1996)¹¹⁵
ReaxFF	$E = E_{bond}(BO) + E_{over}(BO) + E_{under}(BO) + E_{val}(BO) + E_{pen}(BO) + E_{tors}(BO) + E_{conj}(BO) + E_{vdWaals} + E_{Coulomb}$	• Elastic constants, thermal expansion, and melting point (2021)¹¹⁶ • Self-diffusivity (2015)¹¹⁷

Neural network potential (NNP) is one of the most developed techniques to mimic the potential energy surface of the chemical system. Unlike classic potentials, NNP can be free from physics-driven parameterizations and thus, be trained against data sets of immense variety, allowing these potentials to achieve accuracy comparable to quantum mechanical calculations. Moreover, various molecular structure description protocols are allowed in NNP-based MD frameworks, including the high-dimensional neural network potential (HDNNP) by Behler and Parrinello^{25, 26}, or the deep potential for molecular dynamics (DeePMD)^{119, 120} based on deep learning and end-to-end neural networks¹²¹. NNPs have been successfully applied to carbon-based materials¹²², Ag₂S¹²³, Zr¹²⁴, metallic Al-Ce liquid¹²⁵, Ag-Au alloy¹²⁶, condensed phase water^{127, 128}, H₂ reactions on Pt crystals¹²⁹, the Zr-Rh system¹³⁰, oxygen reduction reactions (ORR) on TiAl₂O₅ surfaces¹³¹, Ga¹³², and liquid SiO₂¹³³.

A unified HDNNP was developed to describe the solid and liquid-phase Ni behaviors and the phase transition process. Trained using the dataset generated from *ab initio* molecular dynamics (AIMD) simulations, our potential enables simulations over a temperature range from 0 to 2307 K, encompassing the melting point of nickel (1726 K). The performance of the NNP developed herein was also calibrated with a series of simulations to determine lattice constants, thermal expansion coefficients, radial distribution functions (RDF), diffusivities, and the melting point. We believe that the training process presented here can be used to create reliable NNPs to tackle the challenges of predicting crystal nucleation in complex crystals and their phase transitions.

3.2 Methods

3.2.1 Training Data Generation

All datasets were generated by AIMD using Vienna Ab Initio Simulation Package (VASP)^{8, 9}. The projector augmented wave (PAW) pseudopotential⁸¹, and the GGA-PBE functional¹³⁴ were used to account for the core electrons and the electron exchange-correlation effects. The plane-wave basis set was expanded with a cutoff energy of 380 eV. All AIMD simulations were performed using the NVT ensemble with temperatures controlled by the Nose-Hoover thermostat, with a time step of 2 fs. As shown in Figure 3.1, a face-centered cubic (FCC) supercell consisting of 108 Ni atoms was used as the initial structure for AIMD simulations.

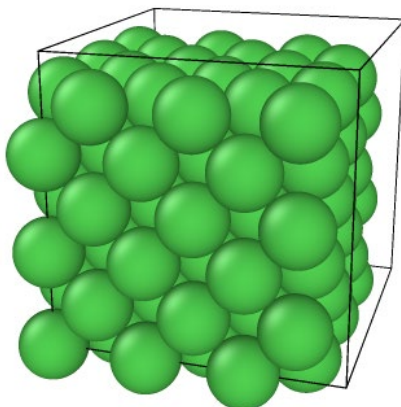


Figure 3.1 An FCC 108-atom Ni supercell. The solid black lines represent the cell boundaries.

Twelve temperatures were chosen between 100 K and 2589 K to adequately sample the targeted temperature regimes for AIMD simulations. The simulation box volume was artificially varied isotropically by scaling the cell dimensions. Three scaling factors were selected for each temperature to represent systems experiencing zero, tensile, and compression stresses. In each simulation, 200 snapshots were extracted from the trajectory upon equilibrium. In total, 7200 data points were included in our dataset. The detailed process of dataset generation using AIMD and

temperatures and scaling factors used in AIMD simulations are reported in the Supporting Information.

The sampling of the T-p space is illustrated by a two-dimensional potential energy diagram (Figure 2). The potential energy (in eV/atom) of each frame is labeled according to the color scheme included in Figure 3.2. At lower temperatures (e.g., 100 K), the fluctuations in both temperature and pressure are limited, manifested by the clustering of data points. In contrast, the T-p range of the frames obtained at a higher temperature, i.e., near or above the melting point temperatures, are far more dispersed and at higher potential energies. Above 2000 K, the trajectories in T-p space display highly chaotic behavior and include the highest potential energy in the training set. According to Niu *et al.*¹³², adequate sampling of T-p space provides critical information to train NNPs so that future interpolation and extrapolation of the potential energy surface can be reliably reproduced. Given the even distribution of our sampling, no extra weight was added to any specific parts of the potential energy diagram.

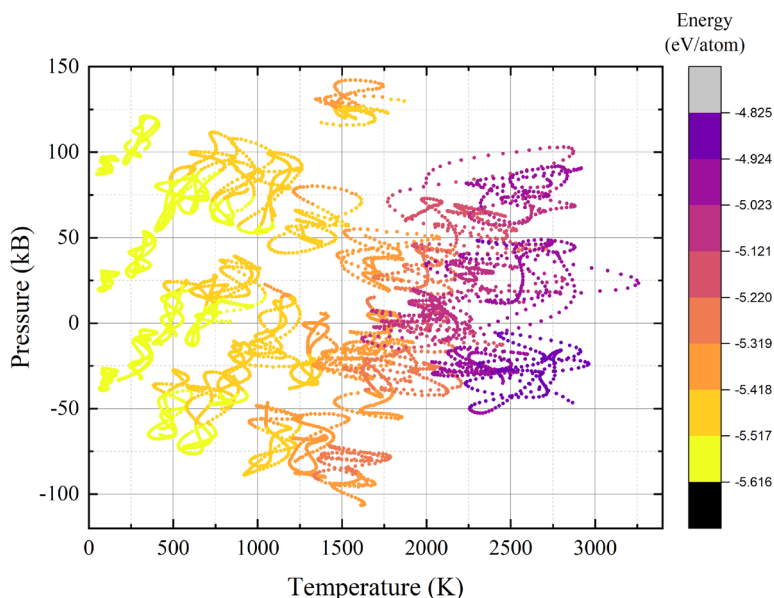


Figure 3.2 A two-dimensional potential energy surface diagram illustrates the T-p parameter space sampling for FCC Ni. The potential energy values (eV/atom) are shown in different colors.

3.2.2 NNP Training

The NNP training was conducted in a multilayer feed-forward (MLFF) neural network implemented in the RuNNer code^{25, 26, 98}. We adopted the atom-centered symmetry function as the structure descriptor¹⁰⁰, as described by Zuo and coworkers¹¹⁸. For Ni systems, two hidden layers were used in the neural network, with 25 neurons in each hidden layer. Other neural network hyperparameters were also tested in this study. The training proceeded for a total of 15 epochs. The weight and bias hyperparameters were determined using the gradient-based iterative optimization approach based on the data from AIMD simulations. The global adaptive Kalman filter^{135, 136} were employed to minimize the root mean squared error (RMSE) of the total energies and forces. 10% of the dataset was used for testing purposes. The atomic coordinates were used as input to calculate energy and forces. Then, the estimated value was validated with quantum mechanical calculations. Lastly, the training and testing set errors were used to determine if overfitting occurred.

3.2.3 Molecular Dynamics Simulations

All MD simulations were performed using the LAMMPS code¹³⁷. The NNP was implemented within LAMMPS using the n2p2 library of functions¹³⁸. Structure visualization and data analyses of phase transition are accomplished by using OVITO¹³⁹. A time step of 0.5 fs was used for all MD simulations. For solid-phase FCC Ni, we used the 5×5×5 Ni supercell (with 500 Ni atoms) as the initial structure (Figure 3.3A). The NPT ensemble was employed to investigate the thermal expansion of the Ni system. Simulations were performed for 2,000,000 steps (1 ns of simulated time) at each temperature; then, the time-averaged values were calculated every 200,000 steps for analyses.

The FCC Ni crystal was first melted at 2295 K using the AIMD method to evaluate liquid-phase Ni structures. Next, a supercell with 2916 Ni atoms was constructed as the initial MD configuration (Figure 3.3B). Again, the NPT ensemble was used in MD simulations at atmospheric pressure, which was maintained with a barostat. The liquid system was equilibrated for 2,000,000 steps at the targeted temperatures. The radial distribution function (RDF) was obtained using the data generated from the last 500,000 steps. The mean square displacement (MSD) was calculated by running an additional 600,000 steps using an NVT ensemble, using the final frame from the prior simulation to resume subsequent simulations.

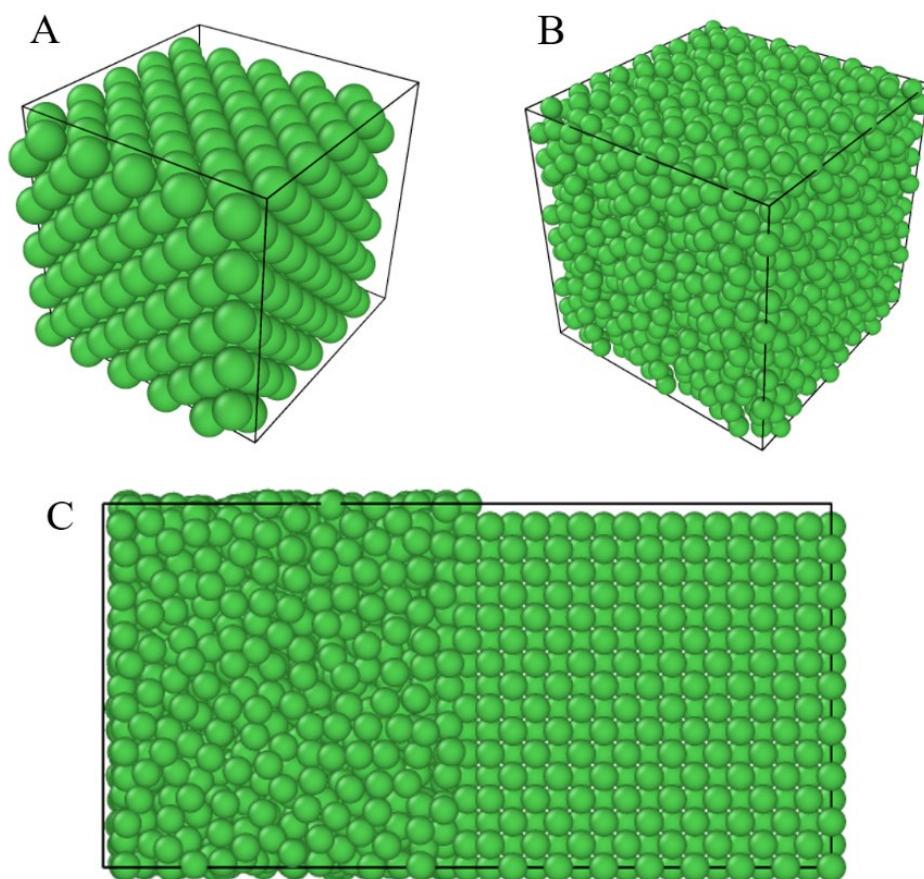


Figure 3.3 Initial structures for solid phase (A), liquid phase (B), and phase transition (C) simulations.

A two-phase interface method was employed to conduct phase transition simulations^{111, 140}. Initially, a solid-liquid interface was generated by melting half of an $8 \times 8 \times 16$ FCC Ni supercell (4096 atoms) at 2200 K with the (100) facet exposed to the liquid phase. However, the orientation of the simulation box has a weak effect on the predicted melting point¹¹¹. The supercell size was adjusted to match the experimental liquid Ni density at 7.81 g/cm^3 ¹⁴¹. The two-phase model and the solid-liquid interface are illustrated in Figure 3.3C. MD simulations based on the NVT ensemble were performed at 1700 K, 1726 K, and 1750 K for 3 ns each to generate the trajectories for structural analyses and melting point determination.

3.3 Results and Discussion

3.3.1 Training Results

Figure 3.4 shows the root mean square error (RMSE) of energy (eV/atom) and forces (eV/Bohr) at different training epochs. The energy RMSE decreases as the training proceeds, while the RMSE of forces reaches the minimum at the third epoch and then rises slightly in the subsequent epochs. Overall, the discrepancies between the training set and testing set errors are minor, proving that no overfitting occurred during training. The trained potential generated from the last epoch was used in MD simulations. The respective RMSE for training and testing sets are 0.811 and 0.831 meV/atom, and force RMSEs for training and testing sets are 75.45 and 74.67 meV/Bohr, respectively. Appendix A reports more detailed comparisons between the reference and predicted data.

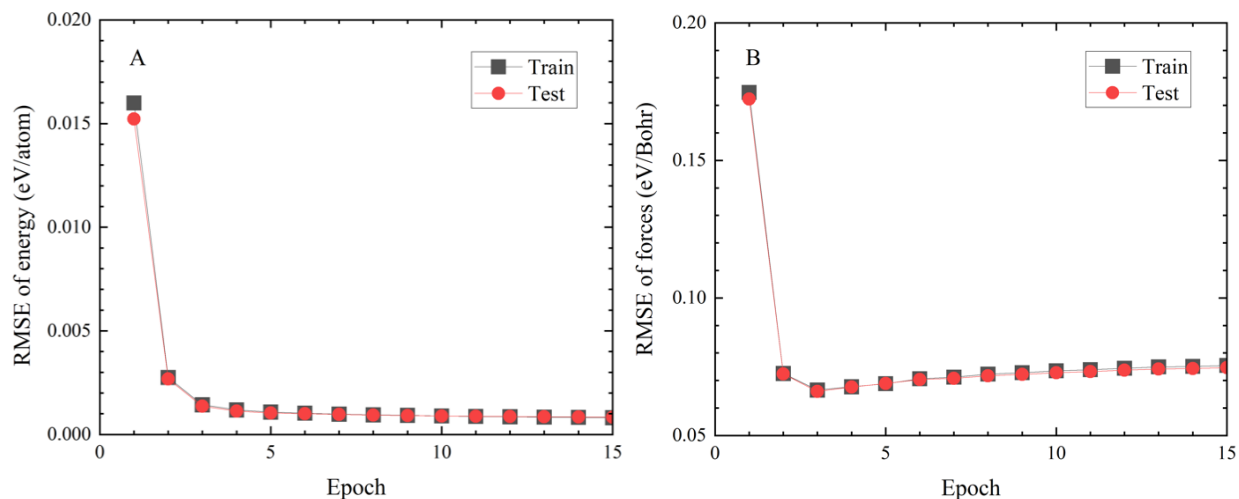


Figure 3.4 RMSE of energy (A) and forces (B) and comparisons between the training (black) and testing (red) sets at different epochs of training.

3.3.2 Structure Properties

First, the lattice constant and thermal expansion coefficient of FCC Ni consisting of 500 atoms were predicted using the MLP developed above. The time-averaged lattice constants collected over 1 ns of simulation trajectories were calculated. The lattice constant fluctuates throughout the trajectory. Hence, the time-averaged lattice constants based on every 0.1 ns interval were also calculated to estimate the errors. The predicted lattice constants over 0–800 K are shown in Figure 3.5. These absolute lattice constant values agree with the literature¹⁴², although we also noted that the fluctuation gradually increases with the simulation temperatures.

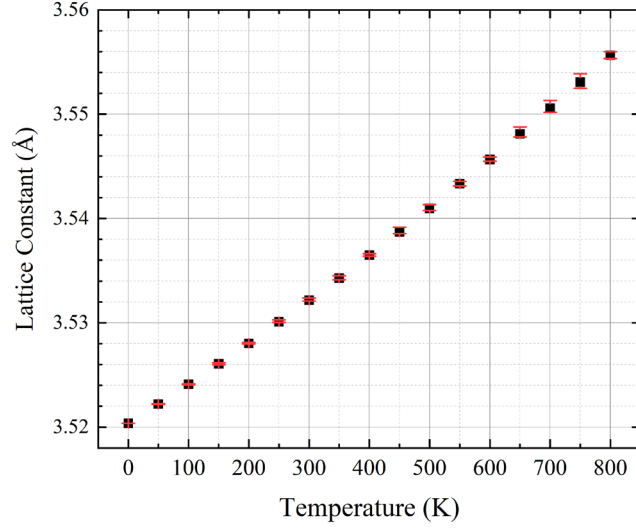


Figure 3.5 Lattice constants (Å) predicted using the NNP for 0–800 K. Simulation results are represented by black squares. The error bars are shown in red.

Using the predicted lattice constants, the thermal expansion coefficient (α) was calculated according to Eqn. (3.1) ¹⁴³:

$$\alpha = \frac{1}{3V_e} \left(\frac{\partial V_e}{\partial T} \right), \quad (3.1)$$

where V_e is the equilibrium volume and T is temperature. Here, the temperature derivative was approximated with Eqn. (3.2):

$$\frac{\partial V_e}{\partial T} = \frac{\Delta V_e}{\Delta T}, \quad (3.2)$$

and estimate the equilibrium volume using the averaged lattice constant over the temperature interval:

$$V_e = \frac{L_2^3 - L_1^3}{2}, \quad (3.3)$$

where L_1 and L_2 represent the lattice constants corresponding to T_1 and T_2 , the two adjacent temperatures.

The calculated thermal expansion coefficients are plotted in Figure 3.6. The error bars for each thermal expansion coefficient account for uncertainties due to lattice fluctuations. For comparison, literature data based on phonon calculations (red dots¹⁴³), the EAM potential (blue triangle¹⁰⁸), and different experimental methods (green inverted triangles and purple diamonds^{142, 144}) are also included. For the predicted thermal expansion coefficients based on NNP, the temperature dependency is not as strong as other methods. Also, the thermal expansion coefficients are underestimated. However, we recognize that the comparison is also sensitive to volume fluctuations. The accompanying error bars indicate that the discrepancies with experiments and other computations fall within the calculation uncertainty.

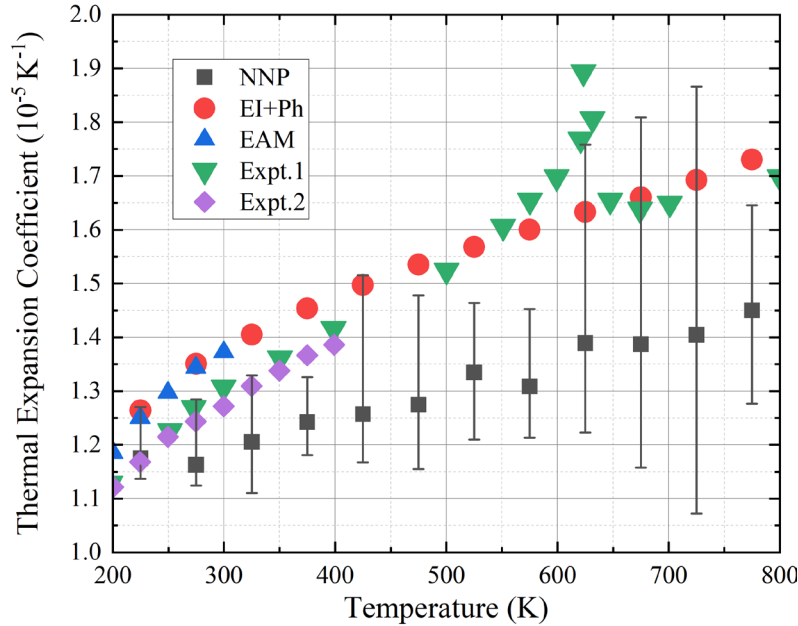


Figure 3.6 Thermal expansion coefficients between 200-800 K obtained using NNP and from literature data: red dots (EI+Ph¹⁴³), blue triangles (EAM¹⁰⁸), green triangles (Expt. 1: Touloukian¹⁴⁴), and purple diamonds (Expt. 2: Touloukian¹⁴²).

The RDFs at different temperatures were obtained to predict the structure of liquid phase Ni using the same MLP. Figure 7 plots RDFs at four temperatures near and above the melting point of Ni. The peaks and valleys predicted by the EAMFS potential¹¹⁴ are labeled with symbols colored according to the temperature. According to Figure 3.7, the trained MLP captures the key features of liquid Ni quite accurately for all temperatures considered. Notably, the magnitudes and locations of each peak and valley fit well with the values predicted by the EAMFS potential. Note as the temperature increases. The peak values decrease while trough values increase, confirming that the structure of Ni becomes more disordered. Moreover, all peaks and troughs except the first peak shift toward the right, consistent with how liquid structures vary with temperature. Trends similar to those observed by these NNP-based MD simulations have also been reported by simulations using the EAMFS potential.

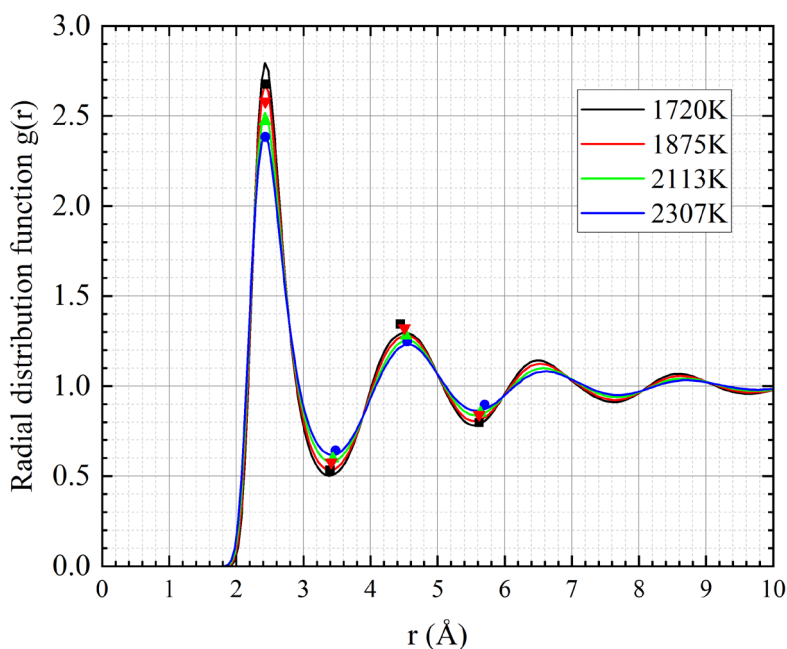


Figure 3.7 RDFs as a function of temperature predicted using the reported NNP. The peak and valley locations predicted for each temperature using the EAMFS potentials¹¹⁴ are included for comparison.

3.3.3 Physical Properties

The diffusivities of Ni atoms in the liquid phase at different temperatures were calculated based on the MSDs expressed by Eqn. (3.4):

$$\langle X^2(t) \rangle = \langle |X(t) - X(0)|^2 \rangle = \frac{1}{N} \sum_{i=1}^N |X^i(t) - X^i(0)|^2, \quad (3.4)$$

where $X^i(t)$ is the position vector of the i -th particle at time t , and N is the number of particles in the system. Then, the diffusivity D is calculated using Eqn. (3.5):

$$D = \frac{\langle X^2(t) \rangle}{6t}, \quad (3.5)$$

where $\frac{\langle X^2(t) \rangle}{t}$ is determined using the slope of the $\langle X^2(t) \rangle$ versus t plot (Figure 3.8A). The slope of the $\langle X^2(t) \rangle$ versus t relationship increases with temperature.

The diffusivity at each temperature was calculated by extracting the slope of the mean square displacement at that temperature (Table 3.2 and Figure 3.8B). The diffusivities predicted by the NNP are systematically lower than the predictions by the EAMFS¹¹⁴ and EAM⁵⁵ potentials, but the discrepancies are much smaller when compared with the experimental results^{145, 146}. The comparison also improves at higher temperatures except at high temperatures (e.g., 2307 K).

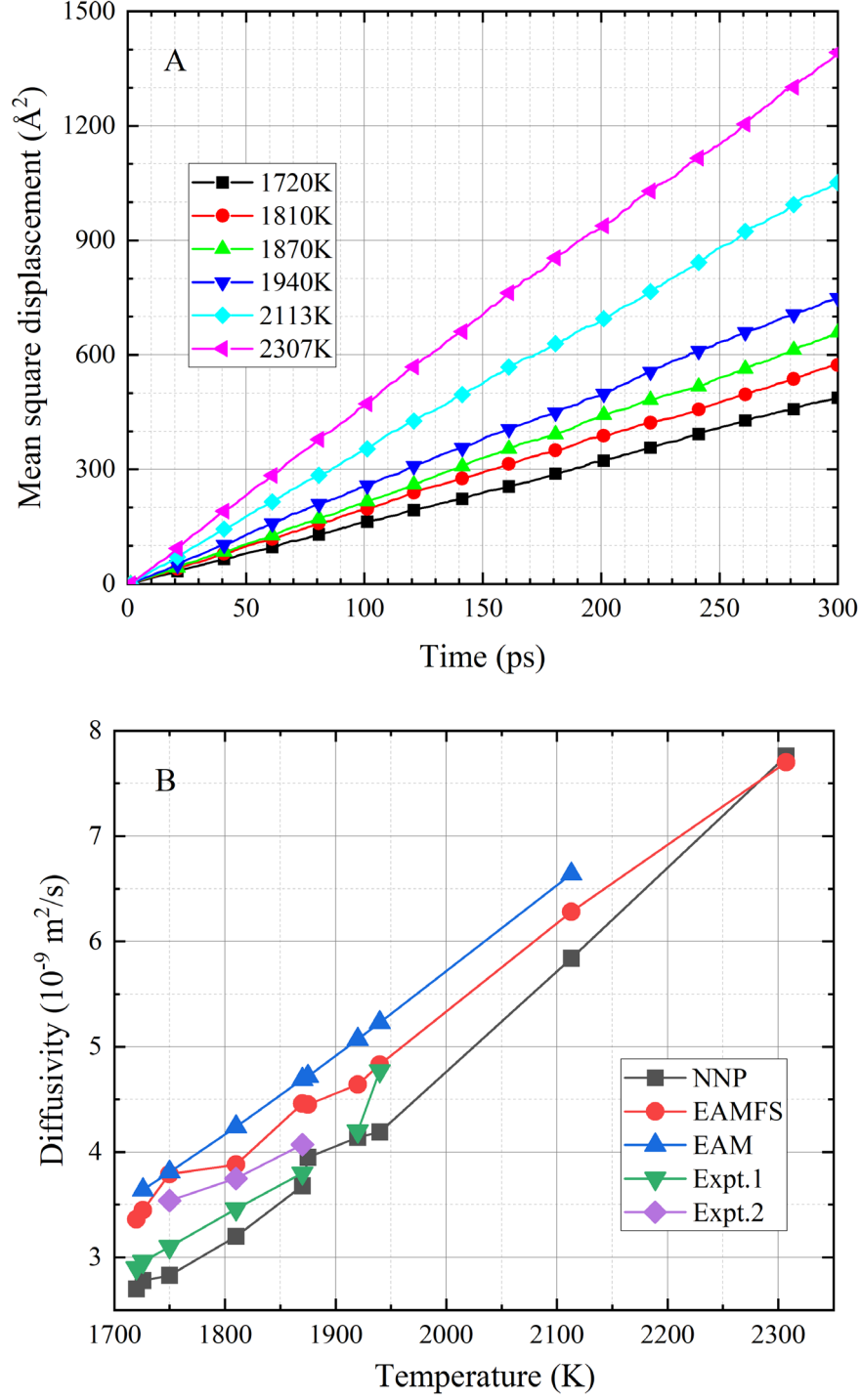


Figure 3.8 (A) Mean square displacements ($\text{\AA}^2$) of Ni atoms in the liquid phase at various temperatures. (B) Calculated diffusivity ($10^{-9} \text{ m}^2/\text{s}$) in this work (black squares) compared with the literature: red dots (EAMFS¹¹⁴), blue triangles (EAM¹⁴⁷), green triangles (Expt. 1: Meyer¹⁴⁶), and purple diamonds (Expt. 2: Meyer¹⁴⁵).

Table 3.2 Diffusivity ($10^{-9} \text{ m}^2/\text{s}$) of Ni atoms in the liquid phase.

Temperatures (K)	NNP	EAMFS or EAM	Experiment
1720	2.70	3.36 ^a	2.90 ^c
1726	2.78	3.45 ^a	2.96 ^c
		3.64 ^b	
1750	2.83	3.79 ^a	3.10 ^c
		3.81 ^b	3.54 ^d
1810	3.20	3.88 ^a	3.46 ^c
		4.24 ^b	3.75 ^d
1870	3.68	4.46 ^a	3.80 ^c
		4.69 ^b	4.07 ^d
1875	3.95	4.45 ^a	
		4.72 ^b	
1920	4.14	4.64 ^a	4.20 ^c
		5.07 ^b	
1940	4.19	4.83 ^a	4.77 ^d
		5.23 ^b	
2113	5.84	6.28 ^a	
		6.64 ^b	
2307	7.76	7.70 ^a	

(^aNenuwe 2018¹¹⁴, ^bLu 2012¹⁴⁷, ^cMeyer 2015¹⁴⁶, ^dMeyer 2008¹⁴⁵)

3.3.4 Melting and Crystallization of FCC Ni

In this section, the NNP was applied to simulate the melting and crystallization of the nickel system using the two-phase interface method^{111, 140}. The phase transition phenomena were analyzed by monitoring the number of atoms along the axial direction (perpendicular to the solid-liquid interface), as the solid phase is identified in regular oscillations. At the same time, the liquid is recognized by irregular fluctuations with smaller amplitudes. This way, the melting point

temperature can be estimated by monitoring the growth or contraction of the liquid fraction during simulations at a series of targeted temperatures.

Snapshots depicting the configurations of the initial and equilibrated states at three temperatures (1700 K, 1726 K, and 1750 K) and the corresponding number of atoms along the axial direction are illustrated in Figure 3.9. By dividing the simulation box into 200 slices equally and calculating the number of atoms inside each bin, the fractions of the solid and liquid phases of the initial state (total axial length is 58.92 Å) could be distinguished via the oscillation pattern and amplitudes described above. For the equilibrated conditions at the targeted temperatures, the fraction of the liquid phase was determined using the distance between the two valleys where the number of atoms begins to be ≤ 5 . The ratio between the liquid and solid dimensions corresponding to each condition is labeled in Figure 3.9.

The liquid phase completely disappears at 1700 K, 26 K below the experimentally determined melting point (1726 K), and the entire Ni system crystallizes into the ordered FCC lattice (Figure 9B). While at 1726 K (Figure 3.9C), the percentage of the liquid fraction is 6.6% lower than the initial structure (Figure 3.9A), which means the melting point predicted by NNP-based MD simulations is at a higher temperature. With the temperature raised to 1750 K (Figure 3.9D), the fraction of the liquid phase expands by approximately 6.0% compared to the initial state. This observation suggests that the melting point is lower than 1750 K. Overall, we can narrow the melting point down to a range between 1726 K and 1750 K, roughly within 24 K of the experimental value.

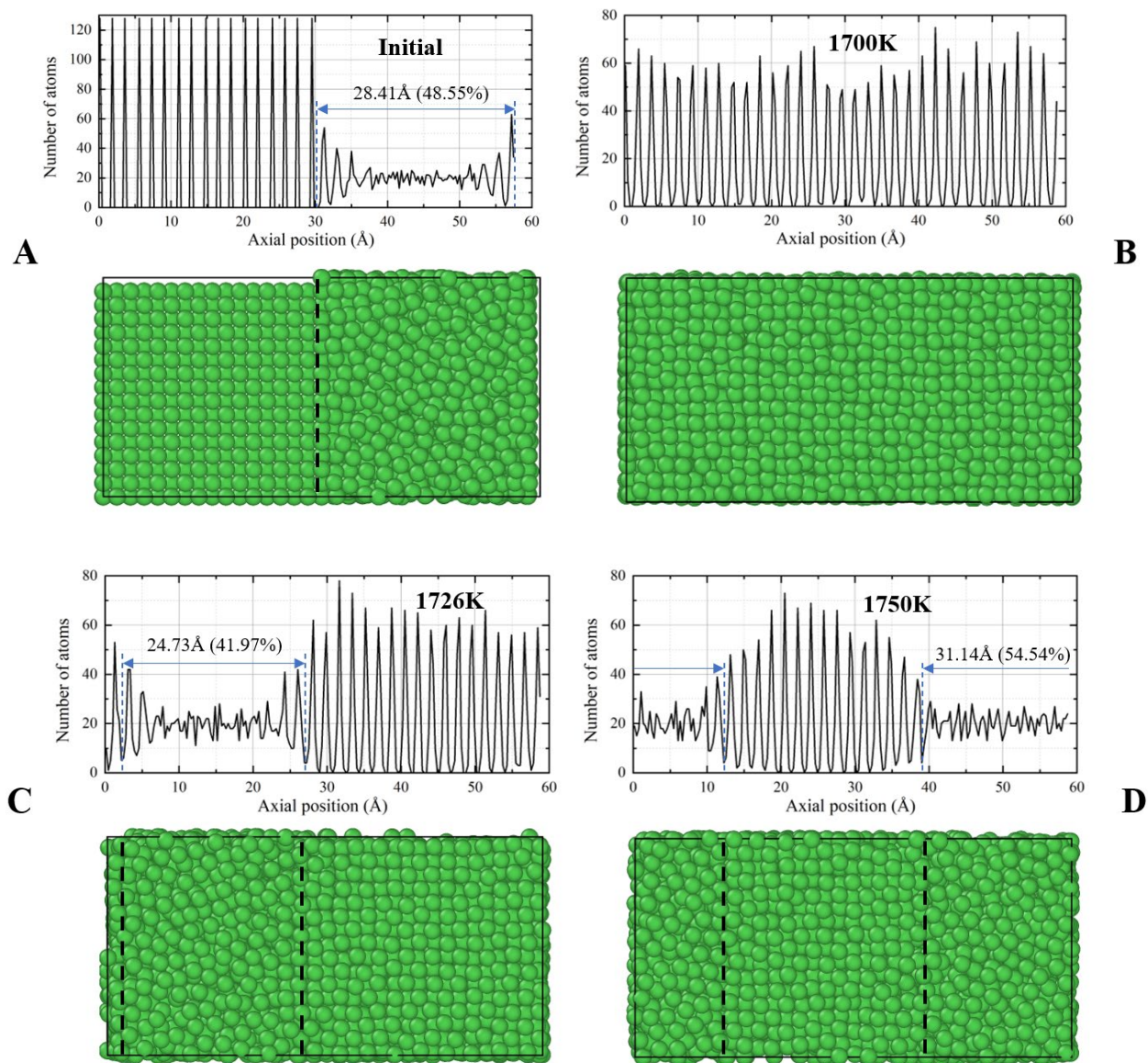


Figure 3.9 The number of atoms along axial direction and snapshots of the simulated solid-liquid two-phase systems for (A) initial configuration, (B) 1700 K, (C) 1726 K, and (D) 1750 K.

3.4 Conclusion

With a data set generated from AIMD, a high-dimensional neural network potential was developed based on a multilayer feed-forward neural network and the atom-centered symmetry function descriptors. This unified NNP was employed to simulate the behaviors of both solid and

liquid phase Ni and the melting and crystallization processes. MD simulations based on the developed potential accurately predict the lattice constants and thermal expansion coefficients in the solid phase. Radial distribution functions and mean square displacements were well described for the liquid phase. Moreover, this potential can predict the melting point within 24 K of the literature value. These validations demonstrated that a stable and reliable NNP could be developed to study crystals under a wide range of thermodynamic conditions. More importantly, the training protocol paves the way to establish MLPs for investigating other materials and thermodynamic behaviors.

Chapter 4 - Thermodynamic Stability Prediction of Icosahedral Boron Allotropes

Chapter 4 is reproduced with permission from:

Deng H.; Liu, B. “Predictions of Boron Phase Stability Using an Efficient Bayesian Machine Learning Interatomic Potential”, The Journal of Physical Chemistry Letters. DOI: 10.1021/acs.jpcclett.4c00322.

4.1 Introduction

Since its first high-purity production in 1909¹⁴⁸, elemental boron has attracted much attention owing to its characteristic icosahedral B₁₂ motif. Moreover, boron compounds derived from the rhombohedral α -boron lattice, such as B₆O¹⁴⁹⁻¹⁵¹, B₁₂P₂^{152, 153}, have broad applications in superhard material manufacturing and energy conversion devices¹⁵².

Among boron's numerous allotropes⁵⁸, α -B (rhombohedral, B₁₂), β -B (rhombohedral, B₁₀₅), and γ -B (orthorhombic, B₂₈) all consist of B₁₂ icosahedra as the fundamental building block. The primitive cells of these boron allotropes are shown in Figure 1.4. In α -B (Figure 1.4A), the B₁₂ icosahedra are located at the corners (green) of the rhombohedral unit cell. The lattice of β -B (Figure 1.4B) contains icosahedra at the edges (red) and corners of rhombohedra, and two triply fused icosahedra (yellow) connected with an interstitial boron atom at the center¹⁵⁴. The γ -B unit cell (Figure 1.4C) comprises B₁₂ icosahedra and B₂ pairs arranged in NaCl-type structure where B₁₂ icosahedra and B₂ pairs are located at ‘anions’ and ‘cations’ positions, respectively⁶¹.

While the γ -B phase dominates at high pressures, the relative phase stability between α -B and β -B at moderate conditions remains debatable⁵⁹. Moreover, the experimental determination

of phase stability between α -B and β -B is challenged by a substantial phase transition kinetic barrier⁶⁰. It has been hypothesized that partial occupancy helps stabilize β -B at low temperatures^{60, 155-158}. The phase diagram produced by Oganov *et al.*⁶¹ suggests that β -B be more stable than α -B at near ambient pressures. White *et al.*¹⁵⁴ performed a thermodynamic cycle analysis, showing that α -B becomes stable only when external pressures are applied. Parakhonskiy *et al.*⁶³ reported a phase diagram based on more than 30 experiments conducted at 3 ~ 14 GPa and 1073 ~ 2423 K to find that, at the ambient pressure, α -B becomes more favorable than β -B below 1000 K. Ogitsu *et al.*¹⁵⁹ tried to reconcile Parakhonskiy's finding by incorporating corrections of partial occupancy. However, β -B remains to be the predicted ground state. Shang *et al.*¹⁵⁵ performed first-principles phonon calculations and confirmed that α -B is more stable than the *defect-free* β -B at temperatures up to 1388 K.

Although density functional theory (DFT) has been widely used to obtain the free energies of pure boron phases^{61, 159}, previous studies showed that standard ground state energies are inadequate to account for the temperature effect on lattice kinetic energies and cell volume expansion (or contraction). *Ab initio* molecular dynamics (AIMD) compensates for such a deficiency. However, the high computational expenses of quantum mechanical calculations often limit the length scale and efficiency to extract key thermodynamic properties, such as the phonon dispersions of β -B lattices. Classical MD simulations considerably improve the computational efficiency of large molecular systems. However, training interatomic potentials for reactive systems can be highly tedious, and the predictivity of physics-based potential deteriorates notably for multi-component, structurally complex systems, including the icosahedral boron allotropes⁴⁸.

In the last decade, interest in machine learning potentials (MLPs) has surged due to their flexibility with training data composition and the techniques employed. Meanwhile, the

computational efficiency is retained, allowing simulations on systems with up to millions of atoms⁴¹. Numerous methodologies, such as neural network potential (NNP)²⁵⁻²⁸, spectral neighbor analysis potential (SNAP)²⁹⁻³², Gaussian approximation potential (GAP)³³⁻³⁵, and moment tensor potential (MTP)³⁶⁻³⁸, have been established and yielded promising outcomes, including pure boron systems. For instance, Podryabinkin *et al.*¹⁶⁰ employed an actively trained MLP to accelerate crystal structure prediction. Huang *et al.*¹⁶¹ constructed a high-dimensional NNP by sampling 15293 distinct configurations to explore the potential energy surface of β -B and elucidated the principles governing the filling of the interstitial sites. Deringer *et al.*¹⁶² constructed a Gaussian approximation potential aided by the random structure searching algorithm to investigate the frustrated β -rhombohedral boron structures. Yang *et al.*¹⁶³ recently studied a novel 24-atom cubic boron and its mechanical properties using the Gaussian approximation potential.

Because of boron's unusual chemical bonding and structural characteristics, the training set must include dedicated structures, energies, and forces representative of each boron phase. Hence, the composition and size of the training set become complex and unmanageable quickly. Recently, a Bayesian model-driven Gaussian process (GP) has shown outstanding performance in MLP-based MD simulations. Most notably, the sparse Gaussian process (SGP) model developed by Kozinsky and coworkers^{41,42} enables a significant reduction of training data size. Furthermore, by incorporating an active learning scheme, the most relevant structures are produced on the fly, accompanied by MD simulations. This practice bypasses an enormous data set while the representation of the targeted systems is still fulfilled. The reported SGP applications include heterogenous catalysis^{41,45}, phase transition of SiC⁴², and crystallization of GaN¹⁶⁴.

This chapter reports the training procedure and MD results based on SGP potential for icosahedral boron crystals. SGP is a Bayesian model that utilizes sparse datasets containing

information representing atomic environments. When calculating the forces acting on a specific atom, the similarity between the environment of the targeted atom and the sparse set is quantified. The forces of interest are the averaged forces in the sparse set, weighted by the distances between the target and sparse set environments. Comparisons between SGP and DFT calculations show that the trained SGP potential accurately predicts the structural, thermodynamic, and vibrational properties of all major boron phases. A P-T phase diagram of α -B, β -B, and γ -B was produced, enabling a deeper understanding of the temperature and pressure effects on boron crystals' thermodynamic stabilities. In particular, the simulations of β -B with B13-type defect (i.e., B₁₀₄) indicate that such defect can stabilize β -B, thus shifting the relative phase stability against α -B. More importantly, this work showed that the SGP potentials trained on a data set devoid of defective structures yield correct thermodynamic stability prediction of defective systems. In this work, the SGP training was conducted using the FLARE program^{39, 40} (version 1.3.3). Detailed information on SGP methodology can be found in papers published by FLARE developers^{41, 42}.

4.2 Methods

The training process is illustrated in Figure 4.1A, which can be described in two stages: (i) on-the-fly (OTF) MD and (ii) offline training. Detailed information, script for OTF and offline training, generated training set, and final SGP potential can be found in Appendix B.

4.2.1 Training Set Generation by OTF MD

To generate a training set, defect-free α -B, β -B, and γ -B supercells with different scaling factors (SFs) were employed to model zero, tensile, and compression stresses. OTF MD simulations were initiated as the engine to generate the training set. All OTF MD simulations were

run for 1,000 steps with an NVE ensemble. The OTF MD simulations also enable the update of SGP potentials. Once the configuration is updated by a new MD step based on the current potential, the algorithm will estimate the uncertainty of the new structure based on the inherent quantification mechanism of SGP and then determine if the predetermined uncertainty threshold ($0.01 \text{ eV/\AA}$) is exceeded. The simulation temperature was raised every 200 steps, starting from 50K, with incremental increases to 500K, 1000K, 1500K, and 2000K. DFT calculations were performed on the newly discovered high-uncertainty structures. The structures and DFT-calculated total energies and forces will be employed to retrain the SGP potential. Representative environments will be appended to the sparse set for future force predictions. DFT calculations involved in OTF MD simulations and validation were performed using the Vienna *Ab Initio* Simulation Package (VASP, version 5.4.4) ^{8, 9}. The computational details of DFT and a brief introduction of the SGP methodology are reported in the Appendix B.

Once OTF MD simulations were completed, all structures with high uncertainty were combined with DFT-calculated energies and forces for training set preparation. The sequence of structures in the training set is shuffled. The training set composition is depicted in Figure 4.1B. The final training set contains 183 structures, significantly fewer than the typical training set size required by MLPs, such as ANN-based potentials ^{28, 122}. The main reason for generating such a small training set is that all OTF MD simulations are implemented on the solid phase. Since temperatures are lower than melting point and supercells are fixed, solid structures didn't show significant deformation during OTF MD simulations. Without high uncertainty structures frequently generated by MD, DFT calls during all OTF training became rare. In this way, there are fewer DFT calculations we can collect from OTF training, and then the size of the final training set becomes small. We noted that structures with larger SFs were generated at higher frequencies

because boron atoms show larger displacements from their equilibrium positions in expanded cells. Also, OTF MDs of β -B invoked more DFT calls than the other two allotropes due to the high complexity of the unit cell. Figure B.1 shows the temperature change and DFT calls during OTF MD of β -B with an SF of 1.05, yielding the largest number of DFT calls. Many DFT calls happen immediately after the temperature increases. Because the kinetic energy of the system increases as temperature rises, boron atoms become likely displaced from their equilibrium positions in the lattice, and so do the uncertainties of the updated structures.

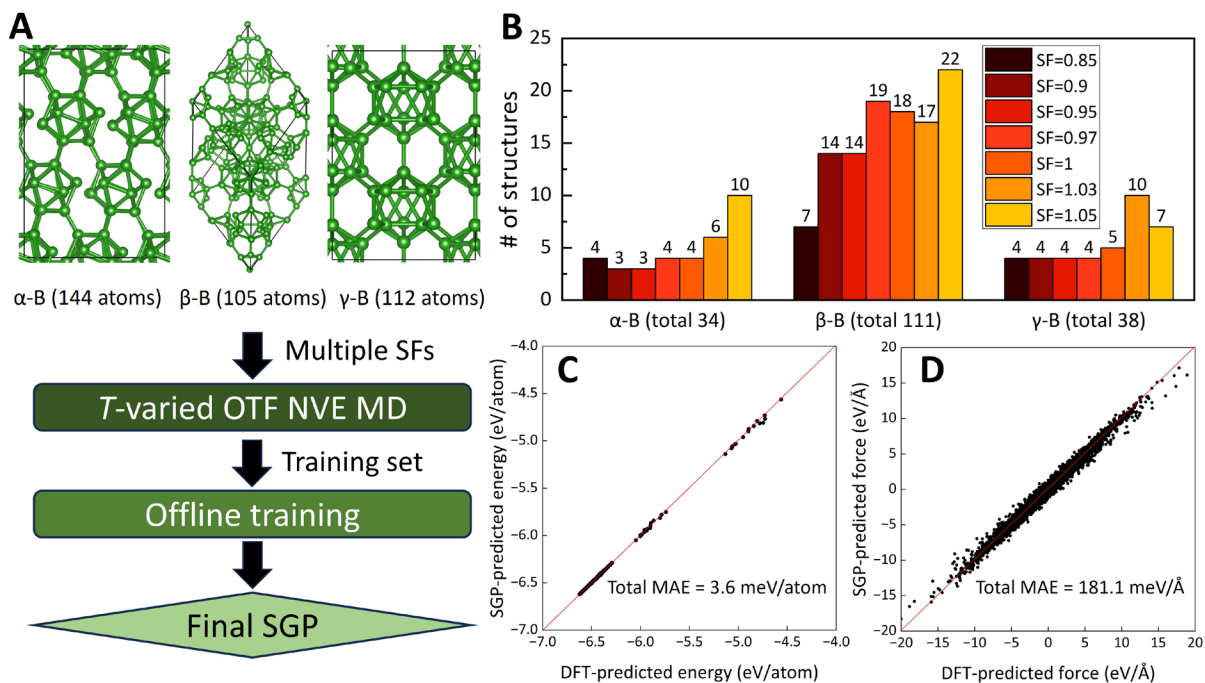


Figure 4.1 (A) Workflow for SGP potential training. Boron crystal structures are visualized using the VESTA software¹⁶⁵. Black solid lines indicate cell boundaries; (B) Composition of the training set for offline training; (C) DFT- versus SGP-predicted energies; (D) DFT- versus SGP-predicted forces.

4.2.2 SGP Potential Offline Training

Offline training was performed based on the training set generated by OTF MD to obtain the final SGP potential. Offline training estimates uncertainties of training set structures in sequence. When a high-uncertainty structure (exceeding the threshold of 0.01 eV/Å) is encountered, the SGP will start a new iteration of training and add representative environments to the sparse set.

To fine-tune the performance of the final SGP potential, we screened three predetermined parameters, including cutoff radius (R_{cut}), radial fidelity (n_{max}), and angular fidelity (l_{max}), which are employed to build the Atomic Cluster Expansion (ACE) as the environment descriptor¹⁰¹. Over 300 offline training were performed, resulting in a parameter set of $R_{cut} = 5.5$ Å, $n_{max} = 14$, and $l_{max} = 6$ based on the force mean absolute error (MAE). Detailed screening results are reported in Figures B.2 and B.3. The energies and forces of the training set structures produced by DFT and the final SGP potential are presented in Figures 4.1C and 4.1D, with the corresponding total MAEs being 3.6 meV/atom and 181.1 meV/Å. To reveal the training process, the frame MAEs based on atomic energies and forces extracted from every frame in the training set are shown in Figures B.4 and B.5 in the Supporting Information. As offline training proceeds, updates to the sparse set become less frequent, and the frame MAEs of atomic energies and forces decrease and become stable. The training input files and final SGP potential are included in Supporting Information.

4.2.3 Molecular Dynamics Simulation

All SGP-based MD simulations were performed using the LAMMPS code¹³⁷. The energy of each boron crystal structure was minimized using the SGP potential to obtain equation of state (EOS), enthalpy differences (ΔH), radial distribution functions (RDF), and phonon dispersions.

The lattice expansions and phase diagram were produced from NPT simulations using a time step of 0.5 fs. Each MD run lasts 1,000,000 steps (i.e., 0.5 ns of simulation time). The time-averaged values were calculated every 200,000 steps, and results from the last interval were collected for analysis.

4.3 Results and Discussion

4.3.1 Thermodynamics and Structural Properties

The performance of the SGP potential was calibrated using DFT and relevant experimental results. The EOS curves for α -, β -, and γ -B are shown in Figure 4.2A. The solid data points were calculated using the SGP potential. Seven volumes were selected for each allotrope. Good agreements of fitted curves indicate that the SGP potential captures the correct total energies of all boron allotropes at varying levels of isotropic compressions and expansions. The fitted 4th-order Birch-Murnaghan (BM) EOS¹⁶⁶ coefficients and MAEs and mean absolute differences (MADs) of MLP predictions are reported in Table S2.

When external pressures were isotropically applied to boron crystals, distinct enthalpic responses are expected from the β - and γ -phases because β -B tends to be more stable at low pressures, whereas γ -B is the high-pressure phase. Oganov *et al.*⁶¹ depicted the relative stability of boron phases using relative enthalpies (ΔH) based on first-principles calculations. As illustrated in Figure 4.2B, when the applied pressures are below 25 GPa, the ΔH yielded by the SGP potential is within 5 meV/atom of the quantum mechanical predictions. Such a pressure range is adequately broad for most practical considerations, such as high-pressure crystal synthesis. Nevertheless, we also noted that ΔH will be overestimated at pressures above 40 GPa for both β -B and γ -B, as indicated by the H-P diagram in Figure B.6. Overestimation is likely associated with insufficient

sampling of high-pressure structures. Figure 4.1B suggested that all boron structures with the smallest scaling factors (SF) are more likely to receive fewer DFT calls during the OTF training.

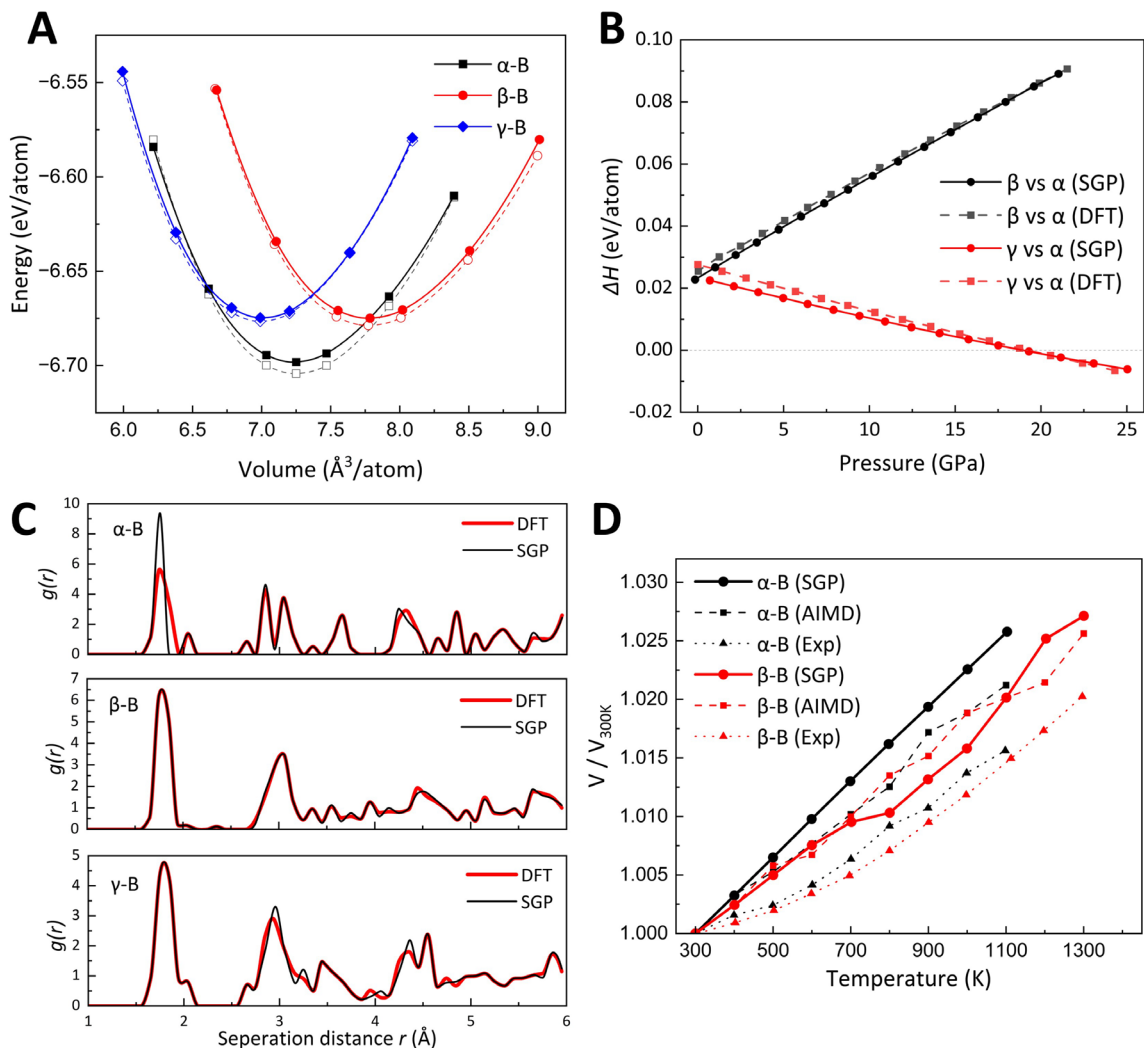


Figure 4.2 (A) Equation of state for α -B, β -B, and γ -B. Dashed lines (DFT) and solid lines (SGP) were fitted to the 4th-order Birch-Murnaghan formula. (B) Enthalpy differences (ΔH) for β -B (black) and γ -B (red) versus α -B. (SGP: solid lines; DFT: dashed lines) (C) Radial distribution functions of solid-state α -B, β -B, and γ -B. RDF analyses were performed using OVITO ¹³⁹. (SGP: black; DFT: red) (D) Volume expansions of α -B (black) and β -B (red). SGP: solid lines with circles; AIMD: dashed lines with squares; experimental volume expansion of α -B ¹⁶⁷ and β -B ¹⁶⁸: dotted lines with triangles.

The SGP potential was also used to compute the RDFs for all boron allotropes as another metric to evaluate the performance of the potential. As shown in Figure 4.2C, the positions of all major peaks match well against DFT calculations, confirming that the SGP potential produces reliable intra- and inter-icosahedral B–B bond lengths.

The volume expansion of α -B and β -B were computed by performing NPT simulations, in which the shapes of the rhombohedral cells are fixed. The LAMMPS simulations (circles in Figure 4.2D) were accompanied by AIMD simulations (squares in Figure 4.2D) to calibrate the predictions of the SGP potential. For AIMD, simulation cells consisting of 144 and 105 B atoms were employed for α -B and β -B respectively, and each simulation lasted 10 ps due to the high computational costs. Reasonable comparisons were achieved. The SGP potential is also able to recognize that α -B has a higher volume expansion rate than β -B, which is a key factor affecting free energy ordering. Since VASP does not support angle-fixed NPT, disagreements on cell volume calculations may occur. Compared with experimental measurements^{167, 168}, we noted that the SGP potential and AIMD overestimate the volume expansions, especially at higher temperatures. This observation demonstrates inaccuracy originating from quantum mechanical calculations can be inherited during the training of machine learning potentials.

4.3.2 Vibrational Properties

The lattice vibrations in boron crystals were simulated to demonstrate that machine learning potentials can replace expensive quantum mechanical calculations to obtain solid-state vibrational entropies (S) and zero-point energies (ZPE). Herein, phonon dispersions and phonon density of states (PDOS) were calculated using Phonopy¹⁶⁹.

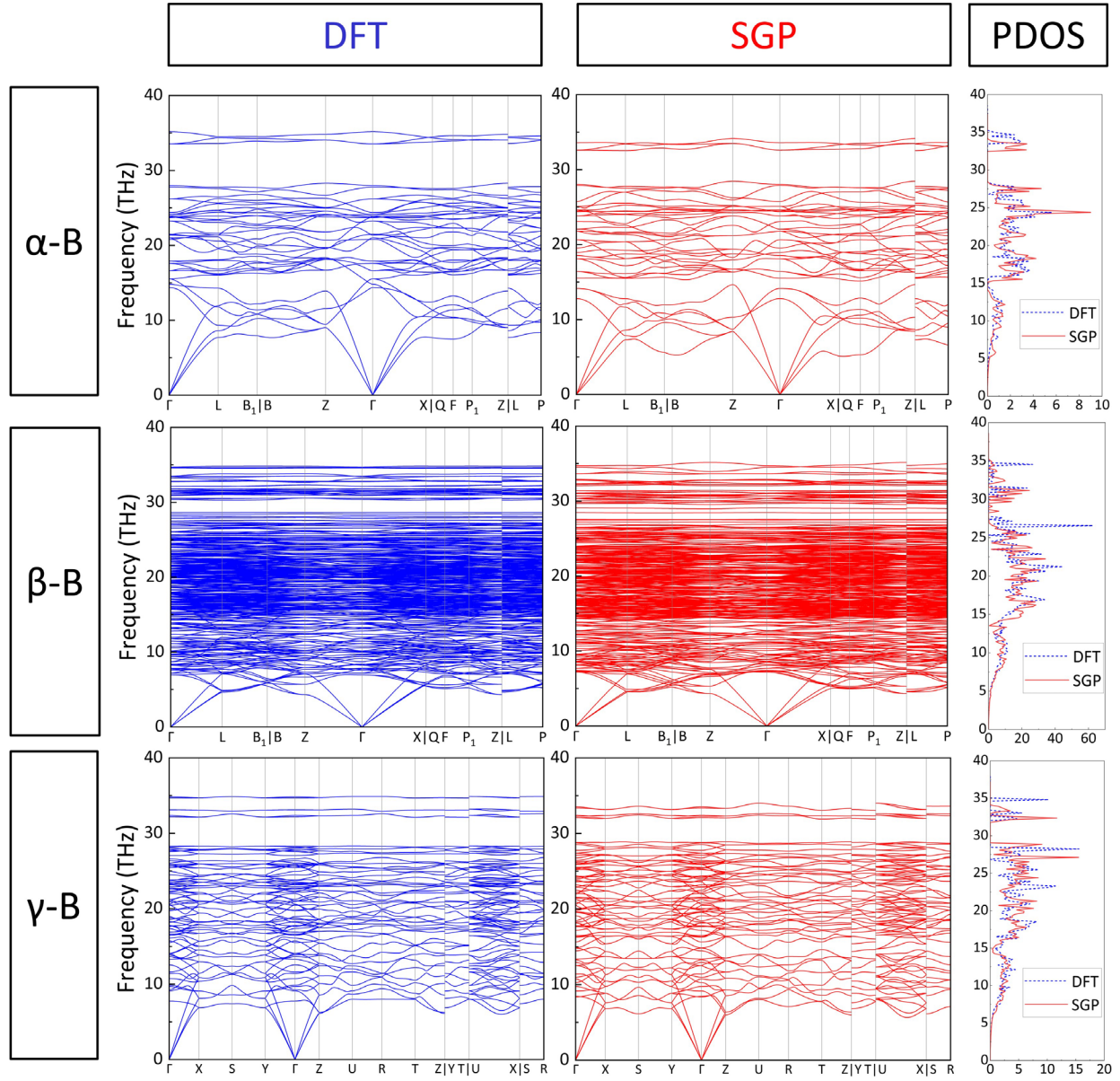


Figure 4.3 Phonon dispersion and phonon density of states (PDOS) of α -B, β -B, and γ -B calculated from DFT (blue) and MD simulations (red).

As shown in Figure 4.3, we observed qualitative resemblance with DFT-calculated phonon dispersions for all boron allotropes. Then, the PDOS, aligned with the corresponding phonon dispersion, was utilized to reveal differences in the fine structures of phonon frequency distributions. The characteristic peaks produced by the SGP potential and quantum mechanical

methods agree well for phonon frequencies below 20 THz. At the same time, larger discrepancies are noted at higher frequencies for all boron allotropes. Hence, the accuracy of the SGP potential was calibrated by calculating the heat capacities and vibrational entropies using Phonopy. Figures B.7A-C indicate that thermodynamic calculations based on the SGP potential will not be significantly impacted even though the PDOS are perfectly aligned with quantum mechanical calculations.

4.3.3 Phase Stability Predictions

For defect-free systems, vibrational entropy plays a central role in dictating the free energies of solid-state boron crystals. Here, the vibrational entropies and zero-point energies were extracted from the MD simulation trajectories using Dump2phonon and its auxiliary post-processing tool Phana^{170, 171}. Unlike Phonopy, where atom displacements are generated manually, Dump2phonon constructs a dynamical matrix from the actual displacements occurring during MD simulations. This approach allows the calculations of vibrational entropies and *ZPEs* corresponding to the prescribed temperatures and pressures. A (2×2×2) supercell was used for MD simulations of all boron allotropes and the accompanying vibrational entropies and *ZPEs* calculations.

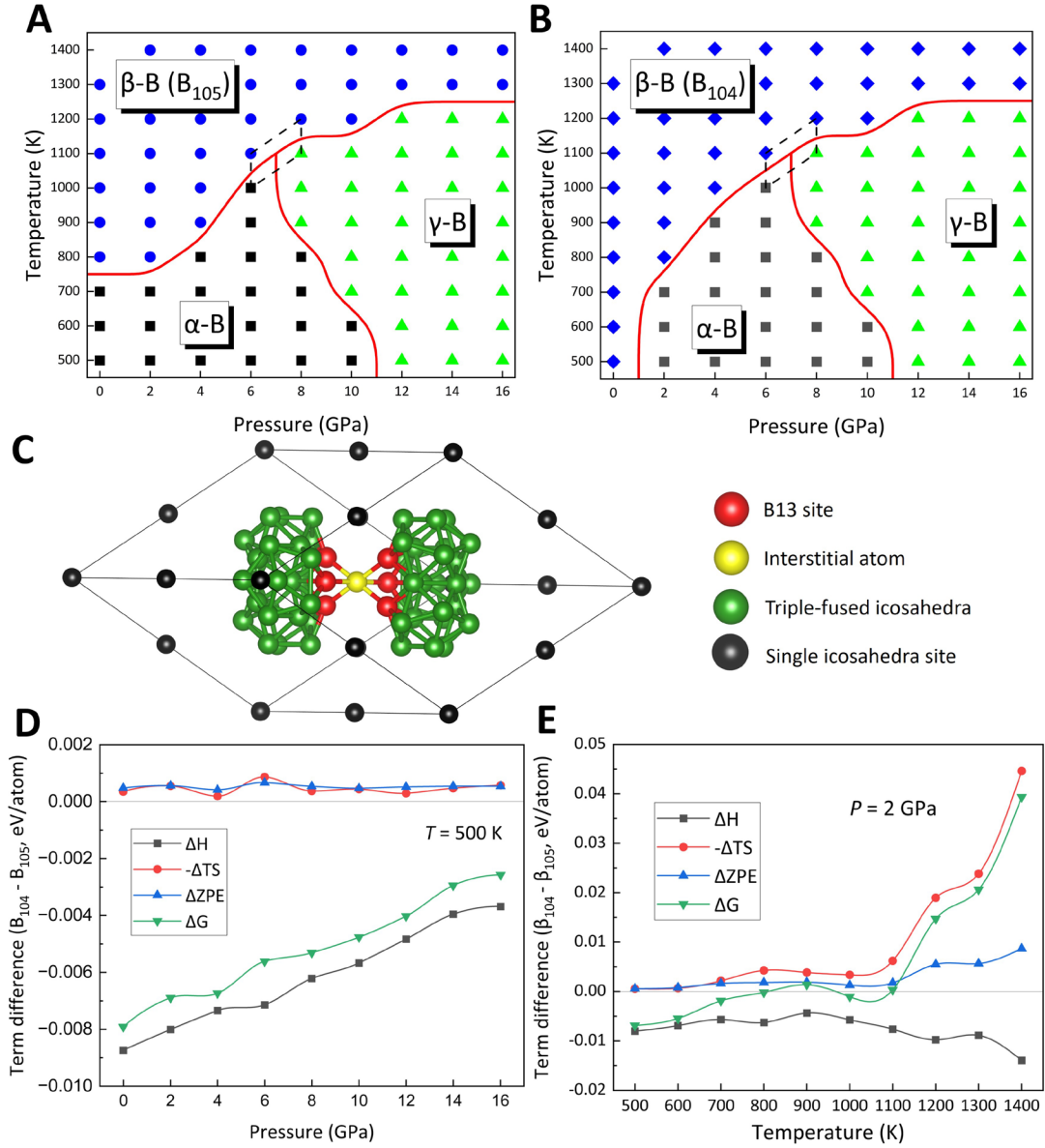


Figure 4.4 Phase diagram predicted by SGP-based MD simulation and phonon calculations considering (A) defect-free β -phase (β - B_{105}) and (B) β -phase with one B13 defect (β - B_{104}). Squares, dots, diamonds, and triangles represent thermodynamically preferable conditions of α -B, β - B_{105} , β - B_{104} , and γ -B, respectively. Red curves and black dashed lines indicate the respective phase boundaries and possible locations of the α - β - γ triple point. (C) Schematic representation of defect-free β -B (B_{105}). Thermodynamic differences (ΔX , $X = H, -TS, ZPE, G$) between B_{105} and B_{104} at (D) $T = 500$ K and (E) $P = 2$ GPa.

The phase diagram of defect-free boron allotropes predicted by SGP-based MD simulations is shown in Figure 4.4A. Qualitatively, the SGP-predicted phase diagram shows agreement with reported experimental and calculational results^{61, 63}. The β -phase dominates the area corresponding to high temperature and low pressure, while γ -B occupies the high-pressure region. The α -B phase shows higher stability (over β -B) under low temperatures and pressures, which has been controversial in some reported phase diagrams^{61, 159}. This phase behavior is mainly due to the use of defect-free β_{105} in the MD simulations. As discussed previously, the introduction of partial occupancy sites stabilizes β -B. First-principles studies investigating the stability of α -B and defect-free β -B showed a higher stability of α -B at 0 K^{155, 157}. The area circled with black dashed lines indicates the possible location of α - β - γ triple point. The α - β - γ triple points reported by Oganov *et al.*⁶¹ and Parakhonskiy *et al.*⁶³ are 7.7 GPa and 1300 K, 7.6 GPa and 1880 K, respectively. Pressures of the triple point predicted by previous studies also fall into this range, whereas our predicted triple point temperature is lower than other studies.

Ogitsu *et al.*⁵⁹ pointed out that partial occupancy at the B13 site in B_{105} (red spheres in Figure 4.4C) can stabilize β -B. Here, we investigated the thermodynamic stability of β -B with a single B13 vacancy (denoted as β - B_{104}) using the SGP potential. The fundamental properties related to β - B_{104} used to calibrate the SGP potential are shown in Figure B.8. It can be noted that the predicted enthalpies agree well with quantum mechanical calculations (Figure B.8B). Furthermore, good agreements on the phonon PDOS (Figure B.8D) and vibrational entropies (Figure B.7D) were also obtained. The structural extrapolation from B13 defect was quantified using principal component analysis (PCA) projection with code ASAP¹⁷², as shown in Figure B.9. Clearly, β - B_{104} has similar features as defect-free B_{105} in compressed dimensions, and both defective and defect-free β -B significantly differ from α - and γ -B. Afterward, the phase diagram

was updated using the free energies of B₁₀₄, as shown in Figure 4.4B. As expected, β -B₁₀₄ is stabilized by the B13 vacancy and becomes the ground-state phase at 0 bar and low temperature (500~700 K). Given that our training set consists of only defect-free structures, there is an indication that the potential training might be capable of extrapolating predicting thermodynamic properties despite the structures consisting of certain types of defects. At low temperatures (e.g., 500 K), Figure 4.4D reveals that the B13 vacancy in β -B₁₀₄ will not significantly alter $-TS$ and ZPE when compared with β -B₁₀₅. The enthalpies of β -B₁₀₄ guarantee the overall free energy of β -B₁₀₄ is lower than that of β -B₁₀₅. However, as the pressure increases, the thermodynamic advantage of β -B₁₀₄ continues to weaken, and eventually, β -B₁₀₅ becomes more stable at high pressures (i.e., $\Delta G_{104-105} > 0$). At constant pressure ($P = 2$ GPa), Figure 4.4E shows that the enthalpy differences (black squares) initially fluctuate as temperature increases until they drift downward, while the ZPE differences (blue triangles) trend in the opposite direction. We anticipate that the defect-free β -B₁₀₅ becomes thermodynamically more favored over β -B₁₀₄ at $T > 1100$ K. Hence, stabilization due to the B13 vacancy will likely occur in relatively low temperatures and pressures.

The individual free energy components (i.e., H , $-TS$, and ZPE) were examined as functions of temperature and pressure to reveal the most influencing thermodynamic contribution during the phase transitions of boron crystals. In Figure 4.5, the relative differences between two boron phases (denoted as $\Delta X_{\beta-\alpha}$, $\Delta X_{\gamma-\alpha}$, and $\Delta X_{\beta-\gamma}$, where X refers to H , $-TS$, ZPE , and G) are presented. Also, only one thermodynamic variable, i.e., pressure or temperature, is allowed to change while the other is fixed at a specific value (i.e., $T = 900$ K and $P = 2$ GPa). Other T and P choices lead to similar conclusions. With α -B as the reference, $\Delta H_{\beta-\alpha}$ (black squares) increases monotonously within 0~16 GPa at 900 K (Figure 4.5A). In contrast, $\Delta H_{\gamma-\alpha}$ exhibits the opposite trend (Figure 4.5C), suggesting that higher pressures destabilize β -B but stabilize γ -B enthalpically. Both $-TS$

(red dots) and ΔZPE (blue triangles) are associated with lattice vibrations and remain nearly constant regardless of pressure. Since these quantities are numerically negative, the inclusions of $-\Delta TS$ and ΔZPE shift ΔG (green triangles) downward. $\Delta G_{\beta-\alpha}$ becomes negative when $P < 4$ GPa, enabling β -B to be the more stable phase over α -B. Meanwhile, the thermodynamic stability transitions from α -B to γ -B at $P > 8$ GPa.

Figures 4.5B and 4.5D show the variations of $\Delta X_{\beta-\alpha}$ and $\Delta X_{\gamma-\alpha}$ with temperatures at $P = 2$ GPa. As expected, the pressure-dependent ΔH (black squares) is almost invariant. Instead, the downward trends of $-\Delta TS$ (red dots) manifest strongly the growing influence of vibrational entropies, which dominate the behaviors of ΔG (green triangles). As illustrated in Figure 4.5B, $-\Delta TS_{\beta-\alpha}$ displays a sharp decrease at $T > 1000$ K. Given that the variations of ΔZPE (blue triangles) occur to a much lesser extent, the entropy is instrumental in stabilizing β -B at high temperatures. While $\Delta X_{\gamma-\alpha}$ in Figure 4.5D follows the same trends as for β -B, we did not observe a sharp decrease in $-\Delta TS$. Moreover, the magnitudes of $-\Delta TS_{\gamma-\alpha}$ and $\Delta ZPE_{\gamma-\alpha}$ are comparable or smaller than $\Delta H_{\gamma-\alpha}$. Thus, the transition from α -B to γ -B is not sensitive to the temperature.

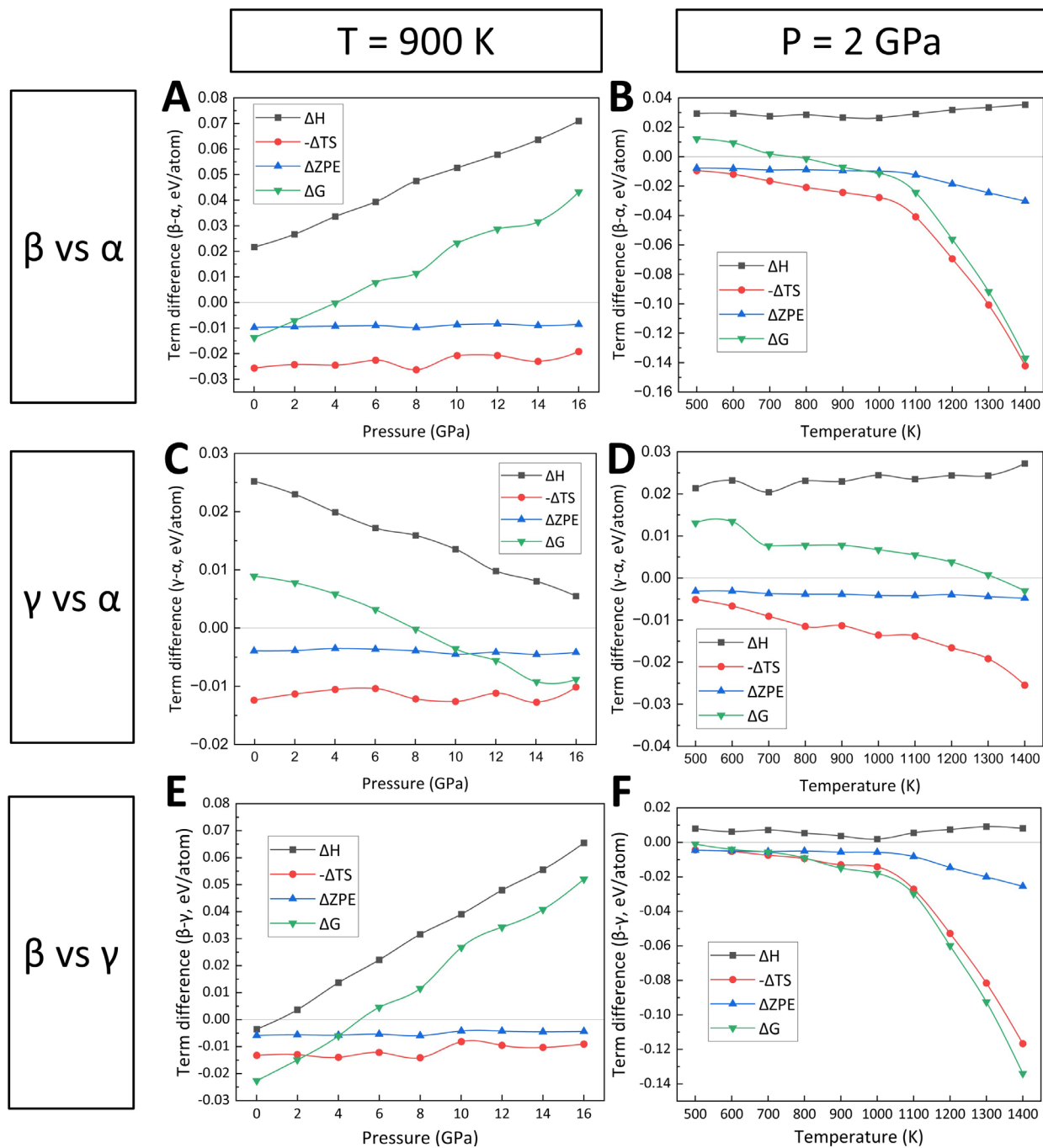


Figure 4.5 Thermodynamic differences between β -B and α -B (A) $T = 900$ K and (B) $P = 2$ GPa; between γ -B and α -B (C) $T = 900$ K and (D) $P = 2$ GPa; between β -B and γ -B (E) $T = 900$ K and (F) $P = 2$ GPa.

Using γ -B as a reference, the same approach was adopted to understand the phase transition from β -B to γ -B. As shown in Figure 4.5E, at constant temperatures, $\Delta H_{\beta-\gamma}$ (black squares) dictates the behavior of $\Delta G_{\beta-\gamma}$ (green triangles), which increase with pressure monotonously, as for β -B versus α -B. Isotropic compressions will more likely disfavor β -B over both α -B and γ -B enthalpically. Such a trend will support the fact that γ -B is the more stable phase at high pressures, as reported in the literature¹⁷³. Next, when the pressure is fixed at 2 GPa, the behavior of $\Delta G_{\beta-\gamma}$ follows closely to $-\Delta TS_{\beta-\gamma}$ and the decrease accelerates as temperatures reach 1000 K, as depicted in Figure 4.6F, favoring the stability of β -B at high temperatures.

Thus far, we generated the phase diagram to establish the relative phase stability of the three boron allotropes. Lastly, we utilized the phase diagrams reported in the literature¹⁷³ to analyze how the inherent discrepancies embedded in the SGP potential and the MD simulations impact the precision of phase boundary locations. We noted that the boundary separating α -B and β -B at 0 bar is 750 K (Figure 4.4A), which is lower than that reported by Parakhonskiy *et al.* (950 K)⁶³. The source of this discrepancy may originate from the larger overestimation of the α -B volume expansion over β -B (0.65% for α -B versus 0.44% for β -B). At a given temperature, the free energy of α -B will be inflated over β -B, making the latter relatively more stable and shifting the phase boundary downward. In addition, both Figures 4.4A-B show that β -B can outcompete γ -B in the higher-pressure regime (i.e., 10~16 GPa). Figure 4.3B rules out that the discrepancies in enthalpic calculations result in favoring β -B. Moreover, Figure B.7 provide evidence that the SGP potential faithfully produce the entropies. Hence, we recognize that, besides the interatomic potential, MD simulations can also introduce uncertainties and in this case overestimate the stabilities of β -B. These issues allude to the ongoing challenges in accurately predicting thermodynamic phase behaviors facing molecular simulations.

4.4 Conclusion

In summary, we presented a protocol including SGP training and thermodynamics stability prediction of elemental boron allotropes. OTF MD simulation, an active learning process, was employed as the engine to generate a training set containing a set of condensed but relevant training information. The trained SGP potential is capable of predicting the structural, thermodynamic, and vibrational properties of elemental boron. Using MD simulations and phonon calculations based on this SGP potential, the P-T phase diagram of α -B, β -B, and γ -B was generated. Defect-free β -B and B13-site-deficient β -B₁₀₄ structures were employed to investigate the impact on the eventual phase diagram. A notable stabilization due to the B13 vacancy was observed, which validates the extrapolation ability of the SGP potential trained from the defect-free training set. By analyzing free energy components, we found that vibrational entropy plays a significant role in energetic ordering, especially during temperature variation. The presented protocol can be extended to other complex crystal systems to study the phase stability and transition mechanisms.

Chapter 5 - Lattice Expansion Investigation of $\text{PrNi}_{1-\delta}\text{Co}_\delta\text{O}_3$ (PNC)

Perovskite induced heating and hydration

Chapter 5 is reproduced with permission from:

Deng H.; Sun Q.; Li M.; Zhao Z.; Bian W.; Liu, B.; Ding D. “Atomistic Simulations of Thermal and Chemical Expansions of $\text{PrNi}_x\text{Co}_{1-x}\text{O}_{3-\delta}$ Accelerated by Machine Learning Potentials”, Small Methods. DOI: 10.1002/smt.202500816.

5.1 Introduction

Protonic ceramic electrochemical cells (PCECs) hold tremendous potential as next-generation electrochemical devices, with several distinct advantages over their competitors, such as enhanced stability¹⁷⁴, reduced operating temperatures^{175, 176}, flexible feedstock options^{177, 178}, high electrochemical efficiency¹⁷⁹, and versatile applications¹⁸⁰⁻¹⁸⁴. Nevertheless, the design and selection of PCEC electrolyte and electrode materials often require a delicate balance to achieve optimal cell performance¹⁸⁵⁻¹⁸⁷. In the past decade, extensive attention has been given to oxygen electrode materials^{175, 188-191}, which must fulfill several critical requirements, including high catalytic activity for electrochemical reactions, durability, suitable hydration ability and charge carrier (i.e., protons, electrons, oxygen-ions) conductivities^{192, 193}. The $\text{PrNi}_x\text{Co}_{1-x}\text{O}_{3-\delta}$ (PNC) perovskite developed by Ding and coworkers^{69, 70, 180} as an oxygen electrode material exhibits remarkable electrochemical performance. In particular, the $\text{PrNi}_{0.7}\text{Co}_{0.3}\text{O}_{3-\delta}$ yields the highest power and current densities of PCEC observed so far⁷⁰. PNC is also shown to mitigate phase segregation and degradation, effectively strengthening its structural stability^{69, 70}.

The mechanical properties of PCEC materials are crucial for large-scale assemblies ¹⁹⁴. Notably, poor oxygen electrode and electrolyte interface stability results in rapid deterioration ¹⁹⁵, ¹⁹⁶. Interfacial rejuvenation ¹⁸⁰ and thermal expansion offset ⁶⁵ are among the established methods for improving the stability of the electrode-electrolyte interface. A primary influencing factor on interface instability is the mismatch between the oxygen electrode and the electrolyte that occurs during lattice expansion. Under *operando* conditions, the oxygen electrodes are exposed to elevated temperatures (400°C to 600°C) and steam concentrations (3% to 50%), which could lead to significant thermal expansion (TE) and hydration-induced chemical expansion (CE). The hydration of electrodes, depicted as a proton uptake pathway, is expressed by Eqn (5.1):



where two protonic defects ($\text{OH}_{\text{O}}^{\bullet}$) are generated in the lattice upon the hydration by one molecule of water, filling one oxygen vacancy ($\text{V}_{\text{O}}^{\bullet\bullet}$) while protonating one lattice oxygen ($\text{O}_{\text{O}}^{\times}$). Typically, electrode materials capable of high hydration are advantageous because they enable more efficient proton uptake, particularly in electrolysis mode. However, when oxygen electrodes are operated in high steam concentration scenarios, unwanted expansion mismatch occurs. For PNC-based electrodes, it has been reported that a higher Ni occupancy leads to lower TE ⁷⁰, but the effect of Ni/Co occupancy on CE remains unclear. Moreover, understanding the complex interactions between TE and CE is crucial to improving oxygen electrode assemblies for PCEC development. Factors influencing TE (e.g., temperature, chemical composition) can also affect CE and vice versa ^{197, 198}. Specifically, CE due to the formation and annihilation of defects is affected by temperatures, which, in turn, alter the atomic environments surrounding the defects ⁷⁶. Moreover, $\text{V}_{\text{O}}^{\bullet\bullet}$ filling can change the atomic configuration and alter the internal stress in the lattice, influencing the TE behavior ¹⁹⁹.

Because TE and CE both occur under *operando* conditions, decoupling their behaviors helps to understand how the PNC lattices, consisting of OH_O and $\text{V}_\text{O}^\bullet$, respond to temperature increases. Experimental measurements of TE are mature ²⁰⁰, whereas CE measurements are demanding ²⁰¹. According to Equation 5.1, the CE due to hydration involves the filling of $\text{V}_\text{O}^\bullet$ and the formation of OH_O . Hence, contributions from different defects cannot be decoupled by simply interpreting measured volume expansions. Computationally, previous efforts to decouple TE and CE in perovskites have relied on first-principles phonon calculations. For instance, Løken *et al.* ¹⁹⁹ computed the phonon dispersions of BaCeO_3 to decouple TE and CE and revealed the effect of lattice vacancies on the linear thermal expansion coefficients. A similar methodology was used by Bjørheim and coworkers ²⁰² to predict the hydration entropies of undoped and acceptor-doped BaZrO_3 . The relationship between CE and the hydration thermodynamics of four proton-conducting perovskites, i.e., BaZrO_3 , BaSnO_3 , BaCeO_3 , and SrZrO_3 , was investigated ²⁰³. Nevertheless, the high computational cost significantly burdens first-principles phonon calculations on perovskites, such as PNC, with much more complex elemental compositions and defect configurations.

Machine learning potentials (MLPs) are being recognized as a powerful surrogate for complex potential energy surfaces, enabling the understanding of material properties and providing a reliable and scalable avenue for studying multi-element, multi-phase systems ^{41, 204, 205}. For material property simulations, many MLP architectures have been widely developed and applied ^{25-38, 41, 42, 53, 206}. In recent studies, the MLP-based phonon calculation has been shown to be a highly efficient approach for thermal property predictions ²⁰⁷⁻²⁰⁹. In this work, we employed MLP to calculate the phonon properties of PNC with selected structural defects and compositional

variations and then to decouple the TE and CE to alleviate the computational burden while retaining the level of accuracy of first-principles methods.

Herein, we report a method that integrates on-the-fly molecular dynamics (OTF-MD) simulation³⁹⁻⁴² with the Allegro^{43, 44, 210, 211}. OTF-MD is an active learning process based on the sparse Gaussian process (SGP), which was employed for the generation of a high-quality training set across diverse composition and defect configurations of PNC. Allegro employs an E(3)-equivariant convolution for the interactions of geometric tensors and can be trained over more information-rich and faithful representations of atomic environments. After validating the trained MLP, we developed a phonon-based workflow to disentangle individual contributions to TE and CE in PNC. The effects of V_{O} and OH_{O} The TE of the lattice was investigated, and the combined TE and CE, along with their implications for the performance of PNC electrodes, were discussed. The MLP-based phonon calculations offer a new computational paradigm for lattice expansion in complex oxide electrodes, thereby facilitating the evaluation of interfacial stability in PCEC systems.

5.2 Methods

We built an inclusive dataset for thermal/chemical expansion simulations to encompass the configurational space. Figure 5.1A illustrates the workflow of MLP training. The workflow can be described as (i) MLP training, initiated by training set generation and offline training, conducted with FLARE^{39, 40} and Allegro^{43, 44}, (ii) MLP accuracy validation, and (iii) phonon calculations for lattice thermal and chemical expansions.

5.2.1 Density Functional Theory (DFT)

DFT serves as the first-principles calculation engine for training and benchmark data generation. All spin-polarized DFT calculations were performed using the Vienna *Ab Initio* Simulation Package (VASP)^{8,9}. The projector augmented wave (PAW) pseudopotentials⁸¹, and the GGA-PBE functional¹³⁴ were used to account for the core electrons and the electron exchange-correlation effects. To verify the rationality of choosing GGA-PBE, we compared the $V_{\text{O}}^{\bullet\bullet}$ formation energies ($\Delta E_{V_{\text{O}}^{\bullet\bullet}}$) predicted by GGA-PBE and $r^2\text{SCAN}+\text{rVV10}$ ²¹², which is a up-to-date methodology combining meta-GGA and non-local van der Waals (vdW) correction. As shown in Table C.9, GGA-PBE gives the same energetic ordering of four $V_{\text{O}}^{\bullet\bullet}$ sites as $r^2\text{SCAN}+\text{rVV10}$, and the predicted $\Delta E_{V_{\text{O}}^{\bullet\bullet}}$ of GGA-PBE is $\sim 10\%$ lower. Considering the lower expense and adequate accuracy, GGA-PBE was chosen as the functional for all DFT calculations. A $5 \times 5 \times 3$ k-point mesh was automatically generated using the Monkhorst-Pack scheme²¹³ to sample the Brillouin zone. The Hubbard U method²¹⁴ was employed to correct the self-interaction effects of the d -electrons in Ni and Co with $(U-J)_{\text{Ni}} = 6.4$ eV and $(U-J)_{\text{Co}} = 3.3$ eV⁶⁹. The plane-wave basis set was expanded for geometry optimizations with a cutoff energy of 520 eV. A slightly smaller cutoff energy of 450 eV was used for single point energy calculations during OTF-MD simulations.

5.2.2 Dataset Generation

As shown in Figure 5.1A, three pristine PNC unit cells with different Ni/Co occupancies, i.e., $\text{PrNi}_{0.25}\text{Co}_{0.75}\text{O}_3$ (PNC37), $\text{PrNi}_{0.5}\text{Co}_{0.5}\text{O}_3$ (PNC55), and $\text{PrNi}_{0.75}\text{Co}_{0.25}\text{O}_3$ (PNC73), were considered. Also, unit cells containing a single $V_{\text{O}}^{\bullet\bullet}$ or $\text{OH}_{\text{O}}^{\bullet}$ defect at all distinct sites were considered and optimized. Supercells consisting of more than one $V_{\text{O}}^{\bullet\bullet}$ were not considered for this work. The lattice parameters of optimized unit cells were summarized in Table C.3. The defective

structures for phonon calculations are illustrated in Figure C.1. Next, all DFT-optimized unit cells, with and without defects, were passed onto the FLARE package as initial structures for OTF-MD. All OTF-MD simulations were run with an NVT ensemble under 1300K, a temperature high enough to trigger lattice distortion and defect migration during OTF-MD. Once the uncertainty of a structure in the trajectory exceeds the predetermined threshold (see Table C.10 in Supporting Information), standard DFT will be called to calculate the single point energy and interatomic forces, which will be stored as the dataset. The detailed information on OTF-MD setting is summarized in Table C.10 of Supporting Information.

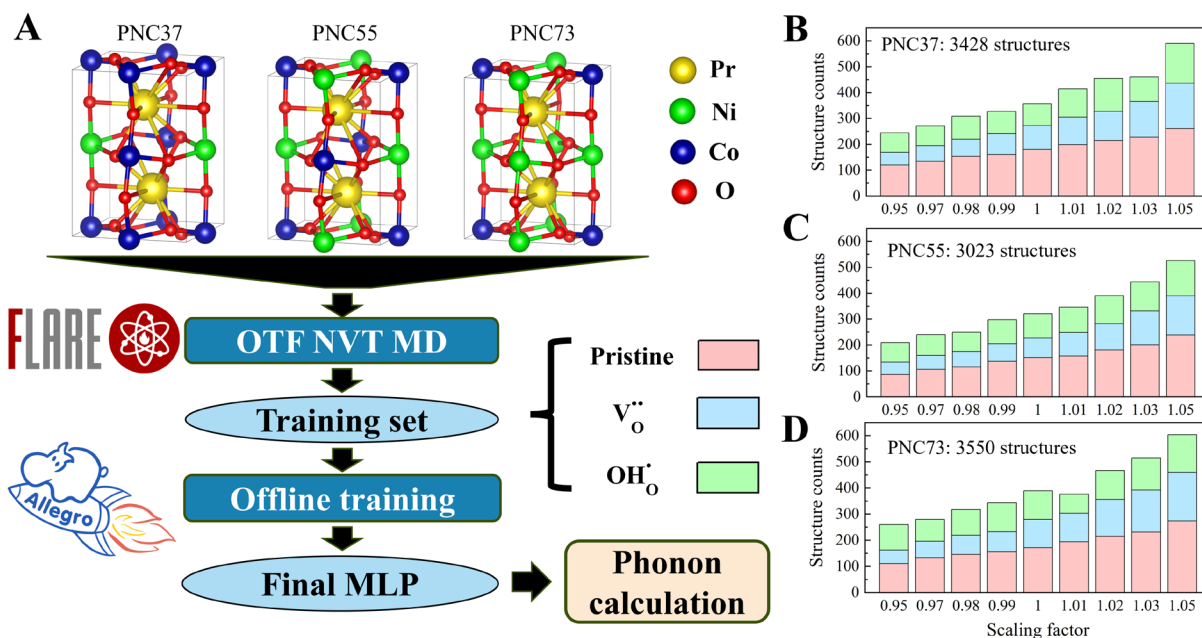


Figure 5.1 The workflow of PNC MLP training and detailed dataset distribution of (B) PNC37, (C) PNC55, and (D) PNC73.

Structure candidates collected for PNC37 (3428 structures), PNC55 (3023 structures), and PNC73 (3550 structures) from OTF-MD were compiled into a dataset pool. Figures 5.1B, 5.1C, and 5.1D show the compositions of the training set. For each composition, nine scaling factors

(0.95, 0.97, 0.98, 0.99, 1.0, 1.01, 1.02, 1.03, 1.05) were employed to scale the PNC unit cells. We noted that, for each scaling factor, the number of structures is evenly distributed between pristine and defective structures (i.e., V_{O} and OH_{O} combined). As the scaling factor increases, the number of structures generated by OTF-MD increases, suggesting more configurations were sampled for the expanded cells, where larger atom displacements from their equilibrium positions are more likely to occur.

5.2.3 Machine Learning Potential (MLP) Training

Once the dataset was generated, it was transferred for offline training implemented in Allegro. The final dataset consists of 10001 structures and was randomly divided into the training set (8500 structures), validation set (750 structures), and testing set (751 structures). Table C.11 summarizes the training setup for the final MLP production. Figure C.10 presents the parity plots for the predicted energies and forces based on the testing set structures. The corresponding MAEs are 7.16 meV/atom and 102.08 meV/Å. To explicitly demonstrate the high efficiency of MLP compared with DFT, we summarize the calculation wall times for single-point energy calculations and geometry optimizations using DFT and MLP in Table C.12. We can anticipate that the calculation speed of MLP is roughly 10,000 times faster than DFT.

The training set contains PNC unit cells with specific Ni/Co ordering (Figure 5.1A). To validate the MLP can extrapolate to other orderings, ten 2x2x1 PNC73 supercells were generated with random Ni/Co ordering and optimized by DFT and MLP, respectively. As shown in Table C.8, optimizations by DFT and MLP give very similar results including energy errors of < 0.5 meV/atom and lattice constant errors of < 1%. The result also proves the ordering change doesn't affect the energy of the system significantly (maximum energy difference is 0.48 meV/atom).

Therefore, to maximize the accuracy of MLP, the same ordering in the training set was employed in subsequent analysis.

5.2.4 MLP-based Phonon Calculation and Expansion Prediction

The phonon calculations were carried out using the Phonopy^{107, 169, 215} and LAMMPS¹³⁷ programs. To determine the equilibrium volumes at different temperatures, the unit cells were scaled with 20 scaling factors with a fine interval of 0.005 in a range between 0.95 and 1.05. For each unit cell, a 4×4×4 supercell with finite displacements of ± 0.01 Å was generated. The forces were calculated using the MLP to yield the dynamical matrix for the estimations of thermodynamic properties. Finally, based on the quasiharmonic approximations (QHA)²¹⁶, the equilibrium volumes under 300-1000K were determined using the G - V relationships, as shown in Figures C.11-C.13.

5.3 Results and Discussion

5.3.1 MLP Benchmark

The equation of states (EOS) for PNC unit cells, with varying Ni/Co occupancy and V_{O} or OH_{O} defects, were examined to reveal the relationships of ground state energies under varying compressive or expansive strains. Energy minimizations were performed using DFT and MLP to benchmark the trained MLP. The EOS based on MLP and DFT calculations were compared in Figure C.3, represented by the red and black curves, respectively. As shown, MLP simulations agree well with DFT calculations on the total energies of unit cells with varied Ni/Co occupancy and defects. The equilibrium volumes corresponding to each energy minima are also in good agreement. Phonon dispersion and phonon density of states (PDOS) of PNC55 were also

calculated using DFT and MLP, as shown in Figure C.5 in the Supporting Information. Again, good agreements were achieved between DFT and MLP, especially in the low-frequency region.

5.3.2 Thermal Expansion

The thermal expansions of pristine PNC unit cells, and those with $V_O^{\bullet\bullet}$ and $OH_O^{\bullet}$ defects at three Ni/Co occupancies, are summarized in Figure 5.2. The lattice dimension (l) of each cell is determined by $l = V^{1/3}$ based on the approach reported by Løken and coworkers¹⁹⁹. Under all scenarios, the values and rates of thermal expansion both follow the trend of PNC37 > PNC55 > PNC73, indicating that higher Co occupancy contributes to a higher thermal expansion of PNC, agreeing with reported experimental results⁷⁰.

The thermal expansions of pristine (Figure 5.2A) and $OH_O^{\bullet}$ -defective (Figure 5.2C) PNC exhibit approximate linear behaviors. The introduction of $OH_O^{\bullet}$ will also increase the rate of thermal expansions for PNC of all Ni/Co occupancies, indicating that $OH_O^{\bullet}$ only has a limited effect on thermal expansion. Nevertheless, Figure 5.2B suggests that the $V_O^{\bullet\bullet}$ defects significantly impact thermal expansions at all Ni/Co occupancies, especially for PNC37 and PNC55. Figure 5.2B also shows that PNC's lattice parameter varies nonlinearly with temperature, and the slopes accelerate at higher temperatures. The fact that the $V_O^{\bullet\bullet}$ defects have more influence over the thermal expansion of PNC than $OH_O^{\bullet}$ has been reported by Løken and coworkers on the orthorhombic $BaCeO_3$ ^{199, 203}, where the perovskite's lattice shrinks in the presence of $V_O^{\bullet\bullet}$ and $OH_O^{\bullet}$.

The volumes of all DFT optimized PNC cells increase upon the formations of $V_O^{\bullet\bullet}$ and $OH_O^{\bullet}$, as shown in Figures C.9B and C.9C. The behaviors associated with volume changes during defect formation and TE are consistent with the findings by Løken and coworkers¹⁹⁹. The $V_O^{\bullet\bullet}$ - and

OH_O -induced cell volume expansions suggest that the bonds between metal cations and lattice oxygen are weakened. To validate this hypothesis, Bader charge analyses were performed on all distinct defective sites (Tables C.2-C.7). Charge accumulations at the Pr, Ni, and Co atomic species suggest that these cations experience varying degrees of reduction when a $\text{V}_\text{O}^{\bullet\bullet}$ is generated. Additionally, we observed that the energies of pristine PNC lattices undergoing volume compression and expansion always vary faster than their defective counterparts under the same strains, indicating that higher energy is required to overcome the stronger metal-oxygen bonds.

The total pressure (P_{total}) and their normal components (P_{xx} , P_{yy} , and P_{zz}) of PNC bulk with and without defects under expansion are summarized in Figure C.7. It was found that the magnitude of the stress grows when the Ni occupancy increases. This observation agrees with the observed TE trend, where higher Ni occupancy leads to lower TE, as shown in Figure 5.2. Furthermore, the generations of $\text{V}_\text{O}^{\bullet\bullet}$ and OH_O defects lower the internal stress and agree with higher TE values observed for defective structures. There are also discernible differences among P_{xx} , P_{yy} , and P_{zz} . All normal pressure components follow the trend of $P_{xx} > P_{zz} > P_{yy}$, which suggests that lattice expansion is not perfectly isotropic.

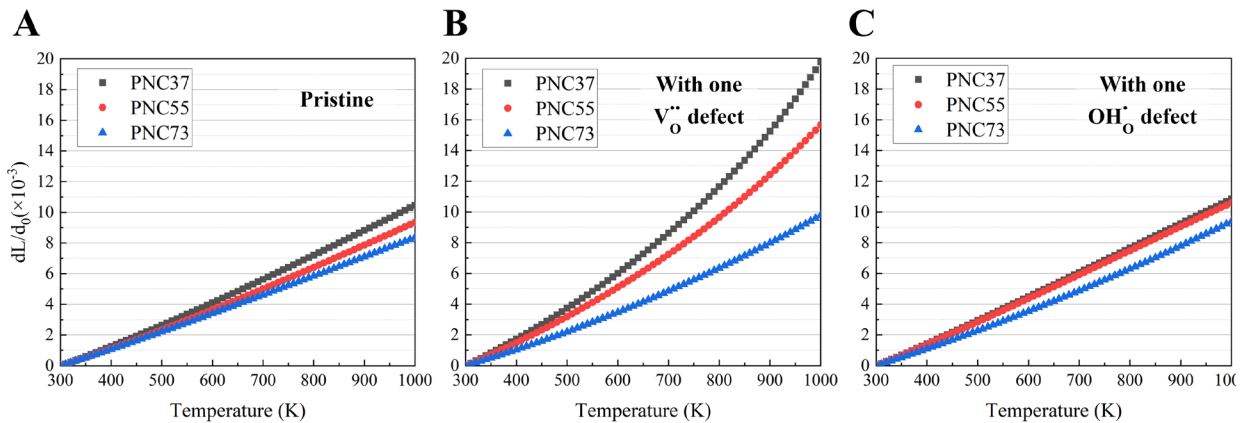


Figure 5.2 Thermal expansion (TE) predicted by MLP-based phonon calculations on (A) pristine, (B) OH_O -defective, and (C) $\text{V}_\text{O}^{\bullet\bullet}$ -defective PNC.

While the simulated thermal expansions for the $V_{\text{O}}^{\bullet\bullet}$ -defective PNC cells agree well with the measurements (Figure C.8), the values for the pristine and the $\text{OH}_{\text{O}}^{\bullet}$ -defective PNC are systematically lower than those reported by Tang *et al.*⁷⁰. We attributed these phenomena to the fact that synthesized PNC perovskites already consist of intrinsic $V_{\text{O}}^{\bullet\bullet}$ in the lattice, as suggested by the oxygen temperature programmed desorption (TPD)⁷⁰. Noted that the thermal expansions of $V_{\text{O}}^{\bullet\bullet}$ -defective PNC agree remarkably well with experiments below 650 K, as shown in Figure C.8. Such deviations become especially notable at higher temperatures. Moreover, the disagreements are more evident as the Co occupancy increases and can be understood as the discrepancies in $V_{\text{O}}^{\bullet\bullet}$ concentrations between the model and synthesized PNC samples.

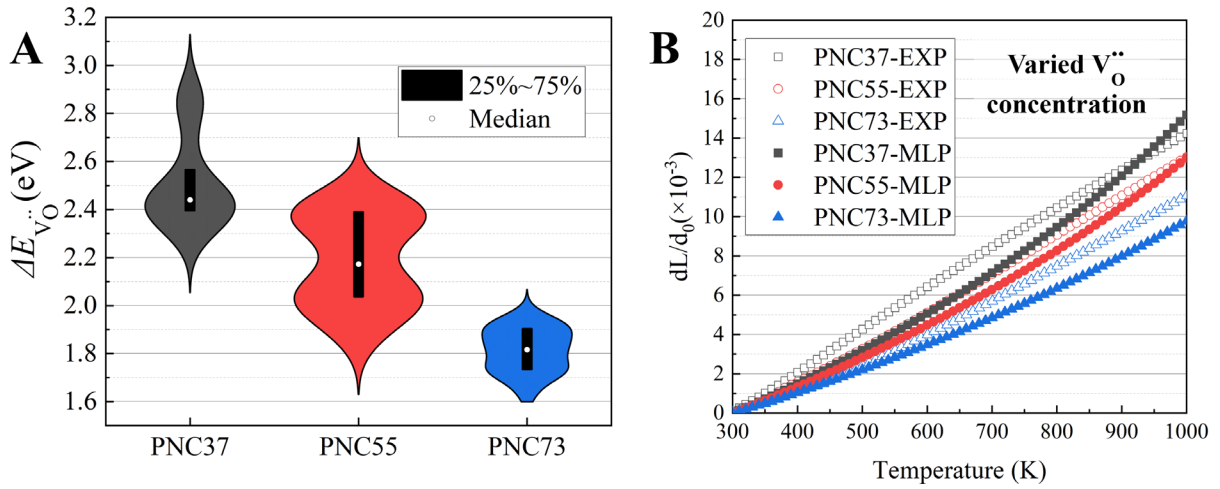


Figure 5.3 (A) The oxygen vacancy formation energies ($\Delta E_{V_{\text{O}}^{\bullet\bullet}}$) of PNC37 (grey), PNC55 (red), PNC73 (blue). (B) Thermal expansions from experiments (hollow symbols) and MLP phonon calculations (solid symbols) for $C_{V_{\text{O}}^{\bullet\bullet}, \text{PNC37}} = 4.17\%$, $C_{V_{\text{O}}^{\bullet\bullet}, \text{PNC55}} = 4.76\%$, and $C_{V_{\text{O}}^{\bullet\bullet}, \text{PNC73}} = 8.33\%$. The experimental results were adapted from Ref.⁷⁰.

Informed by oxygen TPD ⁷⁰, oxygen vacancies occur at three distinct oxygen sites in PNC73 and PNC55 but only two sites in PNC37, indicating that the $V_{\text{O}}^{\bullet\bullet}$ concentration ($C_{V_{\text{O}}^{\bullet\bullet}}$) increases with a higher Ni occupancy. The $V_{\text{O}}^{\bullet\bullet}$ formation energies ($\Delta E_{V_{\text{O}}^{\bullet\bullet}}$) in PNC37, PNC55, and PNC73 were calculated using DFT on all possible lattice sites according to Eqn. (5.2):

$$\Delta E_{V_{\text{O}}^{\bullet\bullet}} = E_{V_{\text{O}}^{\bullet\bullet}} + \frac{1}{2} E_{\text{O}_2} - E_{\text{bulk}} \quad (5.2)$$

where $E_{V_{\text{O}}^{\bullet\bullet}}$, E_{O_2} , and E_{bulk} are the total energies of the unit cell with an oxygen vacancy, the oxygen molecule, and the pristine unit cell, respectively. E_{O_2} was obtained by performing geometry optimization on O_2 molecule using DFT in a $13 \text{ \AA} \times 14 \text{ \AA} \times 15 \text{ \AA}$ periodic box. As shown in Figure 5.3A, $\Delta E_{V_{\text{O}}^{\bullet\bullet}}$ follows the trend of $\text{PNC37} > \text{PNC55} > \text{PNC73}$ as the formation of $V_{\text{O}}^{\bullet\bullet}$ becomes increasingly difficult with the Co occupancy.

For a PNC supercell with one $V_{\text{O}}^{\bullet\bullet}$, the $C_{V_{\text{O}}^{\bullet\bullet}}$ is approximately 8.33% based on Eqn. (5.3). Thus, the overestimation of thermal expansion can be attributed to the higher $C_{V_{\text{O}}^{\bullet\bullet}}$ than the PNC samples used in experimental measurements. The overestimation will be more severe as Co occupation increases, consistent with the deviations observed in Figure 5.2B. Hence, we adjusted the volume of the defective cell (V_{def}) by averaging the volumes of the pristine cell (V_{bulk}) and that of the $V_{\text{O}}^{\bullet\bullet}$ -defective cell ($V_{V_{\text{O}}^{\bullet\bullet}}$) according to Eqn. (5.4):

$$C_{V_{\text{O}}^{\bullet\bullet}} = \frac{n_{V_{\text{O}}^{\bullet\bullet}}}{12(n_{\text{bulk}} + n_{V_{\text{O}}^{\bullet\bullet}})} \quad (5.3)$$

$$V_{\text{def}} = \frac{n_{\text{bulk}} \cdot V_{\text{bulk}} + n_{V_{\text{O}}^{\bullet\bullet}} \cdot V_{V_{\text{O}}^{\bullet\bullet}}}{n_{\text{bulk}} + n_{V_{\text{O}}^{\bullet\bullet}}} \quad (5.4)$$

where n_{bulk} and $n_{V_{\text{O}}^{\bullet\bullet}}$ are the numbers of pristine and $V_{\text{O}}^{\bullet\bullet}$ -defective cells, respectively. The thermal expansion was updated using the adjusted volume of the defective cells, as shown in Figure 5.3B. The MLP-based thermal expansions show improved agreement with experimental results. The

volume of $V_{\text{O}}^{\bullet\bullet}$ -defective PNC73 was not adjusted and remained the same (Figure 5.2B) because the MLP was already trained using the $C_{V_{\text{O}}^{\bullet\bullet}}$ value comparable to the experimental samples.

5.3.3 Chemical Expansion

The chemical expansion (ζ) due to the defect i ($V_{\text{O}}^{\bullet\bullet}$ or $\text{OH}_{\text{O}}^{\bullet}$) were calculated using Eqn. (5.5):

$$\zeta_i = \frac{1}{\delta_i} \left(\frac{l_i(T) - l_0(T)}{l_0(T)} \right) \quad (5.5)$$

where δ_i is the number of defects per formula unit (i.e., 4, due to the single defect formation in the unit cell), $l_i(T)$ and $l_0(T)$ are the dimensions of respective defective and pristine unit cells¹⁹⁹. The chemical expansions related to hydration (ζ_{hyd}) were estimated with Eqn. (5.6), accounting for the simultaneous filling of one $V_{\text{O}}^{\bullet\bullet}$ and the formation of two $\text{OH}_{\text{O}}^{\bullet}$:

$$\zeta_{\text{hyd}} = 2\zeta_{\text{OH}_{\text{O}}^{\bullet}} - \zeta_{V_{\text{O}}^{\bullet\bullet}} \quad (5.6)$$

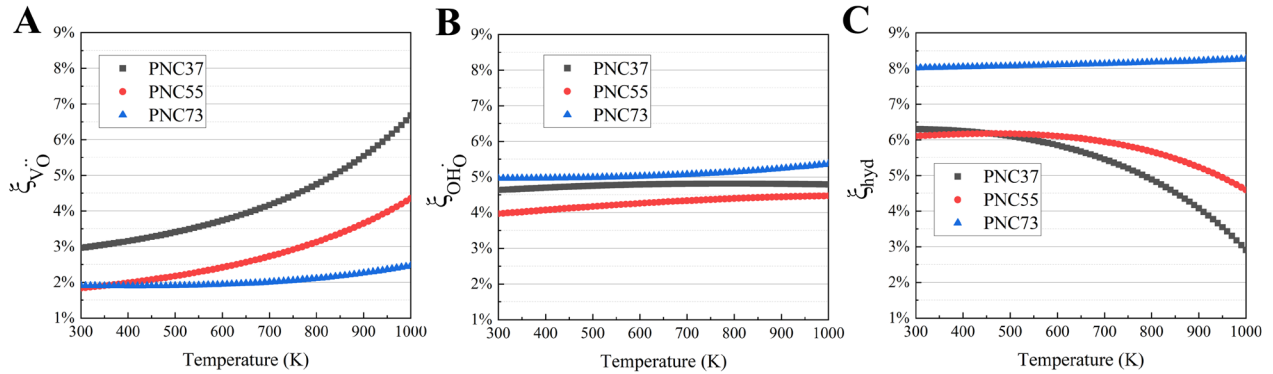


Figure 5.4 Chemical expansions (ζ) of PNC37 (black), PNC55 (red), and PNC73 (blue) from MLP phonon calculations for (A) $V_{\text{O}}^{\bullet\bullet}$ formation ($\zeta_{V_{\text{O}}^{\bullet\bullet}}$), (B) $\text{OH}_{\text{O}}^{\bullet}$ formation ($\zeta_{\text{OH}_{\text{O}}^{\bullet}}$), and (C) hydration (ζ_{hyd}).

Figure 5.4 summarizes the chemical expansions of PNC37, PNC55, and PNC73 as a function of temperature. The initial ξ at 300 K does not correlate with the Ni/Co occupancy because ξ is strongly dependent on the site of defect formation. Different oxygen sites in the same PNC lattice are distinct in chemical expansions. The calculated distributions are shown in Figures C.9B-9D. Here, all V_{O} and OH_{O} were created at the same oxygen site in different PNC lattices, and the discussion primarily focuses on the chemical expansion sensitivity to the temperature.

As shown in Figure 5.4A, the $\xi_{V_{\text{O}}}$ of PNC increases with temperature. However, $\xi_{V_{\text{O}}}$ responds differently to the temperature variation as the Ni/Co occupancy varies. As the Co occupancy increases, $\xi_{V_{\text{O}}}$ changes more rapidly as temperature increases. In PNC37, $\xi_{V_{\text{O}}}$ changes from 2.97% to 6.68% between 300 K and 1000 K, whereas PNC73 varies more slowly, from 1.91% to 2.47%. As temperature increases, the difference of $\xi_{V_{\text{O}}}$ among PNCs becomes evident, with $\xi_{V_{\text{O}}}$ following the trend of PNC37 > PNC55 > PNC73. In contrast, $\xi_{\text{OH}_{\text{O}}}$ (Figure 5.4B) is relatively insensitive to the temperature. The $\xi_{\text{OH}_{\text{O}}}$ of PNC37 is almost invariant with temperature, and $\xi_{\text{OH}_{\text{O}}}$ of PNC55 and PNC73 only increase slightly. Moreover, the ordering of $\xi_{\text{OH}_{\text{O}}}$ does not change within the temperature range. The behaviors ξ_{hyd} at temperatures are displayed in Figure 5.4C, based on Eqn. (5.6). Expectedly, ξ_{hyd} varies inversely concerning $\xi_{V_{\text{O}}}$. The temperature-dependence of ξ_{hyd} varies similarly to that of $\xi_{V_{\text{O}}}$, increasing with higher Co occupancy in the lattice. ξ_{hyd} of PNC37 and PNC55 decrease with temperature. Also, ξ_{hyd} of PNC73 exhibits invariance to temperature due to the relationship with $\xi_{V_{\text{O}}}$ and $\xi_{\text{OH}_{\text{O}}}$.

Based on Figure 5.4, we made two observations. (i) The ξ_{hyd} of PNC with high Ni occupancy will be less sensitive to temperature because $\xi_{V_{\text{O}}}$ and $\xi_{\text{OH}_{\text{O}}}$ will remain mostly constant. (ii) The temperature-sensitivity of ξ_{hyd} for high Co occupancy PNC is highly dependent on the

variation of $\zeta_{V_{\text{O}}^{\cdot\cdot}}$. Quantitatively, the ζ_{hyd} of PNC73 are much higher than PNC37 and PNC55, especially at high temperatures. Figure C.7D validates the statement based on DFT calculations under 0 K. Hence, we concluded that as the temperature rises, ζ_{hyd} are larger and vary slowly with the temperature for PNC with a higher Ni occupancy. On the other hand, ζ_{hyd} will be low and decrease with temperatures for PNC with a high Co occupancy.

5.3.4 Simultaneous Effect of Ni/Co Occupancy on Expansion and Performance

The impact on the combined thermal-chemical expansions and performance are discussed in this section. Figure 5.5 displays contour plots illustrating the variations of expansions with temperature and hydration. The volume of hydrated lattice (V_{hyd}) was obtained using Eqn. (5.7):

$$V_{\text{hyd}}(T) = \frac{(1-x) \cdot V_{\text{bulk}}(T) + (1-x) \cdot V_{V_{\text{O}}^{\cdot\cdot}}(T) + 2x \cdot V_{\text{OH}_2}(T)}{2} \quad (5.7)$$

where V_{OH_2} is the volume of lattice with a single OH_2 , and x is the fraction of hydrated $V_{\text{O}}^{\cdot\cdot}$. Eqn. (5.7) defines the initial $C_{V_{\text{O}}^{\cdot\cdot}}$ being 4.17% (i.e., 1/24) for all PNCs, where $V_{\text{hyd}} = (V_{\text{bulk}} + V_{V_{\text{O}}^{\cdot\cdot}})/2$ when $x = 0$.

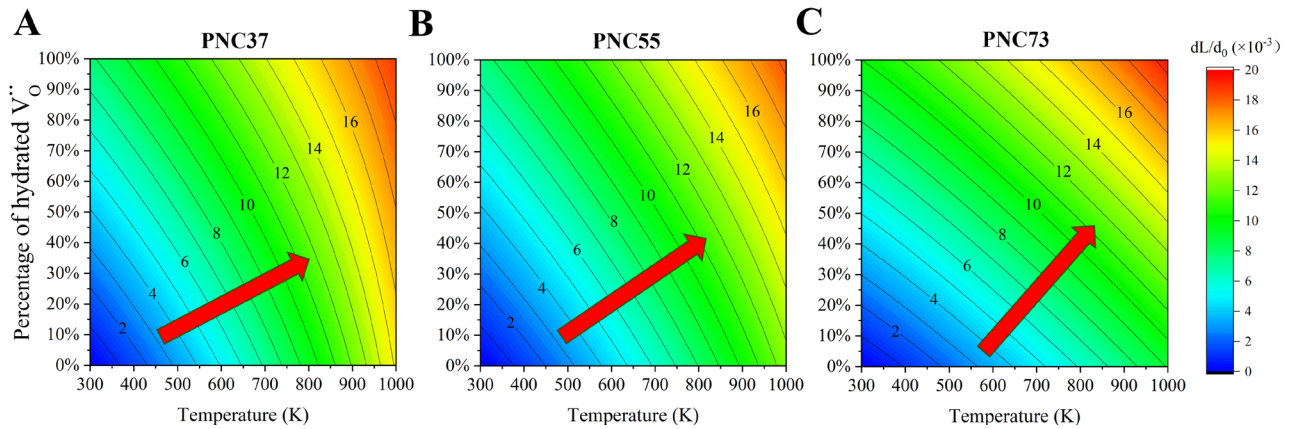


Figure 5.5 Contour plots for lattice expansions of (A) PNC37, (B) PNC55, and (C) PNC73, varying with temperature (x-axis, 330-1000 K) and hydration (y-axis, 0-100%).

Thus far, we have shown that Ni/Co occupancy variations lead to different expansion behaviors. By increasing the Ni occupancy, the expansions (thermal and chemical) of PNC are more sensitive to the extent of $V_{\text{O}}^{\bullet}$ and hydration, but is less sensitive to the temperature increase. Nevertheless, all three PNC compositions ultimately reach similar expansion values at the highest temperature (1000 K) and hydration (100%), that is, $dL/d_0, \text{PNC37} = 18.77 \times 10^{-3}$, $dL/d_0, \text{PNC55} = 18.28 \times 10^{-3}$, $dL/d_0, \text{PNC73} = 19.47 \times 10^{-3}$. However, PNC55 and PNC73 are expected to exhibit higher dL/d_0 , because, as discussed previously, the actual oxygen vacancy concentrations in experimentally prepared perovskites should be higher than $C_{V_{\text{O}}^{\bullet}}$ calculated with Eqn. (5.7). The two-dimensional polynomials were employed to fit the data shown in Figure 5.5. The fitted parameters were summarized in Table C.8 as empirical equations for further predictions.

Experimentally, the PNC perovskites have higher thermal expansion coefficients than the commonly used protonic electrolyte BCZYYb4411⁷⁰. This work suggests that a milder expansion is preferable to ensure mechanical stability at the electrode-electrolyte interface. As such, PNC with the Ni/Co=7:3 occupancy should be considered ideal, and in practice, PNC73 indeed produced the most optimal electrochemical performance when used as the oxygen electrode⁷⁰. From the perspective of lattice expansion, a higher Ni occupancy is beneficial to mitigate the thermal expansion but detrimental to reducing the material's chemical expansion. However, the overall expansion will remain high because of the higher $C_{V_{\text{O}}^{\bullet}}$ that facilitates the hydration of the electrode.

One can also learn that the Ni/Co occupancy in a PNC electrode may bring two opposing effects on electrochemical performance and interfacial stability. The high hydration-induced chemical expansion is the main contributor to weakening interfacial stability. Hydration-induced

chemical expansion is anticipated to exhibit the same trend as electrochemical performance when more Ni is added to PNC. Because a higher $C_{V_O^{\bullet\bullet}}$ causes more sites to be hydrated, the lattice volume and proton concentration will increase simultaneously.

Therefore, to enhance the stability of interfaces between the Ni-abundant PNC electrode and electrolyte, one can mainly focus on mitigating the effect of hydration-induced chemical expansion. From the operation perspective, lowering the steam concentration on the oxygen electrode is one effective strategy. Alternatively, electrolytes with higher chemical expansions can also be deployed to mitigate the expansion mismatch.

5.4 Conclusion

In summary, this paper reports a computational investigation on lattice expansion of PNC affected by temperature, hydration, and Ni/Co occupancy with MLP-driven phonon calculations. Initially, thermal and chemical expansions were investigated separately. It is revealed that the Ni/Co occupancy contributes to thermal and chemical expansions differently, where thermal expansions are related to Co occupancy. In contrast, chemical expansions are more closely associated with the Ni occupancy. Both $V_O^{\bullet\bullet}$ and $OH_O^{\bullet}$ lead to higher thermal expansion than the pristine PNC. The temperature increase will produce a negative impact on hydration-induced chemical expansions. In practice, efforts that boost the hydration of the PNC electrodes at higher Ni occupancies are likely to complicate the electrode-electrolyte interfacial stability due to severe chemical expansions. Strategies that mitigate chemical expansions of the Ni-abundant PNC must also be considered.

The reported methodology can be further applied to predict future thermal and chemical expansions of new proton-conducting perovskites. Moreover, this work also demonstrated a

protocol to train the MLP efficiently and accurately. Hence, this work will accelerate investigations of new oxygen electrodes and electrolyte designs and development. For example, calculations of intrinsic properties like proton and oxygen vacancy diffusivity can also be achieved by performing MLP-based MD simulations, which has already been applied to investigations of anomalous diffusion of proton ²¹⁷, and hydration-induced mechanical degradation of Y-doped BaZrO₃ (BZY) electrolyte ²¹⁸. The extracted properties can be utilized for large-scale simulations to predict cell performance metrics such as Faradic efficiency.

Chapter 6 - Revealing the Lattice expansion and Electrode

Incompatibility in Y-doped BaZrO₃ Electrolytes

6.1 Introduction

Protonic ceramic electrochemical cells (PCECs) have demonstrated outstanding potential for hydrogen production via intermediate-temperature (350–600°C) and high-temperature (>600°C) water electrolysis^{67, 185}, which boosts the energy efficiency (lower heating value > 80%) and exhibits the superiority comparing to low-temperature (<200 °C) techniques (lower heating value < 70%)²¹⁹. A typical PCEC architecture is a sandwich structure with a dense electrolyte sandwiched between porous oxygen and fuel electrodes. Under *operando* conditions, the oxygen electrode is exposed to steam and decomposes water into oxygen molecules and protons. Then, the generated protons transfer through a proton-conducting electrolyte and form hydrogen gas at the hydrogen electrode. Therefore, a high proton conductivity of the electrolyte and activity of the oxygen evolution reaction (OER) on the oxygen electrode are essential to PCEC water electrolysis.

A primary challenge that hinders the scaling-up of the PCEC is the poor oxygen electrode and electrolyte interface stability, which often leads to rapid deterioration^{195, 196}. Strategies including interfacial rejuvenation¹⁸⁰, thermal expansion offset⁶⁵, and soft lithography²²⁰ have been established for improving the stability of the electrode-electrolyte interface. The large variations in lattice thermal and chemical expansions between the oxygen electrode and the electrolyte should not be overlooked for long-term operations. The lattice expansion of PCEC is caused by two aspects²²¹, which are common phenomena that will impact the electrode and electrolyte's thermal behaviors for high-temperature electrolysis. The first is the thermal expansion (TE) caused by lattice vibration during the heating up of PCEC. Second, the lattice will also

experience the hydration-induced chemical expansion (CE), where the water molecules are incorporated into the lattice, filling oxygen vacancies (V_O) and generating protonic defects (OH).

The ABO_3 type proton-conducting perovskites have been extensively investigated^{69, 175, 192, 222, 223}. The Y-doped $BaZrO_3$ (BZY) perovskite is one of the earliest identified candidates and has the highest conductivity among $BaZrO_3$ -based compounds⁷²⁻⁷⁵. The substitutions of Zr^{4+} with trivalent Y^{3+} at its B-sites yield oxygen vacancies in the lattice to enhance hydration and higher protonic conductivity. When compared to $BaCeO_3$, the chemical stability of BZY is higher and more resistant to an acidic environment. Recently, the co-sintering strategy was shown to lower the sintering temperature of $BaZr_{0.8}Y_{0.2}O_{3-\delta}$ (BZY20), which enhanced the Faradaic efficiency and stability²²⁴.

Recently, PCECs were assembled using BZY20 and $PrNi_{0.7}Co_{0.3}O_{3-\delta}$ (PNC73), which has been developed as promising oxygen electrodes^{69, 70, 180}. Such BZY20-PNC73 assemblies were then treated at different temperatures and steam concentrations to test the interface stabilities. From the top view, we observed that when the dry BZY20-PNC37 assembly (Figure 6.1A) is heated to 600°C (Figure 6.1B) and additionally hydrated by 3% steam (Figure 6.1C), the surfaces remain pristine. However, when the steam concentration is increased to 50%, PNC73 notably delaminates from BZY20 after the treatment (Figure 6.1D). Scanning electron microscope (SEM, Figures 6.1E to 6.1H) revealed that separation between BZY20 and PNC73 has already appeared in 3% steam (Figure 6.1G). The lower amplification SEM images were summarized in Figure D.1.

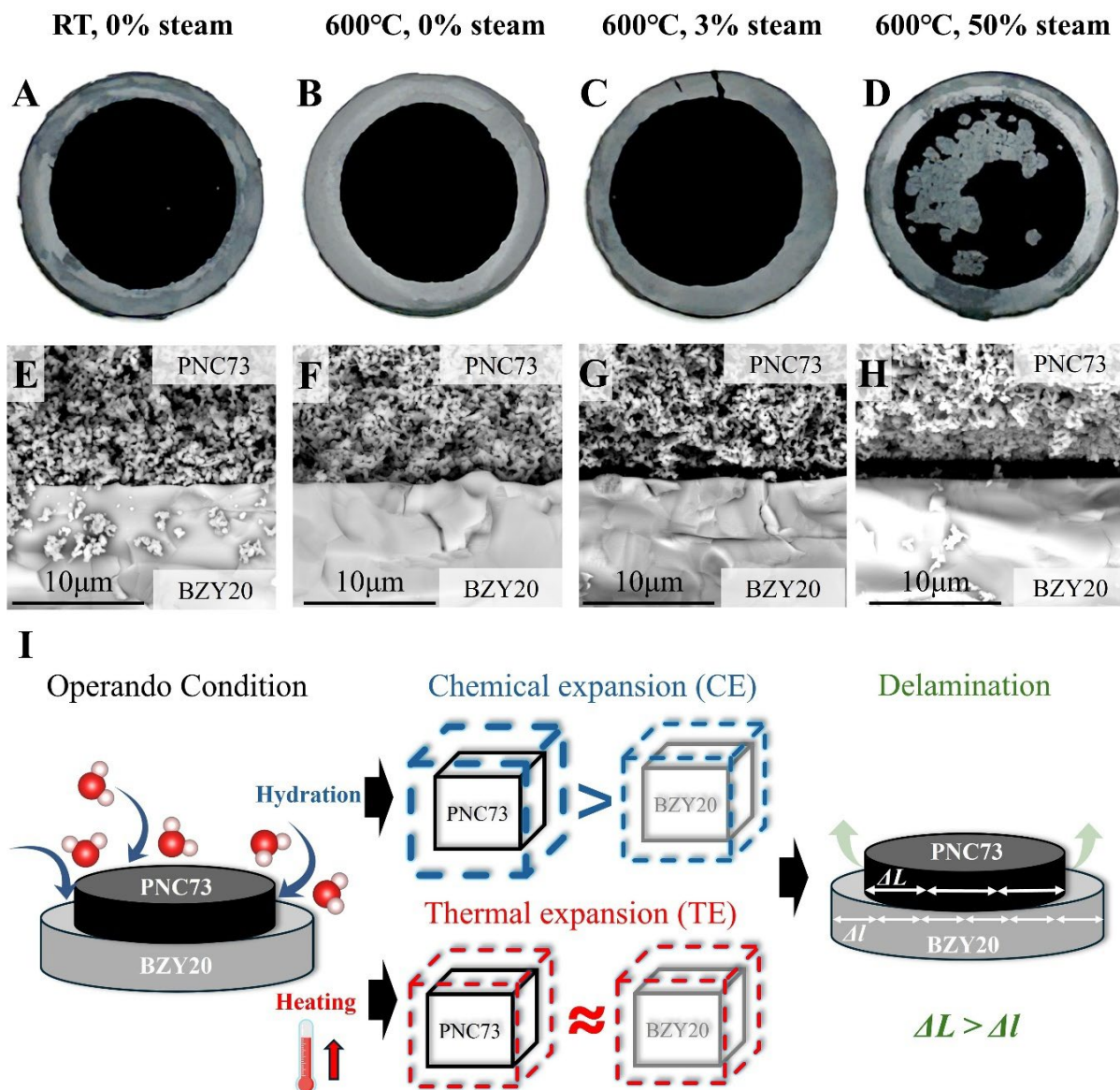


Figure 6.1 Top view (A-D) and cross-sectional SEM images of PNC73-BZY20 cells treated under (A, E) room temperature (RT) and 0% steam, (B, F) RT and 50% steam, (C, G) 600°C and 0% steam, and (D, H) 600°C and 50% steam. The black region is PNC73, and the gray region with large radius is BZY20. (I) Illustration of expansion mismatch between BZY20 and PNC73 under operando condition. The contribution from chemical expansion is higher than thermal expansion.

Based on the observed delamination, we can deduce that the chemical expansion of PNC73 under high humidity exceeds that of BZY20, despite similar thermal expansion behaviors. For high-temperature electrolysis, the instability of the BZY20-PNC73 interface highlights the need to address the electrolyte and electrode materials mismatch at high humidity (~30% steam concentration), as illustrated in Figure 6.1I. At the fundamental level, evidence suggests that thermal and chemical expansions can be traced back to the atomic environment^{202, 225-227}.

Currently, techniques to determine TE are mature and accessible. The determination of CE requires special processes such as *in-situ* X-ray diffraction (XRD)^{76, 226}. Computationally, TE are typically calculated with NPT molecular dynamics (MD) simulations,^{28, 53} coupled with phonon calculations based on the quasiharmonic approximation (QHA) methods^{199, 216}. CE (at 0 K) can be estimated directly with geometry optimizations based on first-principles methods^{203, 225, 228}. Phenomena with both TE and CE have been probed by performing QHA-based phonon calculations on pristine and defective structures¹⁹⁹. However, the existence of defects and dopants breaks the symmetry of the lattice, accompanied by significant computational burdens for phonon calculations^{107, 215}. Such a burden makes first-principles-based phonon calculations almost unachievable on doped and defective perovskite structures.

Fortunately, mathematical surrogates of quantum mechanical models in the form of machine learning potentials (MLPs) have matured in recent years^{47, 48, 99}. MLPs serve as a high-fidelity function describing complex potential energy surfaces. The computational costs of MLPs are more comparable to classical interatomic force fields, and thus, provide the opportunity to reduce computational expenses for large-scale molecular systems with multiple phases and

components. Recently, MLP was employed for BZY20 simulation^{217,218}, and phonon calculations to study their thermal properties²⁰⁷⁻²⁰⁹.

In this study, we performed MLP-based phonon calculations on BZY20 to estimate the expansion mismatch between BZY20 and the oxygen electrode. The atomic environments that are most impactful on TE and CE were identified. The thermodynamically stable 3×3×3 BZY20 supercell was screened out. Then, the contribution of TE and CE to the expansion mismatch was quantified and compared. This research can guide the manipulation of electrolyte-electrode stability and provide a roadmap to computational investigation of interfacial expansion mismatch.

6.2 Methods

6.2.1 Density Functional Theory (DFT)

Density functional theory (DFT) was employed to generate the geometries and forces included in the training set data. Vienna *Ab Initio* Simulation Package (VASP)^{8,9} was employed as the platform for all DFT calculations in this paper. The projector augmented wave (PAW) pseudopotentials⁸¹, and the GGA-PBE functional¹³⁴ were used to account for the core electrons and the electron exchange-correlation effects. For geometry optimizations, the plane-wave basis set was expanded with a cutoff energy of 500 eV and Gaussian smearing ($\sigma = 0.05$ eV), and a slightly smaller cutoff energy of 450 eV was used for single point energy calculations during OTF-MD simulations. A 4×4×4 k-point mesh was automatically generated using the Monkhorst-Pack scheme²¹³ to sample the Brillouin zone. The electronic and ionic convergence criteria are 10^{-6} eV and 0.02 eV/Å, respectively.

6.2.2 Machine Learning Potential (MLP)

We integrated on-the-fly molecular dynamics (OTF-MD) simulation³⁹⁻⁴² with the E(3)-equivariant graph neural network implemented in Allegro^{43, 44, 210, 211}. The training set was generated from the OTF-MD simulation, an active learning process based on the sparse Gaussian process (SGP). The uncertainties associated with predicted structures from the OTF-MD trajectory were quantified by the SGP's inherent mechanism. Only the structures with uncertainties higher than the preset threshold (see Table D.1) will be passed onto DFT calculations and included in the subsequent Allegro training set. The optimized $2\times 2\times 2$ BaZrO₃ (BZO) and BZY (Figure 6.2), and OTF-MD simulations were performed using all distinct defective supercells as initial input. The training set contains a total of 15026 structures. The specific numbers of structures collected from each composition are shown in Figure 6.3A. The OTF-MD of BZO resulted in fewer structures than the Y-doped BZO structures. The relatively small number of the pristine BZO component is due to its high symmetry. The data distribution for each pristine BZY system is even. For defective systems, the 2Y-sub-1 configuration owns the largest share of data since it has more distinct defective sites than other defective BZYs.

The training set was employed for final training using Allegro. The training parameters used in this study are summarized in Table D.2. We allocated 1500 structures in the data set for validation, and another 1526 structures for testing. The parity plots of energy and force predictions based on the testing set are shown in Figures 6.3B and 6.3C. The high parities and low mean absolute error (MAE) confirm the training quality. To ensure the accuracy of TE and CE simulations, the MLP was employed to predict defect formation energies (Figure 6.3D), energies under tensile and compressive strains (Figure 6.3E), and the phonon dispersion (Figure 6.3F). The MLP-predicted defect formation energies in different compositions agree with DFT (Figure 6.3D). Due to the abundance of structures in the training set, the 2Y-sub BZYs show a better agreement

than 1Y-sub BZY and BZO. The MLP-predicted equations of state (EOS) of pristine BZO, 1Y-sub, and 2Y-sub-1 BZY also fit with DFT results (Figure 6.3E). The comparisons of the EOS of defective structures (Figure D.2) show that the MLP can replicate the strain-energy relationships. Due to the high computational expense, the phonon density of states (PDOS) was only compared on BZO. Still, the MLP can yield reasonable agreement with DFT (Figure 6.3F), although only 3.3% the training set data are related to the pristine BZO.

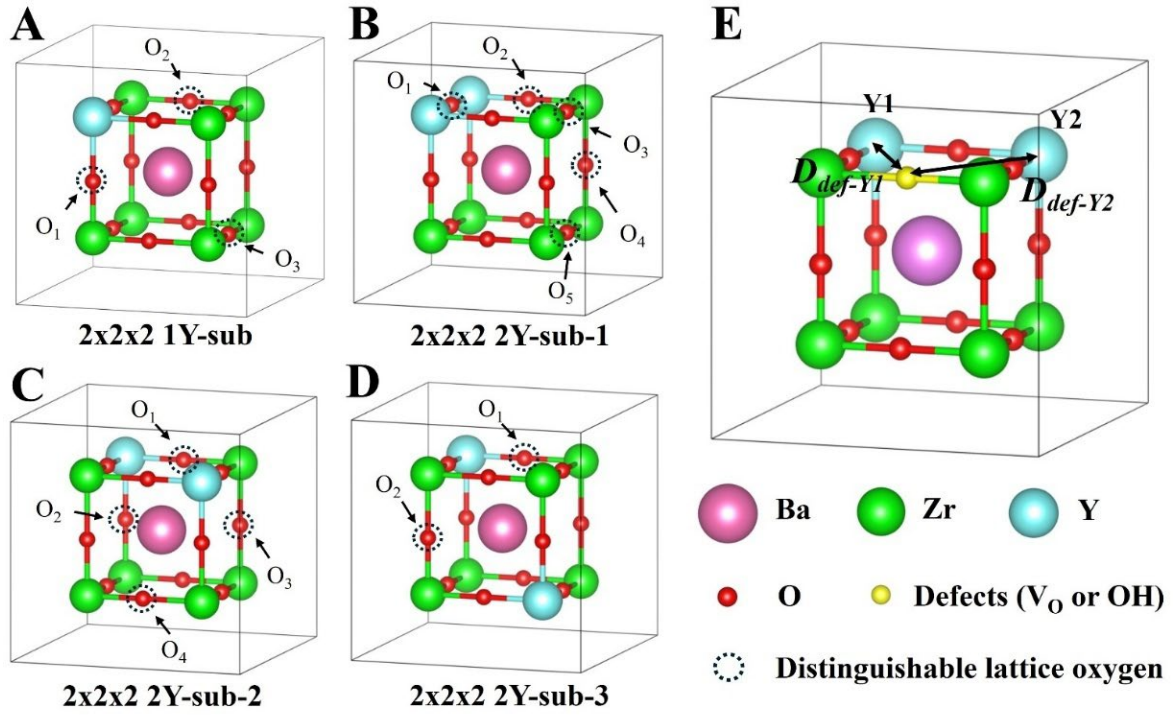


Figure 6.2 Illustration of four $2 \times 2 \times 2$ 40-atom supercells including one 1-Y substitution (A, 1Y-sub) and three 2-Y substitutions (B, 2Y-sub-1, C, 2Y-sub-2, and D, 2Y-sub-3) configurations. (E) Schematic illustration of components involved in D_{def-Y} determination as described in Eqn. 6.1. Atoms existing on periodic boundaries are omitted. Distinguishable lattice oxygen sites in supercell are indicated and indexed in the sequence of increasing D_{def-Y} .

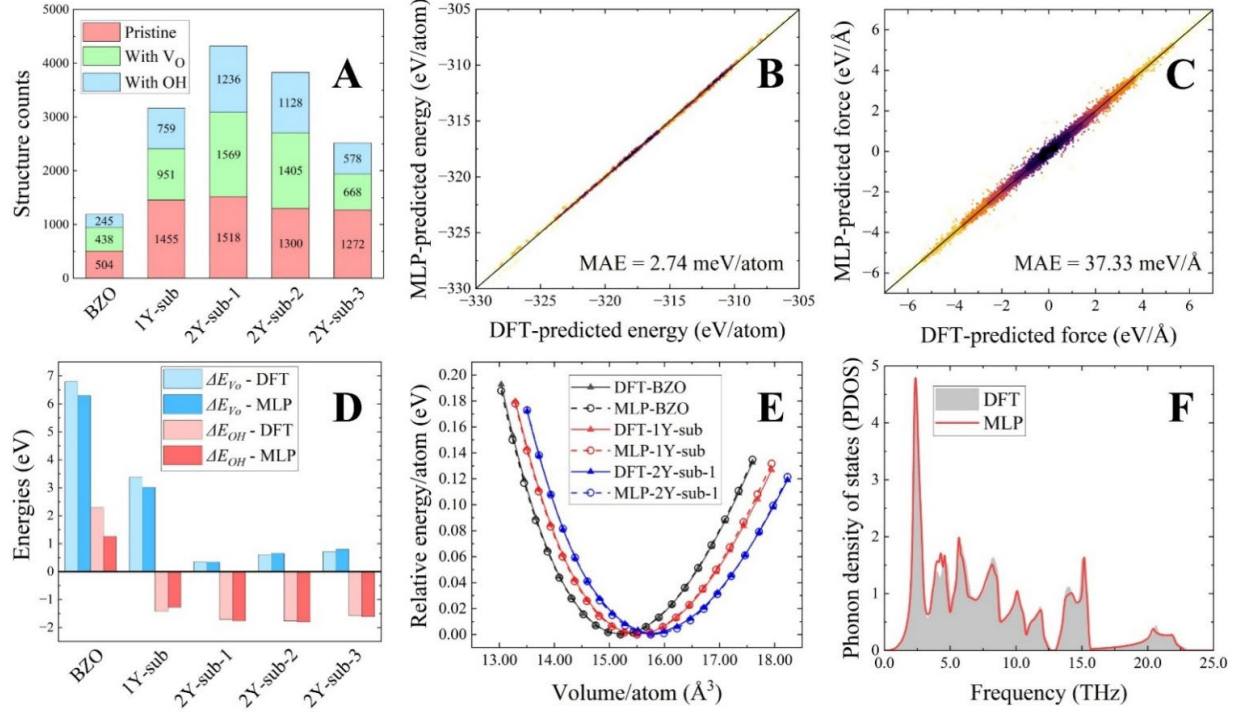


Figure 6.3 Training and benchmarking of BZY-MLP. (A) Training set composition, parity plots and mean absolute error (MAE) of (B) energy and (C) force, darker color represents higher data density region, comparison with DFT on predictions of (D) defect formation energy (ΔE), (E) equation of state (EOS), and (F) BaZrO₃ phonon density of states (PDOS).

6.2.3 Energy Minimization and Phonon Calculations

The energy minimization and force calculations were conducted based on $3 \times 3 \times 3$ supercells using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS)^{137, 229}. The energies were minimized based on the conjugate gradient (CG) algorithm with a convergence criterion of 10^{-6} eV. Phonon calculations were performed using Phonopy^{107, 169, 215}. For each configuration, a $3 \times 3 \times 3$ supercell with finite displacements of ± 0.01 Å was generated. A total of 21 scaling factors (0.95 and 1.05 with an interval of 0.005) were used to sample the compressed and tensile structures. The quasi-harmonic approximation (QHA) methods²¹⁶ were employed to

determine the equilibrium volume at temperatures between 300 and 1000 K. The equilibrium volume was finally determined by identifying the minima in the relationship between free energies (G) and cell volumes (V).

6.2.4 Material Preparation and Characterizations

The *in-situ* expansions (Figure 6.1) were probed on the PNC73-BZY20 cells fabricated in the configuration of NiO+BZCYYb1711/NiO+BZY20/BZY20/PNC73 at the Idaho National Laboratory. The synthesis details of BaZr_{0.8}Y_{0.2}O_{3-δ} (BZY20), BaZr_{0.1}Ce_{0.7}Y_{0.1}Yb_{0.1}O_{3-δ} (BZCYYb1711), and PrNi_{0.7}Co_{0.3}O_{3-δ} (PNC73) were summarized in Appendix 1 of Supporting Information. The methods of PNC73-BZY20 cell fabrication were summarized in Appendix 2.

The thermal expansion was measured on four BZY20 bar samples fabricated under different conditions, involving varied ratio of sintering aid (NiO), sintering temperature, and sintering time. Four samples were sintered at 1500°C for 20 h with 1wt% NiO (BZY20-1), 1600°C for 20 h with 1wt% NiO (BZY20-2), 1600°C for 24 h with 1wt% NiO (BZY20-3), and 1600°C for 24 h with 2wt% NiO (BZY20-4), respectively. BZY20 samples fabricated using the dry-pressed method with a 20 x 4 mm rectangle die by applying uniaxial pressure of 200 MPa. A small amount of BZY20 sacrificial powder was used to compensate the barium evaporation during the high temperature sintering process. The SEM images of BZY20 samples were shown in Figure D.3. The grain size of four samples increases in the trend of BZY20-1 < BZY20-2 < BZY20-3 < BZY20-4. The thermal expansion coefficients (TECs) of the sintered BZY20 bar samples were characterized using NETZSCH dilatometers (DIL 402 Expedit Classic).

6.3 Results and Discussion

6.3.1 Effects of Atomic Environments on Defect Formation and CE

The energetic ordering and CE of the distinct defects were probed using DFT. As shown in Figures 6.2A to 6.2D, $2 \times 2 \times 2$ BZY supercells were created with one and two Y atom substitutions, i.e., 1Y-sub and 2Y-sub BZY, corresponding to compositions of $\text{BaZr}_{0.875}\text{Y}_{0.125}\text{O}_{3-\delta}$ (BZY12.5) and $\text{BaZr}_{0.75}\text{Y}_{0.25}\text{O}_{3-\delta}$ (BZY25), respectively. The V_O -defective 1Y-sub and 2Y-sub BZY have 0.5 and 1 V_O , respectively, based on the charge neutrality. The only 1Y-sub BZY configuration possesses three distinct defective sites, and three 2Y-sub BZY configurations (2Y-sub-1, 2Y-sub-2, and 2Y-sub-3) will yield eleven defective sites. The Y-Y separation increases in the order of $2\text{Y-sub-1} < 2\text{Y-sub-2} < 2\text{Y-sub-3}$. 2Y-sub-3 has the lowest total energy, while the total energies of 2Y-sub-2 and 2Y-sub-1 are 0.08 eV and 0.29 eV higher than 2Y-sub-3, respectively. Such small energy differences and energetic ordering have been reported in the literature²³⁰. We calculated the defect formation energy (ΔE) and chemical expansion (ϵ) of $2 \times 2 \times 2$ BZY supercells based on the equation summarized in Appendix 3 in Supporting Information. In this study, the defect-yttrium distance ($D_{\text{def-Y}}$) (see Figure 6.2E), defined in Eqn. 6.1, is employed as a descriptor to understand the relationships of ΔE_i and ξ_i with the local environment in the supercell.

$$D_{\text{def-Y}} = \sqrt{\sum_i D_{\text{def-Y}_i}^2} \quad (6.1)$$

where $D_{\text{def-Y}_i}$ is the distance between the defect and the i^{th} Y atom in the supercell. Only a single defect in the supercell is considered. Therefore, $D_{\text{def-Y}_i}$ can comprehensively depict the atomic environment in a $2 \times 2 \times 2$ BZY supercell. Distinguishable lattice oxygen sites in each BZY supercell are indexed by increasing $D_{\text{def-Y}}$.

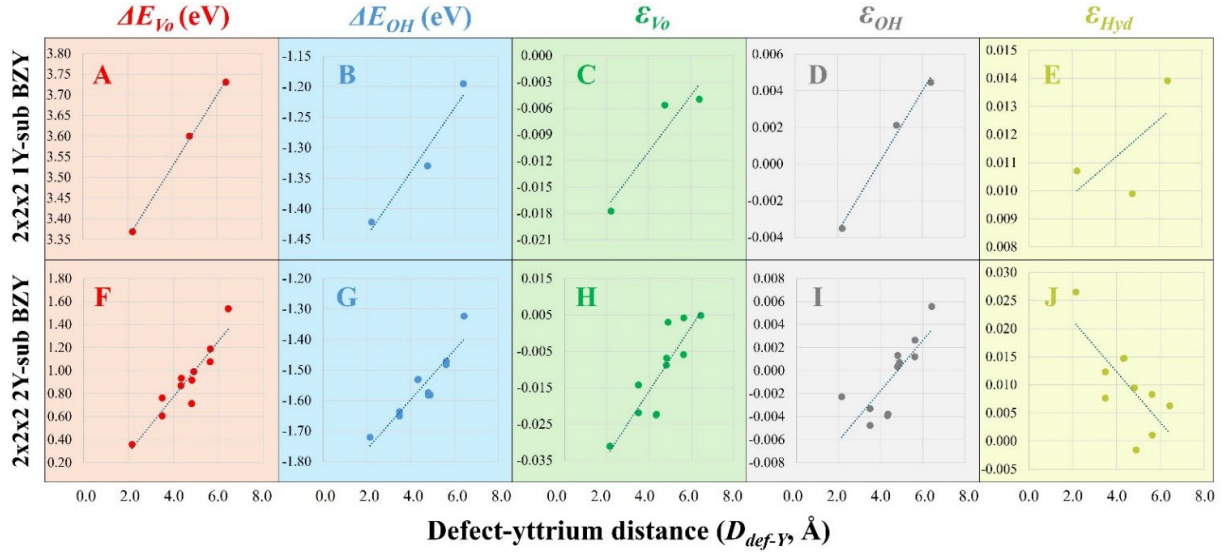


Figure 6.4 Relationships between the averaged defect-yttrium distance (D_{def-Y}) and defect formation energy (ΔE_i) or chemical expansion (ϵ_i) in $2 \times 2 \times 2$ supercells.

The ΔE and ϵ of 1Y-sub and 2Y-sub BZY at all distinct defective sites are summarized in Figure 6.4, with the numerical data and fitting parameters tabulated in Tables D.3 and D.4, respectively. From Figures 6.4A and 6.4F, it can be found that 2Y-sub BZYS have notably lower ΔE_{V_o} than 1Y-sub BZYS, confirming that higher Y concentrations are beneficial to the formation of V_o . Hence, we expect that the ΔE_{OH} of 2Y-sub BZYS (Figure 6.4G) will be lower than that of 1Y-sub BZYS (Figure 6.4B) at high Y doping levels. ϵ_{V_o} of 1Y-sub BZYS (Figure 6.4C) are negative, indicating the lattice shrinkage due to V_o formation. Meanwhile, 2Y-sub BZYS (Figure 6.4H) have three V_o sites with positive formation volumes. ϵ_{OH} can be positive or negative for either 1Y-sub (Figure 6.4D) or 2Y-sub BZYS (Figure 6.4I), and the range of ϵ_{OH} is similar in 1Y-sub or 2Y-sub BZYS. Moreover, ϵ_{V_o} is negative and has a larger magnitude than ϵ_{OH} , especially at

the low D_{def-Y} region. Therefore, the resulting ε_{hyd} is almost constantly positive, indicating a lattice expansion after hydration. This is consistent with the findings from phonon calculations for a $3\times3\times3$ supercell.

We observe that linear correlations exist between ΔE and ε and D_{def-Y} . Specifically, ΔE increases with D_{def-Y} , indicating that V_O and OH formations near Y atoms are energetically favorable. ε also exhibits a positive relationship with D_{def-Y} , indicating that the supercell volumes become smaller due to defect formations in the proximity of Y dopants. We noted that, unlike others, ε_{hyd} does not have a linear relationship as strong as other quantities, especially for 1Y-sub BZY. When calculating ε_{hyd} based on Eqn. 3 in Supporting Information, scatter of ε_{Vo} and ε_{OH} are combined, leading to a less obvious linear trend for ε_{hyd} . Smaller D_{def-Y} corresponds to higher ε_{hyd} in 2Y-sub BZYS, indicating that hydration of V_O near Y has a stronger influence over ε_{hyd} .

In summary, we can conclude that near-Y V_O sites are abundant due to higher energetic stability. Also, the OH defects are prone to form at near-Y V_O sites. Subsequently, the hydration of near-Y V_O result in larger ε_{hyd} . Therefore, we can conclude that CE is mainly contributed by the hydration of V_O near Y atoms.

6.3.2 Effects of Atomic Environments on TE

The TE of pristine and defective $2\times2\times2$ BZY supercells (Figure 6.2) is investigated with phonon calculations. As shown in Figure 6.5, the simulation results are then compared with the experimental measurements obtained for BZY12.5 and BZY25. The lattice constant of 1Y-sub BZY was calculated by averaging the lattice constant of the pristine and 1- V_O supercell, whereas the lattice constant of the 2Y-sub BZY was calculated purely based on the 1- V_O supercell.

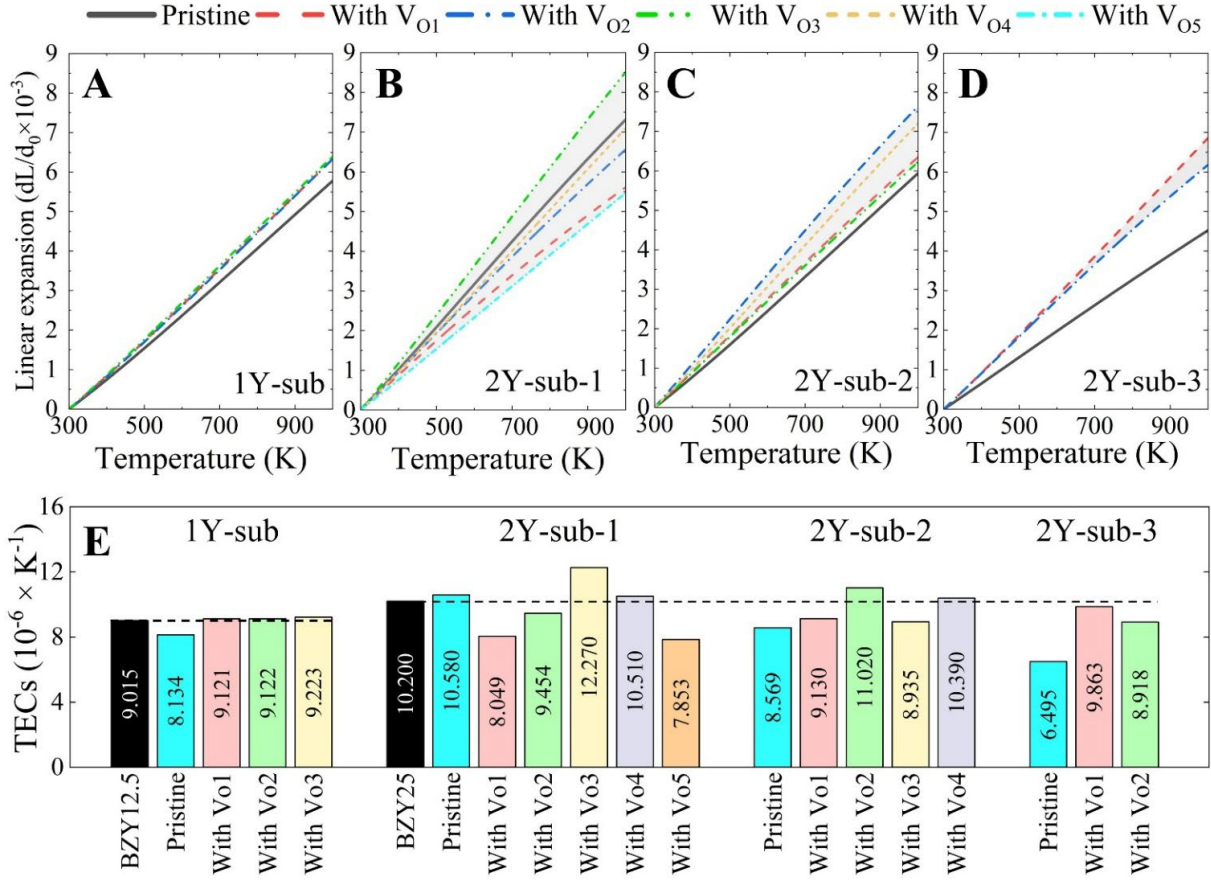


Figure 6.5 Linear expansion of pristine and V_O -defective (A) 1Y-sub, (B) 2Y-sub-1, (C) 2Y-sub-2, (D) 2Y-sub-3 2x2x2 BZY supercells, and (E) corresponding thermal expansion coefficients (TECs) compared with literature ⁷⁶ which are shown as black dash baselines.

For 1Y-sub BZY, the TE of V_O -defective structures are in excellent agreement with the experimental measurements (Figure 6.5A) ⁷⁶, proving the accuracy of MLP predictions. Upon the introduction of V_O at all distinct sites, similar TEs were observed, indicating that the TE of 1Y-sub BZY is not affected by the specific locations of V_O . However, the effects of V_O locations on the TE of 2Y-sub BZY are more complex. At increasing Y-Y distances, the TE of pristine BZY structures decreases notably (solid black lines in Figures 6.5B-D). For 2Y-sub with V_O -defects, the variations of TE become more apparent than 1Y-sub. In 2Y-sub-1, only V_{O3} results in a higher

TE, while the TE for the other four V_O sites is lower. V_O in the 2Y-sub-2 and 2Y-sub-3 cells leads to TE increase. These observations indicate that the expansions of the BZY lattices with a higher Y occupancy are more sensitive to the atomic environment, which can be characterized by the Y-Y distance and the elements surrounding the V_O .

The lattice constants of pristine and V_O -defective $2 \times 2 \times 2$ supercells were summarized in Table D.3. When the Y dopants are spatially separated (1Y-sub, 2Y-sub-2, 2Y-sub-3), the lattice retains near-cubic symmetry ($a \approx b \approx c$) and exhibits increased thermal expansion due to local lattice softening originating from weakened B–O bonding and enhanced vibrational anharmonicity, regardless of their specific locations. In contrast, when two adjacent Y dopants are aligned along the a-axis with an intervening oxygen (i.e., the Y–O–Y configuration in 2Y-sub-1), the lattice becomes anisotropic ($a < b \approx c$), and the thermal expansion becomes site-dependent. The oxygen vacancies adjacent to the Y atoms (i.e., O_1 and O_2 in Figure 6.2B) form tightly bound Y– V_O defect complexes that will increase local stiffness and suppress lattice vibrations. Hence, the observed TEC is lower. Interestingly, even for the V_O between two Zr atoms, the TEC trend diverges depending on positions of Y atoms: when Zr– V_O –Zr lies in the same a,b-plane as the Y–O–Y chain (i.e., V_O at O_3 in Figure 6.2B), the configuration promotes in-plane lattice softening and enhances thermal expansion. However, the TEC is smaller when the vacancy is positioned between Zr atoms along the a-axis but in a different a,b-plane, i.e., below the Y–O–Y layer (V_O at O_4 and O_5 in Figure 6.2B). Such an opposite trend suggests that out-of-plane strain compensation occurs in the latter case, where the lattice constrains atomic vibrations by tightening or reorienting nearby bonds to preserve structural stability, thereby reducing vibrational freedom and suppressing thermal expansion. These findings underscore the configuration dependent interplay between local lattice distortions and vibration dynamics.

6.3.3 Ordering Screening and Lattice Expansion Investigation of BZY20

The larger 3×3×3 supercells were then employed to study the simultaneous TE and CE of BZY20. Five Zr atoms were replaced by Y atoms, and three V_O were created in the supercell to match the experimental BZY stoichiometry. The resulting structure has a slightly higher Y occupation and V_O concentration than the BZY20 samples⁷⁶. The positions of Y and V_O determine the atomic ordering of 5Y-sub BZY supercells. To find the ordering that can faithfully represent the expansion behavior of BZY, we performed energy screening of 98 5Y-sub BZY supercells. The atomic ordering is characterized using geometric descriptors, including the averaged V_O-V_O distance ($D_{V_O-V_O}$), averaged Y-Y distance (D_{Y-Y}), and averaged Y coordination ($Coord_Y$) number defined by Eqns. 6.2-6.4.

$$D_{Y-Y} = \frac{1}{2} \overline{\sum_{i,j, i \neq j} D_{Y_i-Y_j}} \quad (6.2)$$

$$D_{V_O-V_O} = \frac{1}{2} \overline{\sum_{i,j, i \neq j} D_{V_{O_i}-V_{O_j}}} \quad (6.3)$$

$$Coord_Y = \sum_i \overline{Coord_{Y_i}} \quad (6.4)$$

The determinations of $D_{V_O-V_O}$ and D_{Y-Y} are illustrated in Figure D.4. $Coord_Y$ depends on the number of V_O surrounding the Y atom. If no V_O exists at the Y neighbor position, $Coord_Y$ is six.

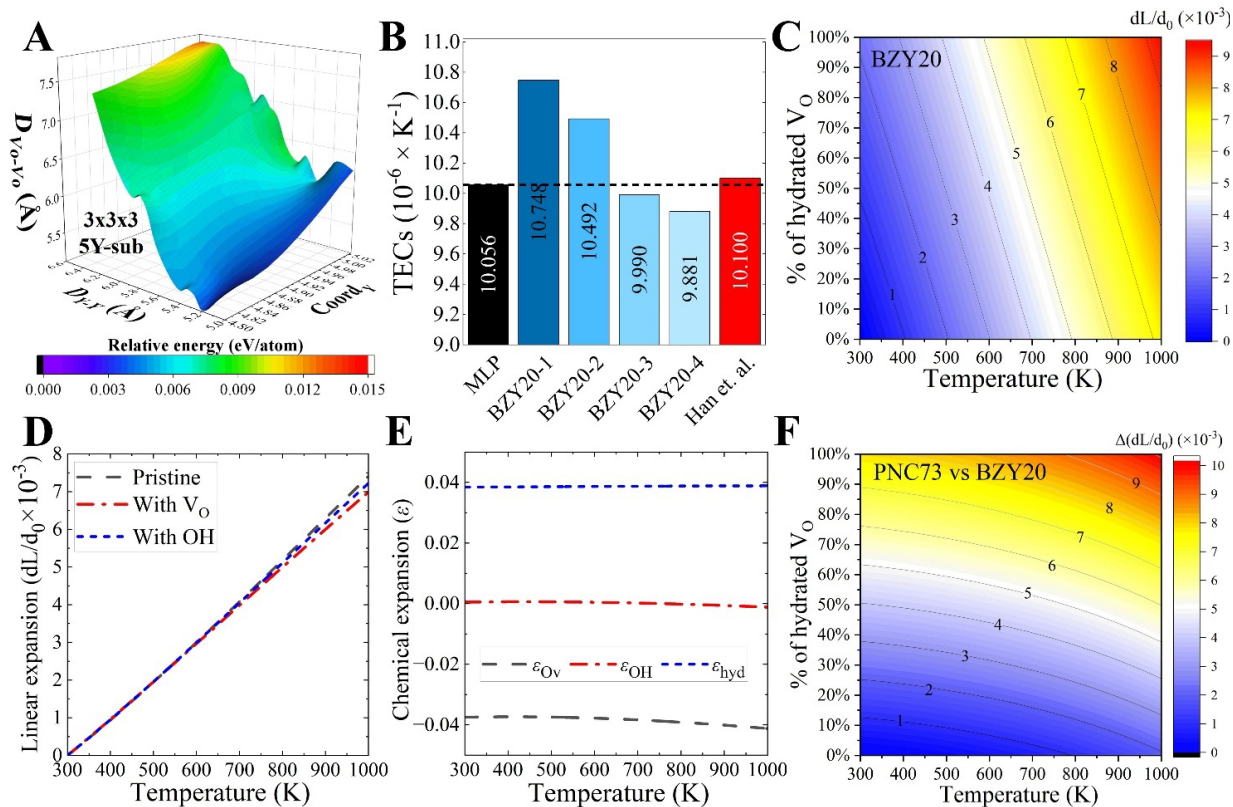


Figure 6.6 (A) Three-dimensional colormap surface illustrating energies of 5Y-sub $3 \times 3 \times 3$ BZY supercells. (B) Thermal expansion coefficients (TECs) of BZY20 ($3 \times 3 \times 3$ 5Y-sub-1 supercell) predicted by MLP and experiments from this work and literature ⁷⁶ (C) Lattice expansion of BZY20 affected by temperature and hydration, calculated based on 5Y-sub-1 supercell. (D) Thermal expansion and (E) chemical expansion of 5Y-sub-1 supercell. (F) Lattice expansion mismatch between PNC73 and BZY20.

As shown in Figure 6.6A, the energies of 98 5Y-sub BZY supercells were calculated and mapped onto a surface in the 3D space using the descriptors defined above. The surface exhibits a positive correlation between all descriptors, i.e., descriptors increase or decrease simultaneously. The highest energy was observed at the corner where all three descriptors are maximized. And the point corresponding to the lowest energy was located at the edge where $D_{V_O-V_O}$ and D_{Y-Y} are

minimized. It reveals that the $3\times3\times3$ 5Y-sub supercell with lower $D_{V_O-V_O}$, D_{Y-Y} , and $Coord_Y$ is energetically favorable. $Coord_Y$ is less impactful on thermodynamic stability than $D_{V_O-V_O}$ and D_{Y-Y} . Figure 6.6A indicates that Y atoms and V_O tend to aggregate, and V_O prefers the location near Y. The same trend was observed on a $3\times3\times3$ 3Y-sub supercell, as shown in Figure D.5.

The supercell with the lowest energy (5Y-sub-1, illustrated in Figure D.6A) was used for phonon calculations to investigate the TE and CE of BZY20. The predicted TEC was compared with experimental measurements and literature, as shown in Figure 6.6B. It can be noted that the TEC predicted from our MLP agrees most closely with the sample BZY20-3. The SEM images of four samples are summarized in Figure D.3. The grain size of samples follows the trend of BZY20-1 < BZY20-2 < BZY20-3 < BZY20-4, while the TEC follows the opposite trend. The relationship between grain size and TEC needs to be revealed in future studies. Moreover, the simulation results are also in very good agreement with the TEC reported by Han and coworkers (red bar in Figure 6.6B).⁷⁶

To further verify the necessity of employing a thermodynamically stable structure, another 5Y-sub supercell with higher energy named 5Y-sub-2 (Figure D.6B) was also employed for phonon calculation. The TECs of 5Y-sub-1 and 5Y-sub-2 are compared in Figure D.7A. The TEC of 5Y-sub-2 is $11.406\times10^{-6} \text{ K}^{-1}$, which is $0.9\times10^{-6} \text{ K}^{-1}$ higher than 5Y-sub-1 and results in notable deviation from experimental measurements. The TECs of pristine and OH-defective 5Y-sub-2 are also significantly different from 5Y-sub-1.

Further investigations on TE and CE of BZY20 were conducted using 5Y-sub-1. As shown in Figure 6.6D, the TEs of pristine and defective BZY20 are extremely close in the temperature range of 300-650 K. At temperature higher than 650 K, the TE follow the order of pristine > OH-

defective > V_O-defective. The TEC of pristine and defective 5Y-sub-1 are summarized in Figure D.7A. The TEC do not vary significantly ($< 0.617 \times 10^{-6} \text{ K}^{-1}$) in the presence of defects.

The CEs (ϵ) caused by V_O/OH formation and hydration at various temperatures are summarized in Figure 6.6E. ϵ_{V_O} remains negative throughout the temperature range, and slightly decreases at higher temperature. Quantitatively, ϵ_{OH} is negligible compared to ϵ_{V_O} . Hence, ϵ_{hyd} nearly equals the opposite of ϵ_{V_O} . This indicates that OH formation barely contributes to lattice expansion during hydration.

The lattice expansion due to simultaneous TE and CE in BZY20 is illustrated in Figure 6.6C. The expansion is determined based on the volume of hydrated lattice (V_{hyd}) according to Eqn. 6.5:

$$V_{hyd}(T) = \frac{(1-x) \cdot V_{bulk}(T) + (1-x) \cdot V_{V_O}(T) + 2x \cdot V_{OH}(T)}{2} \quad (6.5)$$

where V_{bulk} , V_{V_O} , and V_{OH} are the volume of pristine, V_O-defective, and OH-defective supercell, and x is the fraction of hydrated V_O. d_0 is the lattice constant under 300 K without hydration. As temperature increases from room temperature to 1000 K and the lattice becomes fully hydrated, the dL/d_0 increases up to 9.373×10^{-3} . The lattice expansion of 5Y-sub-1 is more sensitive to temperature increase compared to hydration. Complete hydration under 300K results $dL/d_0 = 2.137 \times 10^{-3}$, while a temperature increase of 700 K without hydration leads to $dL/d_0 = 7.197 \times 10^{-3}$. This indicates that lattice expansion BZY20 under operando conditions is mainly due to the contribution from TE. The same analysis was also performed on 5Y-sub-2, and the difference between 5Y-sub-2 and 5Y-sub-1 was illustrated in Figure D.7B. Overall, the lattice expansion predicted based on 5Y-sub-2 is higher than 5Y-sub-1, owing to higher CE contribution.

Finally, we quantified the lattice expansion mismatch ($\Delta(dL/d_0)$) between BZY20 and PNC73, as shown in Figure 6.6F. When BZY20 and PNC73 are heated to 1000K and fully

hydrated, the maximum of $\Delta(dL/d_0) = 10.10 \times 10^{-3}$, which is roughly equal to the dL/d_0 of BZY20 and indicates a severe lattice expansion mismatch. It can be found that the increase of $\Delta(dL/d_0)$ is more likely caused by hydration. Without hydration, 700K temperature increase results in $\Delta(dL/d_0) = 1.88 \times 10^{-3}$. In comparison, full hydration under 300K leads to $\Delta(dL/d_0) = 7.86 \times 10^{-3}$. It has already been shown that PNC73 is prone to expand with hydration instead of heating. Since BZY's expansion is more sensitive to temperature, we anticipate that the lattice expansion behaviors for the electrode and electrolyte will diverge, underscoring the challenge to large PCEC assembly.

6.4 Conclusion

In this paper, we reported research investigating the BZY20 expansion under *operando* conditions and quantifying the electrolyte-electrode expansion mismatch. First, we identified the popular atomic environments in BZY20 and their contribution to lattice expansion. V_O and OH tend to form closely with Y atoms, and the resulting CE is more significant than defects far from Y atoms. Under relatively low Y concentration (12.5%), the location of V_O doesn't affect TE significantly. But when Y concentration is 25%, V_O with different local environments contributes to TE variously, as some increase the TE and the other decrease it. By building three descriptors for $3 \times 3 \times 3$ 5Y-sub supercells, we screened out the most stable configuration and found that the $3 \times 3 \times 3$ 5Y-sub supercell with lower $D_{V_O-V_O}$, D_{Y-Y} , and $Coord_Y$ is more thermodynamically stable. Defect formation in BZY20 doesn't alter TE observably, and the main contributor of CE is filling V_O . Expansion of BZY20 is more sensitive to temperature increase than to hydration. Therefore, when assembled with a hydration-sensitive electrode PNC73, the interface will be destabilized mainly due to the water molecule incorporation.

Chapter 7 - Conclusions and Outlook

This thesis exhibits research based on MLP and molecular simulation techniques, including DFT, MD, and phonon calculations. Three MLP models, including high-dimensional neural network potential (HDNNP), sparse Gaussian process (SGP), and E(3)-equivariant graph neural network (Allegro), were successfully developed and applied. The systems investigated encompass transition metals, semiconductive nonmetal crystals, and proton-conducting perovskites. Vast properties were predicted using MLP-based calculations including thermal expansion (TE), chemical expansion (CE), radial distribution function (RDF), diffusivity, melting point, equation of states (EOS), enthalpy difference, phonon dispersion and density of states (PDOS), configuration thermodynamic stability, and expansion mismatch. The major achievements of this thesis are summarized below:

- An HDNNP was trained on condensed-phase Ni. The training was based on a training set from systematic AIMD simulations. The resulting HDNNP enables accurate MD simulations on the liquid and solid phases of Ni. Moreover, the melting point of condensed-phase Ni can be accurately determined from MD simulations using HDNNP.
- To estimate the thermodynamic phase stability of icosahedral boron allotropes, an SGP potential was trained using the OTF-MD, which enables efficient structural sampling on the characteristic icosahedral boron. The SGP potential showed good accuracy for a range of thermodynamic, structural, and vibrational property predictions. The predicted phase diagram agrees with experimental measurements. The stabilization effect on β -B, originating from the B13 defect, was identified based on thermodynamic calculations using the SGP potential.

- The thermal and chemical expansions of proton-conducting $\text{PrNi}_x\text{Co}_{1-x}\text{O}_{3-\delta}$ (PNC) electrode were predicted by MLP-based phonon calculations. The training set was also generated from OTF-MD. It is revealed that thermal expansions are related to Co occupancy, while chemical expansions are more closely associated with the Ni occupancy. We showed that some strategies solely designed to boost the PCEC's electrochemical performance may be detrimental to the electrode-electrolyte interfacial stability, as high Ni occupancy may result in incongruent chemical expansions with the electrolyte.
- Finally, the proton-conducting electrolyte $\text{BaZr}_{0.8}\text{Y}_{0.2}\text{O}_{3-\delta}$ (BZY20) expansion behaviors were investigated. The popular atomic environments in BZY20 and their contributions to lattice expansion were identified. The lattice expansion of BZY20 is more sensitive to temperature increase than hydration. Therefore, in the assembly consisting of hydration-sensitive electrode PNC73, the interface will likely be destabilized due to the water molecule incorporation.

The thesis demonstrates that, as a mathematical surrogate of first-principles methods, MLP is a powerful tool to accelerate large-scale molecular simulations. We are witnessing a rapid growth of the research community that continuously drives the advancement of MLP methodologies. Nevertheless, there are still critical issues that need to be addressed:

- Extrapolation (i.e., predictability) of MLPs is essential for systems that deviate from the training set during MD simulations. Challenging problems include phase transition, ionic diffusion, adsorption, and so on. Generating a training set encompassing all possible configurations is impractical. Therefore, it is desirable to have MLPs with extrapolation ability to retain accuracy when MD trajectories wander beyond the

training set. We have observed some promising evidence that MLPs are capable of extrapolating configurational states within a reasonable range. As discussed in Chapter 4, the SGP MLP is able to capture the stabilization effect associated with the B13 defect within β -B. Recently, developments of universal MLPs have been reported^{231, 232}. Universal MLPs are trained using the most extensive data set available and cover a larger number of elements. For example, Rodriguez et. al. developed the Elemental Spatial Density Neural Network Force Field, namely Elemental-SDNNFF, which was trained on 9.4 million atomic data spanning 55 elements²³¹. Still, these universal MLP models are limited to a specific class of materials, such as transitional metal oxides.

- It is essential to have a universally accepted protocol for MLP assessment and benchmarking. MLPs are mathematical models, and many trainable parameters are not physically interpretable. A comprehensive validation of MLP accuracy is vital before conducting any production simulations. However, one specific benchmarking calculation can only show the accuracy from a particular aspect. Therefore, the protocol that comprehensively exhibits the MLP accuracy is essential for modelers to develop an MLP with confidence.
- The ease of compilation is ideal for researchers with less knowledge of computer science. For people with backgrounds in chemistry and materials, compiling MLP codes and corresponding MD tools may be challenging. Especially, to achieve high computational efficiency, current MLP training and MD simulations heavily rely on GPUs, which bring more complexity during compilation. Therefore, it's beneficial to users if the compilation is designed to be straightforward during the development process.

Based on the current progression of my research, several other topics can be explored using the MLP techniques. In addition to mechanical stability, surface reactions and conduction mechanisms are crucial for the PCEC assembly to meet performance targets as well.

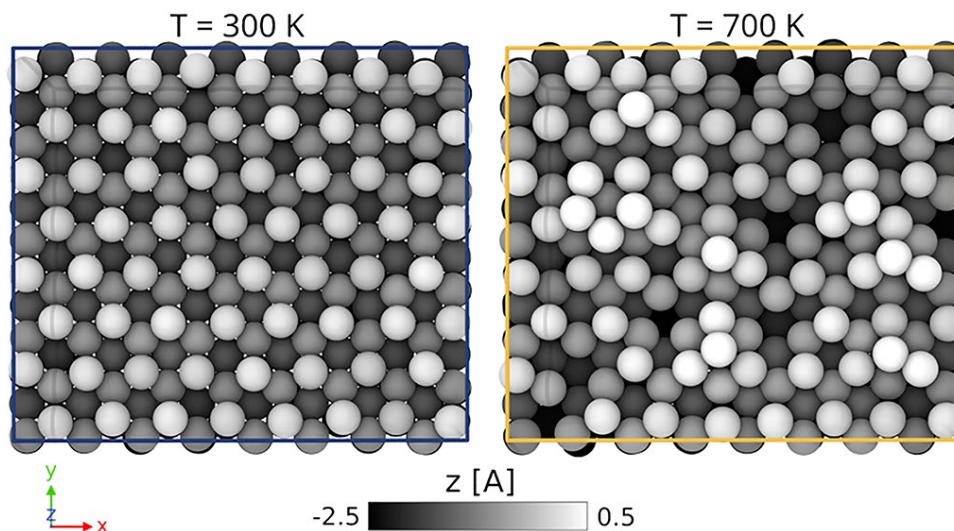


Figure 7.1 Fe(111) morphology: room temperature vs. in operando condition. Top view of final structures obtained from 20-ns molecular dynamics simulations. Bright regions indicate adatoms, while dark regions the formation of vacancies on the surface. Adapted from Ref. [233].

The oxygen evolution reaction (OER) and oxygen reduction reaction (ORR) are essential for the oxygen electrode design. Under operando conditions, the morphology of the oxygen electrode is affected by temperature, steam concentration, and electric field. The altered surface structure will directly affect the reaction mechanism since the surface site will change dramatically. MLP-based MD can probe the surface evolution. As shown in Figure 7.1, the morphology change of Fe(111) was observed by a 20-ns MLP-based MD simulation. We are confident that MLP-based MD simulations will be superior to elucidate the mechanism of OER and ORR. What we can learn from such MLP-based MD simulations includes the sites with the

highest reactivity, the evolution of the surface structures, and the reaction rates under realistic conditions (i.e., temperature, steam concentration, and electric field).

For electrolytes, investigating proton conduction is of high importance. AIMD has been the standard tool for analyzing the proton transportation pathway, as shown in Figure 7.2. The proton can move around lattice oxygens and hop to neighboring oxygens. AIMD simulations are computationally expensive, so they are only used for small, regular systems and short periods. On the other hand, traditional interatomic potentials are highly unlikely to accurately describe the complex interatomic forces of state-of-the-art electrolytes (e.g., BCZYYb). Therefore, MLP-based MD will probably be one of the most appropriate tools for investigating proton conduction in PCEC electrolytes.

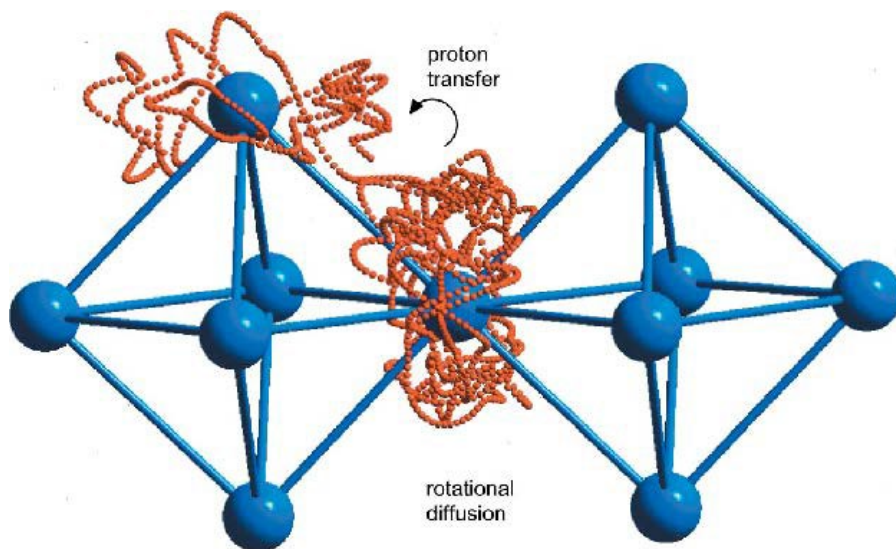


Figure 7.2 The trace of a proton in a perovskite showing the two principle features of proton transport: rotational diffusion and proton transfer (data have been obtained by a AIMD simulation of a proton in cubic BaCeO_3). Adapted from Ref. [198].

References

- (1) Yusuf, V. F.; Malek, N. I.; Kailasa, S. K. Review on Metal–Organic Framework Classification, Synthetic Approaches, and Influencing Factors: Applications in Energy, Drug Delivery, and Wastewater Treatment. *ACS Omega* **2022**, *7* (49), 44507-44531. DOI: 10.1021/acsomega.2c05310.
- (2) Novoselov, K. S.; Geim, A. K.; Morozov, S. V.; Jiang, D.; Zhang, Y.; Dubonos, S. V.; Grigorieva, I. V.; Firsov, A. A. Electric Field Effect in Atomically Thin Carbon Films. *Science* **2004**, *306* (5696), 666-669. DOI: doi:10.1126/science.1102896.
- (3) Bhalla, A. S.; Guo, R.; Roy, R. The perovskite structure – a review of its role in ceramic science and technology. *Material Research Innovations* **2000**, *4* (1), 3-26. DOI: 10.1007/s100190000062.
- (4) George, E. P.; Raabe, D.; Ritchie, R. O. High-entropy alloys. *Nature Reviews Materials* **2019**, *4* (8), 515-534. DOI: 10.1038/s41578-019-0121-4.
- (5) Wang, H.; Yu, Z.; Kong, X.; Kim, S. C.; Boyle, D. T.; Qin, J.; Bao, Z.; Cui, Y. Liquid electrolyte: The nexus of practical lithium metal batteries. *Joule* **2022**, *6* (3), 588-616. DOI: <https://doi.org/10.1016/j.joule.2021.12.018>.
- (6) Clarivate. *Web of Science*. 2025. <https://www.webofscience.com/> (accessed 2025).
- (7) Alder, B. J.; Wainwright, T. E. Phase Transition for a Hard Sphere System. *The Journal of Chemical Physics* **1957**, *27* (5), 1208-1209. DOI: 10.1063/1.1743957 (accessed 4/3/2025).
- (8) Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational materials science* **1996**, *6* (1), 15-50.
- (9) Kresse, G.; Hafner, J. Ab initio molecular dynamics for open-shell transition metals. *Physical Review B* **1993**, *48* (17), 13115.

- (10) Hartree, D. R. The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part II. Some Results and Discussion. *Mathematical Proceedings of the Cambridge Philosophical Society* **1928**, *24* (1), 111-132. DOI: 10.1017/S0305004100011920 From Cambridge University Press Cambridge Core.
- (11) Alder, B. J.; Wainwright, T. E. Studies in Molecular Dynamics. I. General Method. *The Journal of Chemical Physics* **1959**, *31* (2), 459-466. DOI: 10.1063/1.1730376 (accessed 4/3/2025).
- (12) Metropolis, N.; Ulam, S. The Monte Carlo method. *J Am Stat Assoc* **1949**, *44* (247), 335-341. DOI: 10.1080/01621459.1949.10483310 From NLM.
- (13) Parr, R. G. Density Functional Theory of Atoms and Molecules. In *Horizons of Quantum Chemistry*, Dordrecht, 1980//, 1980; Fukui, K., Pullman, B., Eds.; Springer Netherlands: pp 5-15.
- (14) Kohn, W.; Becke, A. D.; Parr, R. G. Density Functional Theory of Electronic Structure. *The Journal of Physical Chemistry* **1996**, *100* (31), 12974-12980. DOI: 10.1021/jp960669l.
- (15) Fernando, G. W. Chapter 6 - Generalized Kohn–Sham Density Functional Theory via Effective Action Formalism. In *Handbook of Metal Physics*, Fernando, G. W. Ed.; Vol. 4; Elsevier, 2008; pp 131-156.
- (16) Guan, H.; Sun, H.; Zhao, X. Application of Density Functional Theory to Molecular Engineering of Pharmaceutical Formulations. *International Journal of Molecular Sciences* **2025**, *26* (7), 3262.
- (17) Shan, N.; Zhou, M.; Hanchett, M. K.; Chen, J.; Liu, B. Practical principles of density functional theory for catalytic reaction simulations on metal surfaces – from theory to applications. *Mol Simulat* **2017**, *43* (10-11), 861-885. DOI: 10.1080/08927022.2017.1303687.

- (18) Jain, A.; Shin, Y.; Persson, K. A. Computational predictions of energy materials using density functional theory. *Nature Reviews Materials* **2016**, *1* (1), 15004. DOI: 10.1038/natrevmats.2015.4.
- (19) Schleder, G. R.; Padilha, A. C. M.; Acosta, C. M.; Costa, M.; Fazzio, A. From DFT to machine learning: recent approaches to materials science—a review. *Journal of Physics: Materials* **2019**, *2* (3), 032001. DOI: 10.1088/2515-7639/ab084b.
- (20) Karplus, M.; McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nature Structural Biology* **2002**, *9* (9), 646-652. DOI: 10.1038/nsb0902-646.
- (21) McCammon, J. A.; Gelin, B. R.; Karplus, M. Dynamics of folded proteins. *Nature* **1977**, *267* (5612), 585-590. DOI: 10.1038/267585a0.
- (22) Jones, J. E.; Chapman, S. On the determination of molecular fields. —II. From the equation of state of a gas. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **1924**, *106* (738), 463-477. DOI: doi:10.1098/rspa.1924.0082.
- (23) Daw, M. S.; Baskes, M. I. Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals. *Physical Review B* **1984**, *29* (12), 6443-6453. DOI: 10.1103/PhysRevB.29.6443.
- (24) Senftle, T. P.; Hong, S.; Islam, M. M.; Kylasa, S. B.; Zheng, Y.; Shin, Y. K.; Junkermeier, C.; Engel-Herbert, R.; Janik, M. J.; Aktulga, H. M.; et al. The ReaxFF reactive force-field: development, applications and future directions. *npj Computational Materials* **2016**, *2* (1), 15011. DOI: 10.1038/npjcompumats.2015.11.
- (25) Behler, J. Constructing high - dimensional neural network potentials: A tutorial review. *International Journal of Quantum Chemistry* **2015**, *115* (16), 1032-1050.

- (26) Behler, J. First principles neural network potentials for reactive simulations of large molecular and condensed systems. *Angewandte Chemie International Edition* **2017**, *56* (42), 12828-12840.
- (27) Huang, X.; Shen, Y.; An, Q. Nanotwinning induced decreased lattice thermal conductivity of high temperature thermoelectric boron subphosphide (B12P2) from deep learning potential simulations. *Energy and AI* **2022**, *8*, 100135. DOI: <https://doi.org/10.1016/j.egyai.2022.100135>.
- (28) Deng, H.; Comer, J.; Liu, B. A high-dimensional neural network potential for molecular dynamics simulations of condensed phase nickel and phase transitions. *Mol Simulat* **2023**, *49* (3), 263-270. DOI: 10.1080/08927022.2022.2156561.
- (29) Thompson, A. P.; Swiler, L. P.; Trott, C. R.; Foiles, S. M.; Tucker, G. J. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *Journal of Computational Physics* **2015**, *285*, 316-330. DOI: <https://doi.org/10.1016/j.jcp.2014.12.018>.
- (30) Chen, C.; Deng, Z.; Tran, R.; Tang, H.; Chu, I.-H.; Ong, S. P. Accurate force field for molybdenum by machine learning large materials data. *Physical Review Materials* **2017**, *1* (4), 043603. DOI: 10.1103/PhysRevMaterials.1.043603.
- (31) Li, X. G.; Hu, C. Z.; Chen, C.; Deng, Z.; Luo, J.; Ong, S. P. Quantum-accurate spectral neighbor analysis potential models for Ni-Mo binary alloys and fcc metals. *Physical Review B* **2018**, *98* (9). DOI: 10.1103/PhysRevB.98.094104.
- (32) Deng, Z.; Chen, C.; Li, X.-G.; Ong, S. P. An electrostatic spectral neighbor analysis potential for lithium nitride. *npj Computational Materials* **2019**, *5* (1), 75. DOI: 10.1038/s41524-019-0212-1.
- (33) Bartók, A. P.; Kondor, R.; Csányi, G. On representing chemical environments. *Physical Review B* **2013**, *87* (18), 184115. DOI: 10.1103/PhysRevB.87.184115.

- (34) Szlachta, W. J.; Bartók, A. P.; Csányi, G. Accuracy and transferability of Gaussian approximation potential models for tungsten. *Physical Review B* **2014**, *90* (10), 104108. DOI: 10.1103/PhysRevB.90.104108.
- (35) Dragoni, D.; Daff, T. D.; Csányi, G.; Marzari, N. Achieving DFT accuracy with a machine-learning interatomic potential: Thermomechanics and defects in bcc ferromagnetic iron. *Physical Review Materials* **2018**, *2* (1), 013808. DOI: 10.1103/PhysRevMaterials.2.013808.
- (36) Shapeev, A. V. Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials. *Multiscale Modeling & Simulation* **2016**, *14* (3), 1153-1173. DOI: 10.1137/15m1054183.
- (37) Podryabinkin, E. V.; Shapeev, A. V. Active learning of linearly parametrized interatomic potentials. *Computational Materials Science* **2017**, *140*, 171-180. DOI: <https://doi.org/10.1016/j.commatsci.2017.08.031>.
- (38) Gubaev, K.; Podryabinkin, E. V.; Hart, G. L. W.; Shapeev, A. V. Accelerating high-throughput searches for new alloys with active learning of interatomic potentials. *Computational Materials Science* **2019**, *156*, 148-156. DOI: <https://doi.org/10.1016/j.commatsci.2018.09.031>.
- (39) Vandermause, J.; Torrisi, S. B.; Batzner, S.; Xie, Y.; Sun, L.; Kolpak, A. M.; Kozinsky, B. On-the-fly active learning of interpretable Bayesian force fields for atomistic rare events. *npj Computational Materials* **2020**, *6* (1), 20. DOI: 10.1038/s41524-020-0283-z.
- (40) Xie, Y.; Vandermause, J.; Sun, L.; Cepellotti, A.; Kozinsky, B. Bayesian force fields from active learning for simulation of inter-dimensional transformation of stanene. *npj Computational Materials* **2021**, *7* (1), 40. DOI: 10.1038/s41524-021-00510-y.

- (41) Vandermause, J.; Xie, Y.; Lim, J. S.; Owen, C. J.; Kozinsky, B. Active learning of reactive Bayesian force fields applied to heterogeneous catalysis dynamics of H/Pt. *Nature Communications* **2022**, *13* (1), 5183. DOI: 10.1038/s41467-022-32294-0.
- (42) Xie, Y.; Vandermause, J.; Ramakers, S.; Protik, N. H.; Johansson, A.; Kozinsky, B. Uncertainty-aware molecular dynamics from Bayesian active learning for phase transformations and thermal transport in SiC. *npj Computational Materials* **2023**, *9* (1), 36. DOI: 10.1038/s41524-023-00988-8.
- (43) Batzner, S.; Musaelian, A.; Sun, L.; Geiger, M.; Mailoa, J. P.; Kornbluth, M.; Molinari, N.; Smidt, T. E.; Kozinsky, B. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature Communications* **2022**, *13* (1), 2453. DOI: 10.1038/s41467-022-29939-5.
- (44) Musaelian, A.; Batzner, S.; Johansson, A.; Sun, L.; Owen, C. J.; Kornbluth, M.; Kozinsky, B. Learning local equivariant representations for large-scale atomistic dynamics. *Nature Communications* **2023**, *14* (1), 579. DOI: 10.1038/s41467-023-36329-y.
- (45) Johansson, A.; Xie, Y.; Owen, C. J.; Lim, J. S.; Sun, L.; Vandermause, J.; Kozinsky, B. Micron-scale heterogeneous catalysis with Bayesian force fields from first principles and active learning. *arXiv preprint arXiv:2204.12573* **2022**.
- (46) Mishin, Y. Machine-learning interatomic potentials for materials science. *Acta Materialia* **2021**, *214*, 116980. DOI: <https://doi.org/10.1016/j.actamat.2021.116980>.
- (47) Jacobs, R.; Morgan, D.; Attarian, S.; Meng, J.; Shen, C.; Wu, Z.; Xie, C. Y.; Yang, J. H.; Artrith, N.; Blaiszik, B.; et al. A practical guide to machine learning interatomic potentials – Status and future. *Current Opinion in Solid State and Materials Science* **2025**, *35*, 101214. DOI: <https://doi.org/10.1016/j.cossms.2025.101214>.

- (48) Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltaysky, I.; Schutt, K. T.; Tkatchenko, A.; Muller, K. R. Machine Learning Force Fields. *Chemical Reviews* **2021**, *121* (16), 10142-10186. DOI: 10.1021/acs.chemrev.0c01111.
- (49) Wang, G.; Wang, C.; Zhang, X.; Li, Z.; Zhou, J.; Sun, Z. Machine learning interatomic potential: Bridge the gap between small-scale models and realistic device-scale simulations. *iScience* **2024**, *27* (5), 109673. DOI: <https://doi.org/10.1016/j.isci.2024.109673>.
- (50) Huu Kien, P. Study of Structural and Phase Transition of Nickel Metal. *International Scholarly Research Notices* **2014**, *2014* (1), 253627. DOI: <https://doi.org/10.1155/2014/253627>.
- (51) Liu, T.; Chen, L.; Li, W.; Liu, Z.; Zhang, J.; Zhang, X.; Zhang, X.; Zhu, S.; Hou, X. Orientation-dependent phase transition pathways of single-crystal nickel over large shock range. *International Journal of Mechanical Sciences* **2024**, *261*, 108689. DOI: <https://doi.org/10.1016/j.ijmecsci.2023.108689>.
- (52) Hamilton, J. C.; Jach, T. Structural Phase Transitions in Nickel at the Curie Temperature. *Physical Review Letters* **1981**, *46* (11), 745-748. DOI: 10.1103/PhysRevLett.46.745.
- (53) Shin, Y. K.; Gao, Y.; Shin, D.; van Duin, A. C. T. Impact of three-body interactions in a ReaxFF force field for Ni and Cr transition metals and their alloys on the prediction of thermal and mechanical properties. *Computational Materials Science* **2021**, *197*, 110602. DOI: <https://doi.org/10.1016/j.commatsci.2021.110602>.
- (54) Guo, J.; Lou, H.; Zhao, H.; Chai, D.; Zheng, X. Dry reforming of methane over nickel catalysts supported on magnesium aluminate spinels. *Applied Catalysis A: General* **2004**, *273* (1), 75-82. DOI: <https://doi.org/10.1016/j.apcata.2004.06.014>.
- (55) Zhou, M.; Liu, B. DFT Investigation on the Competition of the Water–Gas Shift Reaction Versus Methanation on Clean and Potassium-Modified Nickel(1 1 1) Surfaces. *ChemCatChem*

2015, 7 (23), 3928-3935, <https://doi.org/10.1002/cctc.201500547>. DOI: <https://doi.org/10.1002/cctc.201500547> (accessed 2022/06/03).

(56) Liu, S.; He, R.; Xue, L.; Li, J.; Liu, B.; Edgar, J. H. Single Crystal Growth of Millimeter-Sized Monoisotopic Hexagonal Boron Nitride. *Chemistry of Materials* **2018**, 30 (18), 6222-6225. DOI: 10.1021/acs.chemmater.8b02589.

(57) Liu, S.; van Duin, A. C. T.; van Duin, D. M.; Liu, B.; Edgar, J. H. Atomistic Insights into Nucleation and Formation of Hexagonal Boron Nitride on Nickel from First-Principles-Based Reactive Molecular Dynamics Simulations. *ACS Nano* **2017**, 11 (4), 3585-3596. DOI: 10.1021/acsnano.6b06736.

(58) Douglas, B.; Ho, S.-M. *Structure and chemistry of crystalline solids*; Springer Science & Business Media, 2007.

(59) Ogitsu, T.; Schwegler, E.; Galli, G. β -Rhombohedral Boron: At the Crossroads of the Chemistry of Boron and the Physics of Frustration. *Chemical Reviews* **2013**, 113 (5), 3425-3449. DOI: 10.1021/cr300356t.

(60) Ogitsu, T.; Gygi, F.; Reed, J.; Motome, Y.; Schwegler, E.; Galli, G. Imperfect Crystal and Unusual Semiconductor: Boron, a Frustrated Element. *Journal of the American Chemical Society* **2009**, 131 (5), 1903-1909. DOI: 10.1021/ja807622w.

(61) Oganov, A. R.; Chen, J.; Gatti, C.; Ma, Y.; Ma, Y.; Glass, C. W.; Liu, Z.; Yu, T.; Kurakevych, O. O.; Solozhenko, V. L. Ionic high-pressure form of elemental boron. *Nature* **2009**, 457 (7231), 863-867. DOI: 10.1038/nature07736.

(62) Zarechnaya, E. Y.; Dubrovinsky, L.; Dubrovinskaia, N.; Filinchuk, Y.; Chernyshov, D.; Dmitriev, V.; Miyajima, N.; El Goresy, A.; Braun, H. F.; Van Smaalen, S.; et al. Superhard

Semiconducting Optically Transparent High Pressure Phase of Boron. *Physical Review Letters* **2009**, *102* (18), 185501. DOI: 10.1103/PhysRevLett.102.185501.

(63) Parakhonskiy, G.; Dubrovinskaia, N.; Bykova, E.; Wirth, R.; Dubrovinsky, L. Experimental pressure-temperature phase diagram of boron: resolving the long-standing enigma. *Scientific Reports* **2011**, *1* (1), 96. DOI: 10.1038/srep00096.

(64) Duan, C.; Kee, R.; Zhu, H.; Sullivan, N.; Zhu, L.; Bian, L.; Jennings, D.; O'Hayre, R. Highly efficient reversible protonic ceramic electrochemical cells for power generation and fuel production. *Nature Energy* **2019**, *4* (3), 230-240. DOI: 10.1038/s41560-019-0333-2.

(65) Zhang, Y.; Chen, B.; Guan, D.; Xu, M.; Ran, R.; Ni, M.; Zhou, W.; O'Hayre, R.; Shao, Z. Thermal-expansion offset for high-performance fuel cell cathodes. *Nature* **2021**, *591* (7849), 246-251. DOI: 10.1038/s41586-021-03264-1.

(66) Liu, Z.; Zhang, L.; Zheng, C. J.; Zhang, Y.; Chen, B.; Shao, Z.; Ge, J. Advanced Electrode Materials for Efficient Hydrogen Production in Protonic Ceramic Electrolysis Cells. *Advanced Materials* n/a (n/a), 2503609. DOI: <https://doi.org/10.1002/adma.202503609>.

(67) Zhang, W.; Liu, M.; Gu, X.; Shi, Y.; Deng, Z.; Cai, N. Water Electrolysis toward Elevated Temperature: Advances, Challenges and Frontiers. *Chemical Reviews* **2023**, *123* (11), 7119-7192. DOI: 10.1021/acs.chemrev.2c00573.

(68) Cao, J.; Ji, Y.; Shao, Z. Nanotechnologies in ceramic electrochemical cells. *Chemical Society Reviews* **2024**, *53* (1), 450-501, 10.1039/D3CS00303E. DOI: 10.1039/D3CS00303E.

(69) Ding, H.; Wu, W.; Jiang, C.; Ding, Y.; Bian, W.; Hu, B.; Singh, P.; Orme, C. J.; Wang, L.; Zhang, Y.; et al. Self-sustainable protonic ceramic electrochemical cells using a triple conducting electrode for hydrogen and power production. *Nature Communications* **2020**, *11* (1), 1907. DOI: 10.1038/s41467-020-15677-z.

- (70) Tang, W.; Ding, H.; Bian, W.; Regalado Vera, C. Y.; Gomez, J. Y.; Dong, Y.; Li, J.; Wu, W.; Fan, W.; Zhou, M.; et al. An Unbalanced Battle in Excellence: Revealing Effect of Ni/Co Occupancy on Water Splitting and Oxygen Reduction Reactions in Triple-Conducting Oxides for Protonic Ceramic Electrochemical Cells. *Small* **2022**, *18* (30), 2201953. DOI: <https://doi.org/10.1002/sml.202201953>.
- (71) Zhong, Z.; Song, T.; Zhao, S.; Sun, H.; Guo, X.; Feng, Y.; Hu, Q. High-performance BaZr_{0.1}Ce_{0.7}Y_{0.1}Yb_{0.1}O_{3-δ} (BZCYYb) protonic ceramic fuel cell electrolytes by the Ba evaporation inhibition strategy. *Ceramics International* **2024**, *50* (2, Part B), 3633-3640. DOI: <https://doi.org/10.1016/j.ceramint.2023.11.113>.
- (72) Regalado Vera, C. Y.; Ding, H.; Peterson, D.; Gibbons, W. T.; Zhou, M.; Ding, D. A mini-review on proton conduction of BaZrO₃-based perovskite electrolytes. *Journal of Physics: Energy* **2021**, *3* (3), 032019. DOI: 10.1088/2515-7655/ac12ab.
- (73) Kreuer, K. D.; Adams, S.; Münch, W.; Fuchs, A.; Klock, U.; Maier, J. Proton conducting alkaline earth zirconates and titanates for high drain electrochemical applications. *Solid State Ionics* **2001**, *145* (1), 295-306. DOI: [https://doi.org/10.1016/S0167-2738\(01\)00953-5](https://doi.org/10.1016/S0167-2738(01)00953-5).
- (74) Han, D.; Shinoda, K.; Sato, S.; Majima, M.; Uda, T. Correlation between electroconductive and structural properties of proton conductive acceptor-doped barium zirconate. *Journal of Materials Chemistry A* **2015**, *3* (3), 1243-1250, 10.1039/C4TA05701E. DOI: 10.1039/C4TA05701E.
- (75) Yamazaki, Y.; Hernandez-Sanchez, R.; Haile, S. M. High Total Proton Conductivity in Large-Grained Yttrium-Doped Barium Zirconate. *Chemistry of Materials* **2009**, *21* (13), 2755-2762. DOI: 10.1021/cm900208w.

- (76) Han, D.; Hatada, N.; Uda, T. Chemical Expansion of Yttrium-Doped Barium Zirconate and Correlation with Proton Concentration and Conductivity. *Journal of the American Ceramic Society* **2016**, *99* (11), 3745-3753. DOI: <https://doi.org/10.1111/jace.14377>.
- (77) Morrow, J. D.; Gardner, J. L. A.; Deringer, V. L. How to validate machine-learned interatomic potentials. *The Journal of Chemical Physics* **2023**, *158* (12). DOI: 10.1063/5.0139611 (accessed 4/7/2025).
- (78) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review* **1965**, *140* (4A), A1133-A1138. DOI: 10.1103/PhysRev.140.A1133.
- (79) Shan, N.; Mingxia, Z.; K., H. M.; Josephine, C.; and Liu, B. Practical principles of density functional theory for catalytic reaction simulations on metal surfaces – from theory to applications. *Mol Simulat* **2017**, *43* (10-11), 861-885. DOI: 10.1080/08927022.2017.1303687.
- (80) Kittel, C. *Introduction to Solid State Physics, 7th Edition*; New York (NY): Wiley, 1996.
- (81) Blöchl, P. E. Projector augmented-wave method. *Physical review B* **1994**, *50* (24), 17953.
- (82) Perdew, J. P.; Schmidt, K. Jacob's ladder of density functional approximations for the exchange-correlation energy. *AIP Conference Proceedings* **2001**, *577* (1), 1-20. DOI: 10.1063/1.1390175 (accessed 4/17/2025).
- (83) Car, R. Density functional theory: Fixing Jacob's ladder. *Nat Chem* **2016**, *8* (9), 820-821. DOI: 10.1038/nchem.2605 From NLM.
- (84) Perdew, J. P.; Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Physical Review B* **1981**, *23* (10), 5048-5079. DOI: 10.1103/PhysRevB.23.5048.
- (85) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical Review Letters* **1996**, *77* (18), 3865-3868. DOI: 10.1103/PhysRevLett.77.3865.

- (86) Perdew, J. P.; Wang, Y. Accurate and simple analytic representation of the electron-gas correlation energy. *Physical Review B* **1992**, *45* (23), 13244-13249. DOI: 10.1103/PhysRevB.45.13244.
- (87) Hammer, B.; Hansen, L. B.; Nørskov, J. K. Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Physical Review B* **1999**, *59* (11), 7413-7421. DOI: 10.1103/PhysRevB.59.7413.
- (88) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E.; Constantin, L. A.; Zhou, X.; Burke, K. Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces. *Physical Review Letters* **2008**, *100* (13), 136406. DOI: 10.1103/PhysRevLett.100.136406.
- (89) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Physical Review Letters* **2015**, *115* (3), 036402. DOI: 10.1103/PhysRevLett.115.036402.
- (90) Furness, J. W.; Kaplan, A. D.; Ning, J.; Perdew, J. P.; Sun, J. Accurate and Numerically Efficient r2SCAN Meta-Generalized Gradient Approximation. *The Journal of Physical Chemistry Letters* **2020**, *11* (19), 8208-8215. DOI: 10.1021/acs.jpclett.0c02405.
- (91) Kothakonda, M.; Kaplan, A. D.; Isaacs, E. B.; Bartel, C. J.; Furness, J. W.; Ning, J.; Wolverton, C.; Perdew, J. P.; Sun, J. Testing the r2SCAN Density Functional for the Thermodynamic Stability of Solids with and without a van der Waals Correction. *ACS Materials Au* **2023**, *3* (2), 102-111. DOI: 10.1021/acsmaterialsau.2c00059.
- (92) Becke, A. D. Density - functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **1993**, *98* (7), 5648-5652. DOI: 10.1063/1.464913 (accessed 4/21/2025).

- (93) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B* **1988**, *37* (2), 785-789. DOI: 10.1103/PhysRevB.37.785.
- (94) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Hybrid functionals based on a screened Coulomb potential. *The Journal of Chemical Physics* **2003**, *118* (18), 8207-8215. DOI: 10.1063/1.1564060 (accessed 4/21/2025).
- (95) Adamo, C.; Barone, V. Toward chemical accuracy in the computation of NMR shieldings: the PBE0 model. *Chemical Physics Letters* **1998**, *298* (1), 113-119. DOI: [https://doi.org/10.1016/S0009-2614\(98\)01201-9](https://doi.org/10.1016/S0009-2614(98)01201-9).
- (96) Anisimov, V. I.; Zaanen, J.; Andersen, O. K. Band theory and Mott insulators: Hubbard U instead of Stoner I. *Physical Review B* **1991**, *44* (3), 943-954. DOI: 10.1103/PhysRevB.44.943.
- (97) Wang, L.; Maxisch, T.; Ceder, G. Oxidation energies of transition metal oxides within the $\text{GGA}+\text{U}$ framework. *Physical Review B* **2006**, *73* (19), 195107. DOI: 10.1103/PhysRevB.73.195107.
- (98) Behler, J.; Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical review letters* **2007**, *98* (14), 146401.
- (99) Behler, J. Four Generations of High-Dimensional Neural Network Potentials. *Chemical Reviews* **2021**, *121* (16), 10037-10072. DOI: 10.1021/acs.chemrev.0c00868.
- (100) Behler, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *The Journal of chemical physics* **2011**, *134* (7), 074106.
- (101) Drautz, R. Atomic cluster expansion for accurate and transferable interatomic potentials. *Physical Review B* **2019**, *99* (1), 014104. DOI: 10.1103/PhysRevB.99.014104.

- (102) Wang, M.; Ding, J.; Xiong, G.; Zhan, N.; Owen, C. J.; Musaelian, A.; Xie, Y.; Molinari, N.; Adams, R. P.; Haile, S. Revealing the proton slingshot mechanism in solid acid electrolytes through machine learning molecular dynamics. *arXiv preprint arXiv:2503.15389* **2025**.
- (103) Yang, J. H.; Ooi, A. W. S.; Goodwin, Z. A. H.; Xie, Y.; Ding, J.; Falletta, S.; Park, A.-H. A.; Kozinsky, B. Room-Temperature Decomposition of the Ethaline Deep Eutectic Solvent. *The Journal of Physical Chemistry Letters* **2025**, *16* (12), 3039-3046. DOI: 10.1021/acs.jpcclett.4c03645.
- (104) Ab Initio Molecular Dynamics. In *Density Functional Theory*, 2009; pp 193-208.
- (105) Pathria, R. K.; Beale, P. D. 3 - The Canonical Ensemble. In *Statistical Mechanics (Third Edition)*, Pathria, R. K., Beale, P. D. Eds.; Academic Press, 2011; pp 39-90.
- (106) Parrinello, M.; Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Physical Review Letters* **1980**, *45* (14), 1196-1199. DOI: 10.1103/PhysRevLett.45.1196.
- (107) Togo, A. First-principles Phonon Calculations with Phonopy and Phono3py. *Journal of the Physical Society of Japan* **2022**, *92* (1), 012001. DOI: 10.7566/JPSJ.92.012001 (accessed 2024/11/07).
- (108) Sheng, H. W.; Kramer, M. J.; Cadien, A.; Fujita, T.; Chen, M. W. Highly optimized embedded-atom-method potentials for fourteen fcc metals. *Physical Review B* **2011**, *83* (13). DOI: 10.1103/PhysRevB.83.134118.
- (109) Mishin, Y.; Farkas, D.; Mehl, M. J.; Papaconstantopoulos, D. A. Interatomic potentials for monoatomic metals from experimental data and ab initio calculations. *Physical Review B* **1999**, *59* (5), 3393-3407. DOI: 10.1103/PhysRevB.59.3393.
- (110) Adams, J. B.; Foiles, S. M.; Wolfer, W. G. SELF-DIFFUSION AND IMPURITY DIFFUSION OF FCC METALS USING THE 5-FREQUENCY MODEL AND THE

EMBEDDED ATOM METHOD. *Journal of Materials Research* **1989**, 4 (1), 102-112. DOI: 10.1557/jmr.1989.0102.

(111) Asadi, E.; Zaeem, M. A.; Nouranian, S.; Baskes, M. I. Two-phase solid-liquid coexistence of Ni, Cu, and Al by molecular dynamics simulations using the modified embedded-atom method. *Acta Materialia* **2015**, 86, 169-181. DOI: 10.1016/j.actamat.2014.12.010.

(112) Etesami, S. A.; Asadi, E. Molecular dynamics for near melting temperatures simulations of metals using modified embedded-atom method. *Journal of Physics and Chemistry of Solids* **2018**, 112, 61-72. DOI: 10.1016/j.jpcs.2017.09.001.

(113) Lee, B.-J.; Shim, J.-H.; Baskes, M. I. Semiempirical atomic potentials for the fcc metals Cu, Ag, Au, Ni, Pd, Pt, Al, and Pb based on first and second nearest-neighbor modified embedded atom method. *Physical Review B* **2003**, 68 (14), 144112. DOI: 10.1103/PhysRevB.68.144112.

(114) Nenuwe, N.; Agbalagba, E.; Enaibe, E. Comparative study of thermophysical properties of Nickel in liquid phase. *Ruhuna Journal of Science* **2018**, 9 (2), 117-126.

(115) Jacobsen, K. W.; Stoltze, P.; Norskov, J. K. A semi-empirical effective medium theory for metals and alloys. *Surface Science* **1996**, 366 (2), 394-402. DOI: 10.1016/0039-6028(96)00816-3.

(116) Shin, Y. K.; Gao, Y. W.; Shin, D.; van Duin, A. C. T. Impact of three-body interactions in a ReaxFF force field for Ni and Cr transition metals and their alloys on the prediction of thermal and mechanical properties. *Computational Materials Science* **2021**, 197. DOI: 10.1016/j.commatsci.2021.110602.

(117) Zou, C.; Shin, Y. K.; van Duin, A. C. T.; Fang, H.; Liu, Z.-K. Molecular dynamics simulations of the effects of vacancies on nickel self-diffusion, oxygen diffusion and oxidation initiation in nickel, using the ReaxFF reactive force field. *Acta Materialia* **2015**, 83, 102-112. DOI: <https://doi.org/10.1016/j.actamat.2014.09.047>.

- (118) Zuo, Y.; Chen, C.; Li, X.; Deng, Z.; Chen, Y.; Behler, J. r.; Csányi, G.; Shapeev, A. V.; Thompson, A. P.; Wood, M. A. Performance and cost assessment of machine learning interatomic potentials. *The Journal of Physical Chemistry A* **2020**, *124* (4), 731-745.
- (119) Wang, H.; Zhang, L.; Han, J.; Weinan, E. DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Computer Physics Communications* **2018**, *228*, 178-184.
- (120) Jiequn Han, L. Z., Roberto Car & Weinan E. Deep Potential: A General Representation of a Many-Body Potential Energy Surface. *Communications in Computational Physics* **2020**, *23*, 629-639. DOI: 10.4208/cicp.OA-2017-0213.
- (121) Zhang, L.; Han, J.; Wang, H.; Saidi, W.; Car, R. End-to-end symmetry preserving interatomic potential energy model for finite and extended systems. *Advances in Neural Information Processing Systems* **2018**, *31*.
- (122) Wang, J. J.; Shen, H.; Yang, R. Y.; Xie, K.; Zhang, C.; Chen, L. Y.; Ho, K. M.; Wang, C. Z.; Wang, S. Y. A deep learning interatomic potential developed for atomistic simulation of carbon materials. *Carbon* **2022**, *186*, 1-8. DOI: 10.1016/j.carbon.2021.09.062.
- (123) Balyakin, I. A.; Sadovnikov, S. I. Deep learning potential for superionic phase of Ag₂S. *Computational Materials Science* **2022**, *202*. DOI: 10.1016/j.commatsci.2021.110963.
- (124) Okita, T.; Terayama, S.; Tsugawa, K.; Kobayashi, K.; Okumura, M.; Itakura, M.; Suzuki, K. Construction of machine-learning Zr interatomic potentials for identifying the formation process of c-type dislocation loops. *Computational Materials Science* **2022**, *202*. DOI: 10.1016/j.commatsci.2021.110865.

- (125) Tang, L.; Ho, K. M.; Wang, C. Z. Molecular dynamics simulation of metallic Al-Ce liquids using a neural network machine learning interatomic potential. *Journal of Chemical Physics* **2021**, *155* (19). DOI: 10.1063/5.0066061.
- (126) Andolina, C. M.; Bon, M.; Passerone, D.; Saidi, W. A. Robust, Multi-Length-Scale, Machine Learning Potential for Ag-Au Bimetallic Alloys from Clusters to Bulk Materials. *Journal of Physical Chemistry C* **2021**, *125* (31), 17438-17447. DOI: 10.1021/acs.jpcc.1c04403.
- (127) Shi, Y.; Doyle, C. C.; Beck, T. L. Condensed Phase Water Molecular Multipole Moments from Deep Neural Network Models Trained on Ab Initio Simulation Data. *Journal of Physical Chemistry Letters* **2021**, *12* (42), 10310-10317. DOI: 10.1021/acs.jpclett.1c02328.
- (128) Cheng, B. Q.; Engel, E. A.; Behler, J.; Dellago, C.; Ceriotti, M. Ab initio thermodynamics of liquid and solid water. *Proceedings of the National Academy of Sciences of the United States of America* **2019**, *116* (4), 1110-1115. DOI: 10.1073/pnas.1815117116.
- (129) Gerrits, N. Accurate Simulations of the Reaction of H₂ on a Curved Pt Crystal through Machine Learning. *Journal of Physical Chemistry Letters* **2021**, *12* (51), 12157-12164. DOI: 10.1021/acs.jpclett.1c03395.
- (130) Xie, K.; Qiao, C.; Shen, H.; Yang, R. Y.; Xu, M.; Zhang, C.; Zheng, Y. X.; Zhang, R. J.; Chen, L. Y.; Ho, K. M.; et al. Neural network potential for Zr-Rh system by machine learning. *Journal of Physics-Condensed Matter* **2022**, *34* (7). DOI: 10.1088/1361-648X/ac37dc.
- (131) Groenenboom, M. C.; Anderson, R. M.; Wollmershauser, J. A.; Horton, D. J.; Policastro, S. A.; Keith, J. A. Combined Neural Network Potential and Density Functional Theory Study of TiAl₂O₅ Surface Morphology and Oxygen Reduction Reaction Overpotentials. *Journal of Physical Chemistry C* **2020**, *124* (28), 15171-15179. DOI: 10.1021/acs.jpcc.0c02093.

- (132) Niu, H. Y.; Bonati, L.; Piaggi, P. M.; Parrinello, M. Ab initio phase diagram and nucleation of gallium. *Nature Communications* **2020**, *11* (1). DOI: 10.1038/s41467-020-16372-9.
- (133) Balyakin, I. A.; Rempel, S. V.; Ryltsev, R. E.; Rempel, A. A. Deep machine learning interatomic potential for liquid silica. *Physical Review E* **2020**, *102* (5). DOI: 10.1103/PhysRevE.102.052125.
- (134) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Physical review letters* **1996**, *77* (18), 3865.
- (135) Gassner, H.; Probst, M.; Lauenstein, A.; Hermansson, K. Representation of Intermolecular Potential Functions by Neural Networks. *The Journal of Physical Chemistry A* **1998**, *102* (24), 4596-4605. DOI: 10.1021/jp972209d.
- (136) Blank, T. B.; Brown, S. D. Adaptive, global, extended Kalman filters for training feedforward neural networks. *Journal of Chemometrics* **1994**, *8* (6), 391-407. DOI: <https://doi.org/10.1002/cem.1180080605>.
- (137) Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics* **1995**, *117* (1), 1-19.
- (138) Singraber, A.; Behler, J. r.; Dellago, C. Library-based LAMMPS implementation of high-dimensional neural network potentials. *Journal of chemical theory and computation* **2019**, *15* (3), 1827-1840.
- (139) Stukowski, A. Visualization and analysis of atomistic simulation data with OVITO—the Open Visualization Tool. *Modelling and simulation in materials science and engineering* **2009**, *18* (1), 015012.
- (140) Zhang, W. J.; Peng, Y. F.; Liu, Z. L. Molecular dynamics simulations of the melting curve of NiAl alloy under pressure. *Aip Advances* **2014**, *4* (5). DOI: 10.1063/1.4876515.

- (141) Lide, D. R. *CRC handbook of chemistry and physics*; CRC press, 2004.
- (142) Y.S. Touloukian, R. K. K., R.E. Taylor, P.D. Desai. *Thermal Expansion, Metallic Elements and Alloys*; Plenum Press, 1975.
- (143) Shang, S. L.; Wang, Y.; Kim, D.; Liu, Z. K. First-principles thermodynamics from phonon and Debye model: Application to Ni and Ni₃Al. *Computational Materials Science* **2010**, 47 (4), 1040-1048. DOI: 10.1016/j.commatsci.2009.12.006.
- (144) Y.S. Touloukian, C. Y. H. *Properties of Selected Ferrous Alloying Elements*; Hemisphere Publishing Corporation, 1989.
- (145) Meyer, A.; Stuber, S.; Holland-Moritz, D.; Heinen, O.; Unruh, T. Determination of self-diffusion coefficients by quasielastic neutron scattering measurements of levitated Ni droplets. *Physical Review B* **2008**, 77 (9). DOI: 10.1103/PhysRevB.77.092201.
- (146) Meyer, A. The measurement of self-diffusion coefficients in liquid metals with quasielastic neutron scattering. *EPJ Web of Conferences* **2015**, 83, 01002.
- (147) Lu, Y. J.; Cheng, H.; Chen, M. A molecular dynamics examination of the relationship between self-diffusion and viscosity in liquid metals. *Journal of Chemical Physics* **2012**, 136 (21). DOI: 10.1063/1.4723683.
- (148) Weintraub, E. Preparation and properties of pure boron. *Trans. Amer. Electrochem. Soc.* **1909**, 16, 165-184.
- (149) Kobayashi, M.; Higashi, I.; Brodhag, C.; Thevenot, F. Structure of B₆O boron-suboxide by Rietveld refinement. *Journal of materials science* **1993**, 28, 2129-2134.
- (150) Hubert, H.; Devouard, B.; Garvie, L. A. J.; O'Keeffe, M.; Buseck, P. R.; Petuskey, W. T.; McMillan, P. F. Icosahedral packing of B₁₂ icosahedra in boron suboxide (B₆O). *Nature* **1998**, 391 (6665), 376-378. DOI: 10.1038/34885.

- (151) Hubert, H.; Garvie, L. A. J.; Devouard, B.; Buseck, P. R.; Petuskey, W. T.; McMillan, P. F. High-Pressure, High-Temperature Synthesis and Characterization of Boron Suboxide (B₆O). *Chemistry of Materials* **1998**, *10* (6), 1530-1537. DOI: 10.1021/cm970433+.
- (152) Emin, D. Unusual properties of icosahedral boron-rich solids. *Journal of Solid State Chemistry* **2006**, *179* (9), 2791-2798.
- (153) Huber, S. P.; Gullikson, E.; Frye, C. D.; Edgar, J. H.; van de Kruijs, R. W. E.; Bijkerk, F.; Prendergast, D. Self-healing in B₁₂P₂ through Mediated Defect Recombination. *Chemistry of Materials* **2016**, *28* (22), 8415-8428. DOI: 10.1021/acs.chemmater.6b04075.
- (154) White, M. A.; Cerqueira, A. B.; Whitman, C. A.; Johnson, M. B.; Ogitsu, T. Determination of phase stability of elemental boron. *Angewandte Chemie* **2015**, *127* (12), 3697-3700.
- (155) Shang, S.; Wang, Y.; Arroyave, R.; Liu, Z.-K. Phase stability in α - and β -rhombohedral boron. *Physical Review B* **2007**, *75* (9), 092101. DOI: 10.1103/PhysRevB.75.092101.
- (156) van Setten, M. J.; Uijttewaalt, M. A.; de Wijs, G. A.; de Groot, R. A. Thermodynamic Stability of Boron: The Role of Defects and Zero Point Motion. *Journal of the American Chemical Society* **2007**, *129* (9), 2458-2465. DOI: 10.1021/ja0631246.
- (157) Michael, W.; Marek, M. Relative stability of α and β boron. *Journal of Physics: Conference Series* **2009**, *176* (1), 012024. DOI: 10.1088/1742-6596/176/1/012024.
- (158) Ogitsu, T.; Gygi, F.; Reed, J.; Udagawa, M.; Motome, Y.; Schwegler, E.; Galli, G. Geometrical frustration in an elemental solid: An Ising model to explain the defect structure of β -rhombohedral boron. *Physical Review B* **2010**, *81* (2), 020102. DOI: 10.1103/PhysRevB.81.020102.
- (159) Ogitsu, T.; Schwegler, E. The α - β phase boundary of elemental boron. *Solid State Sciences* **2012**, *14* (11), 1598-1600. DOI: <https://doi.org/10.1016/j.solidstatesciences.2012.03.005>.

- (160) Podryabinkin, E. V.; Tikhonov, E. V.; Shapeev, A. V.; Oganov, A. R. Accelerating crystal structure prediction by machine-learning interatomic potentials with active learning. *Physical Review B* **2019**, *99* (6), 064114. DOI: 10.1103/PhysRevB.99.064114.
- (161) Huang, S.-D.; Shang, C.; Kang, P.-L.; Liu, Z.-P. Atomic structure of boron resolved using machine learning and global sampling. *Chemical Science* **2018**, *9* (46), 8644-8655, 10.1039/C8SC03427C. DOI: 10.1039/C8SC03427C.
- (162) Deringer, V. L.; Pickard, C. J.; Csányi, G. Data-Driven Learning of Total and Local Energies in Elemental Boron. *Physical Review Letters* **2018**, *120* (15), 156001. DOI: 10.1103/PhysRevLett.120.156001.
- (163) Yang, Q.; Lv, J.; Tong, Q.; Du, X.; Wang, Y.; Zhang, S.; Yang, G.; Bergara, A.; Ma, Y. Hard and superconducting cubic boron phase via swarm-intelligence structural prediction driven by a machine-learning potential. *Physical Review B* **2021**, *103* (2), 024505. DOI: 10.1103/PhysRevB.103.024505.
- (164) Chen, X.; Shao, W.; Le, N. Q.; Clancy, P. Transferable Force Field for Gallium Nitride Crystal Growth from the Melt Using On-The-Fly Active Learning. *Journal of Chemical Theory and Computation* **2023**. DOI: 10.1021/acs.jctc.3c00587.
- (165) Momma, K.; Izumi, F. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. *Journal of applied crystallography* **2011**, *44* (6), 1272-1276.
- (166) Birch, F. Finite Elastic Strain of Cubic Crystals. *Physical Review* **1947**, *71* (11), 809-824. DOI: 10.1103/PhysRev.71.809.
- (167) Cherednichenko, K. A.; Solozhenko, V. L. Thermal expansion of α -boron and some boron-rich pnictides. *Solid State Communications* **2019**, *303-304*, 113735. DOI: <https://doi.org/10.1016/j.ssc.2019.113735>.

- (168) Lundström, T.; Lönnberg, B.; Bauer, J. Thermal expansion of β -rhombohedral boron. *Journal of Alloys and Compounds* **1998**, 267 (1), 54-58. DOI: [https://doi.org/10.1016/S0925-8388\(97\)00545-8](https://doi.org/10.1016/S0925-8388(97)00545-8).
- (169) Togo, A.; Tanaka, I. First principles phonon calculations in materials science. *Scripta Materialia* **2015**, 108, 1-5. DOI: <https://doi.org/10.1016/j.scriptamat.2015.07.021>.
- (170) Kong, L.; Lewis, L. J. Surface diffusion coefficients: Substrate dynamics matters. *Physical Review B* **2008**, 77 (16), 165422. DOI: 10.1103/PhysRevB.77.165422.
- (171) Kong, L. T. Phonon dispersion measured directly from molecular dynamics simulations. *Computer Physics Communications* **2011**, 182 (10), 2201-2207. DOI: <https://doi.org/10.1016/j.cpc.2011.04.019>.
- (172) Cheng, B.; Griffiths, R.-R.; Wengert, S.; Kunkel, C.; Stenczel, T.; Zhu, B.; Deringer, V. L.; Bernstein, N.; Margraf, J. T.; Reuter, K.; et al. Mapping Materials and Molecules. *Accounts of Chemical Research* **2020**, 53 (9), 1981-1991. DOI: 10.1021/acs.accounts.0c00403.
- (173) Shirai, K. Phase diagram of boron crystals. *Japanese Journal of Applied Physics* **2017**, 56 (5S3), 05FA06. DOI: 10.7567/JJAP.56.05FA06.
- (174) Duan, C.; Tong, J.; Shang, M.; Nikodemski, S.; Sanders, M.; Ricote, S.; Almansoori, A.; O'Hayre, R. Readily processed protonic ceramic fuel cells with high performance at low temperatures. *Science* **2015**, 349 (6254), 1321-1326. DOI: doi:10.1126/science.aab3987.
- (175) Liu, F.; Deng, H.; Diercks, D.; Kumar, P.; Jabbar, M. H. A.; Gumeci, C.; Furuya, Y.; Dale, N.; Oku, T.; Usuda, M.; et al. Lowering the operating temperature of protonic ceramic electrochemical cells to <450 °C. *Nature Energy* **2023**, 8 (10), 1145-1157. DOI: 10.1038/s41560-023-01350-4.

- (176) Duan, C.; Hook, D.; Chen, Y.; Tong, J.; O'Hayre, R. Zr and Y co-doped perovskite as a stable, high performance cathode for solid oxide fuel cells operating below 500 °C. *Energy & Environmental Science* **2017**, *10* (1), 176-182, 10.1039/C6EE01915C. DOI: 10.1039/C6EE01915C.
- (177) Liu, F.; Deng, H.; Ding, H.; Kazempoor, P.; Liu, B.; Duan, C. Process-intensified protonic ceramic fuel cells for power generation, chemical production, and greenhouse gas mitigation. *Joule* **2023**, *7* (6), 1308-1332. DOI: <https://doi.org/10.1016/j.joule.2023.05.009>.
- (178) Meng, Y.; Deng, H.; Wang, L.-C.; Kim, D.; Liu, B.; Liu, F.; Ding, Y.; Ding, D. Highly efficient La/Ni co-doped strontium titanate catalyst for co-production of propylene and hydrogen from propane in protonic ceramic electrochemical cells. *Applied Catalysis B: Environment and Energy* **2024**, *354*, 124111. DOI: <https://doi.org/10.1016/j.apcatb.2024.124111>.
- (179) Zheng, Y.; Wang, J.; Yu, B.; Zhang, W.; Chen, J.; Qiao, J.; Zhang, J. A review of high temperature co-electrolysis of H₂O and CO₂ to produce sustainable fuels using solid oxide electrolysis cells (SOECs): advanced materials and technology. *Chemical Society Reviews* **2017**, *46* (5), 1427-1463, 10.1039/C6CS00403B. DOI: 10.1039/C6CS00403B.
- (180) Bian, W.; Wu, W.; Wang, B.; Tang, W.; Zhou, M.; Jin, C.; Ding, H.; Fan, W.; Dong, Y.; Li, J.; et al. Revitalizing interface in protonic ceramic cells by acid etch. *Nature* **2022**, *604* (7906), 479-485. DOI: 10.1038/s41586-022-04457-y.
- (181) Ding, D.; Zhang, Y.; Wu, W.; Chen, D.; Liu, M.; He, T. A novel low-thermal-budget approach for the co-production of ethylene and hydrogen via the electrochemical non-oxidative deprotonation of ethane. *Energy & Environmental Science* **2018**, *11* (7), 1710-1716, 10.1039/C8EE00645H. DOI: 10.1039/C8EE00645H.

- (182) Wu, W.; Wang, L.-C.; Hu, H.; Bian, W.; Gomez, J. Y.; Orme, C. J.; Ding, H.; Dong, Y.; He, T.; Li, J.; et al. Electrochemically Engineered, Highly Energy-Efficient Conversion of Ethane to Ethylene and Hydrogen below 550 °C in a Protonic Ceramic Electrochemical Cell. *ACS Catalysis* **2021**, *11* (19), 12194-12202. DOI: 10.1021/acscatal.1c03351.
- (183) Li, M.; Hua, B.; Wu, W.; Wang, L.-C.; Ding, Y.; Welander, M. M.; Walker, R. A.; Ding, D. Activating nano-bulk interplays for sustainable ammonia electrosynthesis. *Materials Today* **2022**, *60*, 31-40. DOI: 10.1016/j.mattod.2022.09.011.
- (184) Li, M.; Hua, B.; Wang, L.-C.; Sugar, J. D.; Wu, W.; Ding, Y.; Li, J.; Ding, D. Switching of metal–oxygen hybridization for selective CO₂ electrohydrogenation under mild temperature and pressure. *Nature Catalysis* **2021**, *4* (4), 274-283. DOI: 10.1038/s41929-021-00590-5.
- (185) Wang, Y.; Ling, Y.; Wang, B.; Zhai, G.; Yang, G.; Shao, Z.; Xiao, R.; Li, T. A review of progress in proton ceramic electrochemical cells: material and structural design, coupled with value-added chemical production. *Energy & Environmental Science* **2023**, *16* (12), 5721-5770.
- (186) Su, H.; Hu, Y. H. Degradation issues and stabilization strategies of protonic ceramic electrolysis cells for steam electrolysis. *Energy Science & Engineering* **2022**, *10* (5), 1706-1725. DOI: <https://doi.org/10.1002/ese3.1010>.
- (187) Guo, S.; Jiang, L.; Li, Y.; Zhong, P.; Ismail, S. A.; Norby, T.; Han, D. From Electrolyte and Electrode Materials to Large-Area Protonic Ceramic Fuel Cells: A Review. *Advanced Functional Materials* **2024**, *34* (32), 2304729. DOI: <https://doi.org/10.1002/adfm.202304729>.
- (188) Zhou, Y.; Zhang, W.; Kane, N.; Luo, Z.; Pei, K.; Sasaki, K.; Choi, Y.; Chen, Y.; Ding, D.; Liu, M. An Efficient Bifunctional Air Electrode for Reversible Protonic Ceramic Electrochemical Cells. *Advanced Functional Materials* **2021**, *31* (40), 2105386. DOI: <https://doi.org/10.1002/adfm.202105386>.

- (189) Niu, Y.; Zhou, Y.; Zhang, W.; Zhang, Y.; Evans, C.; Luo, Z.; Kane, N.; Ding, Y.; Chen, Y.; Guo, X.; et al. Highly Active and Durable Air Electrodes for Reversible Protonic Ceramic Electrochemical Cells Enabled by an Efficient Bifunctional Catalyst. *Advanced Energy Materials* **2022**, *12* (12), 2103783. DOI: <https://doi.org/10.1002/aenm.202103783>.
- (190) Liu, Z.; Bai, Y.; Sun, H.; Guan, D.; Li, W.; Huang, W.-H.; Pao, C.-W.; Hu, Z.; Yang, G.; Zhu, Y.; et al. Synergistic dual-phase air electrode enables high and durable performance of reversible proton ceramic electrochemical cells. *Nature Communications* **2024**, *15* (1), 472. DOI: 10.1038/s41467-024-44767-5.
- (191) Pei, K.; Zhou, Y.; Xu, K.; Zhang, H.; Ding, Y.; Zhao, B.; Yuan, W.; Sasaki, K.; Choi, Y.; Chen, Y.; et al. Surface restructuring of a perovskite-type air electrode for reversible protonic ceramic electrochemical cells. *Nature Communications* **2022**, *13* (1), 2207. DOI: 10.1038/s41467-022-29866-5.
- (192) Yuan, B.; Wang, N.; Tang, C.; Meng, L.; Du, L.; Su, Q.; Aoki, Y.; Ye, S. Advances and challenges in high-performance cathodes for protonic solid oxide fuel cells and machine learning-guided perspectives. *Nano Energy* **2024**, *122*, 109306. DOI: <https://doi.org/10.1016/j.nanoen.2024.109306>.
- (193) Wang, N.; Tang, C.; Du, L.; Zhu, R.; Xing, L.; Song, Z.; Yuan, B.; Zhao, L.; Aoki, Y.; Ye, S. Advanced Cathode Materials for Protonic Ceramic Fuel Cells: Recent Progress and Future Perspectives. *Advanced Energy Materials* **2022**, *12* (34), 2201882. DOI: <https://doi.org/10.1002/aenm.202201882>.
- (194) Feng, W.; Wu, W.; Zhao, Z.; Gomez, J. Y.; Orme, C. J.; Tang, W.; Bian, W.; Priest, C.; Stewart, F. F.; Jin, C.; et al. Mathematical Model-Assisted Ultrasonic Spray Coating for Scalable

Production of Large-Sized Solid Oxide Electrochemical Cells. *ACS Applied Materials & Interfaces* **2023**, *15* (26), 31430-31437. DOI: 10.1021/acsami.3c04208.

(195) Li, C.; Tong, X.; Yuan, C.; Tong, Y.; Zhang, Y.; Wang, N.; Li, P.; Pang, S.; Wang, L.; Zhan, Z. Electrolyte-electrode interface: A key factor for advanced protonic ceramic electrochemical cells. *Ceramics International* **2024**, *50* (3, Part A), 4656-4664. DOI: <https://doi.org/10.1016/j.ceramint.2023.11.209>.

(196) Wu, W.; Wang, C.; Bian, W.; Hua, B.; Gomez, J. Y.; Orme, C. J.; Tang, W.; Stewart, F. F.; Ding, D. Root Cause Analysis of Degradation in Protonic Ceramic Electrochemical Cell with Interfacial Electrical Sensors Using Data-Driven Machine Learning. *Adv Sci (Weinh)* **2023**, *10* (30), e2304074. DOI: 10.1002/advs.202304074 From NLM.

(197) Nomura, K.; Shimada, H.; Yamaguchi, Y.; Shin, W.; Okuyama, Y.; Mizutani, Y. Phase Transitions, Thermal Expansions, Chemical Expansions, and CO₂ Resistances of Ba(Ce_{0.8-x}Zr_xY_{0.1}Yb_{0.1})O_{3-δ} (x = 0.1, 0.4) Perovskite-Type Proton Conductors. *Journal of The Electrochemical Society* **2022**, *169* (2), 024516. DOI: 10.1149/1945-7111/ac5480.

(198) Kreuer, K. D. Proton-Conducting Oxides. *Annual Review of Materials Research* **2003**, *33* (Volume 33, 2003), 333-359. DOI: <https://doi.org/10.1146/annurev.matsci.33.022802.091825>.

(199) Løken, A.; Haugrud, R.; Bjørheim, T. S. Unravelling the fundamentals of thermal and chemical expansion of BaCeO₃ from first principles phonon calculations. *Physical Chemistry Chemical Physics* **2016**, *18* (45), 31296-31303, 10.1039/C6CP05710A. DOI: 10.1039/C6CP05710A.

(200) Zheng, H.; Sata, N.; Han, F.; Dayaghi, A. M.; Iguchi, F.; Develos-Bagarinao, K.; Norby, T.; Stange, M.; Costa, R. Hydrogen Production with a Protonic Ceramic Membrane Reactor on Porous Fe–Cr Alloy. *ACS Energy Letters* **2024**, *9* (8), 3962-3969. DOI: 10.1021/acsenergylett.4c01173.

- (201) Anderson, L. O.; Zhang, Q.; Perry, N. H. Tuning Perovskites' Hydration-Induced Chemical Expansion with Octahedral Tilt Angles. *Chemistry of Materials* **2024**, *36* (12), 5953-5964. DOI: 10.1021/acs.chemmater.4c00354.
- (202) Bjørheim, T. S.; Kotomin, E. A.; Maier, J. Hydration entropy of BaZrO₃ from first principles phonon calculations. *Journal of Materials Chemistry A* **2015**, *3* (14), 7639-7648, 10.1039/C4TA06880G. DOI: 10.1039/C4TA06880G.
- (203) Bjørheim, T. S.; Løken, A.; Haugsrud, R. On the relationship between chemical expansion and hydration thermodynamics of proton conducting perovskites. *Journal of Materials Chemistry A* **2016**, *4* (16), 5917-5924, 10.1039/C5TA10090A. DOI: 10.1039/C5TA10090A.
- (204) Cusentino, M. A.; Wood, M. A.; Thompson, A. P. Machine learned interatomic potentials for gas-metal interactions. *Modelling and Simulation in Materials Science and Engineering* **2025**, *33* (1), 015007. DOI: 10.1088/1361-651X/ad93ec.
- (205) Kostiuchenko, T.; Körmann, F.; Neugebauer, J.; Shapeev, A. Impact of lattice relaxations on phase transitions in a high-entropy alloy studied by machine-learning potentials. *npj Computational Materials* **2019**, *5* (1), 55. DOI: 10.1038/s41524-019-0195-y.
- (206) Deng, H.; Liu, B. Predictions of Boron Phase Stability Using an Efficient Bayesian Machine Learning Interatomic Potential. *The Journal of Physical Chemistry Letters* **2024**, *15* (9), 2419-2427. DOI: 10.1021/acs.jpcclett.4c00322.
- (207) Ouyang, Y.; Yu, C.; He, J.; Jiang, P.; Ren, W.; Chen, J. Accurate description of high-order phonon anharmonicity and lattice thermal conductivity from molecular dynamics simulations with machine learning potential. *Physical Review B* **2022**, *105* (11), 115202. DOI: 10.1103/PhysRevB.105.115202.

- (208) Yulou, O.; Zhongwei, Z.; Cuiqian, Y.; Jia, H.; Gang, Y.; Jie, C. Accuracy of Machine Learning Potential for Predictions of Multiple-Target Physical Properties. *Chin. Phys. Lett.* **2020**, *37* (12), 126301-126301. DOI: 10.1088/0256-307X/37/12/126301.
- (209) He, J.; Yu, C.; Lu, S.; Shan, S.; Zhang, Z.; Chen, J. Complex role of strain engineering of lattice thermal conductivity in hydrogenated graphene-like borophene induced by high-order phonon anharmonicity. *Nanotechnology* **2024**, *35* (2), 025703. DOI: 10.1088/1361-6528/ad0127.
- (210) Kozinsky, B.; Musaelian, A.; Johansson, A.; Batzner, S. Scaling the Leading Accuracy of Deep Equivariant Models to Biomolecular Simulations of Realistic Size. In Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, Denver, CO, USA; 2023.
- (211) Tan, C. W.; Descoteaux, M. L.; Kotak, M.; Nascimento, G. d. M.; Kavanagh, S. R.; Zichi, L.; Wang, M.; Saluja, A.; Hu, Y. R.; Smidt, T. High-performance training and inference for deep equivariant interatomic potentials. *arXiv preprint arXiv:2504.16068* **2025**.
- (212) Ning, J.; Kothakonda, M.; Furness, J. W.; Kaplan, A. D.; Ehlert, S.; Brandenburg, J. G.; Perdew, J. P.; Sun, J. Workhorse minimally empirical dispersion-corrected density functional with tests for weakly bound systems: r2SCAN+rVV10. *Physical Review B* **2022**, *106* (7), 075422. DOI: 10.1103/PhysRevB.106.075422.
- (213) Monkhorst, H. J.; Pack, J. D. Special points for Brillouin-zone integrations. *Physical Review B* **1976**, *13*, 5188-5192.
- (214) Dudarev, S. L.; Botton, G. A.; Savrasov, S. Y.; Humphreys, C. J.; Sutton, A. P. Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study. *Physical Review B* **1998**, *57* (3), 1505-1509. DOI: 10.1103/PhysRevB.57.1505.

- (215) Togo, A.; Chaput, L.; Tadano, T.; Tanaka, I. Implementation strategies in phonopy and phono3py. *Journal of Physics: Condensed Matter* **2023**, *35* (35), 353001. DOI: 10.1088/1361-648X/acd831.
- (216) Togo, A.; Chaput, L.; Tanaka, I.; Hug, G. First-principles phonon calculations of thermal expansion in Ti₃SiC₂, Ti₃AlC₂, and Ti₃GeC₂. *Physical Review B* **2010**, *81* (17), 174301. DOI: 10.1103/PhysRevB.81.174301.
- (217) Niu, H.; Jing, Y.; Sun, Y.; Guo, L.; Aluru, N. R.; Li, W.; Yang, J.; Li, X. On the anomalous diffusion of proton in Y-doped BaZrO₃ perovskite oxide. *Solid State Ionics* **2022**, *376*, 115859. DOI: <https://doi.org/10.1016/j.ssi.2022.115859>.
- (218) Wang, Z.; Jing, Y.; Sun, Y.; Li, W.; Yang, J.; Li, X. Hydration induced mechanical degradation in the Y-doped BaZrO₃ solid oxide. *Computational Materials Science* **2024**, *235*, 112824. DOI: <https://doi.org/10.1016/j.commatsci.2024.112824>.
- (219) Li, M.; Liu, F.; Ding, D. Critical insights into the steam electrolysis electrode in protonic ceramic cells for hydrogen production. *Nature Catalysis* **2025**, *8* (4), 293-300. DOI: 10.1038/s41929-025-01313-w.
- (220) Lee, C.; Shin, S. S.; Kim, J.; Choi, J.; Choi, M.; Shin, H. H. Tailoring an Interface Microstructure for High-Performance Reversible Protonic Ceramic Electrochemical Cells via Soft Lithography. *ACS Applied Materials & Interfaces* **2022**, *14* (28), 32124-32133. DOI: 10.1021/acsami.2c08918.
- (221) Løken, A.; Ricote, S.; Wachowski, S. Thermal and Chemical Expansion in Proton Ceramic Electrolytes and Compatible Electrodes. *Crystals* **2018**, *8* (9), 365.
- (222) Zhang, W.; Zhang, Q.; Zhu, P.; Meng, Y.; Zhao, Z.; Wang, W.; Ding, Y.; Sun, Q.; Zhang, Y.; Li, M.; et al. An Active Oxygen Electrode for Proton-Conducting Solid Oxide Electrolysis

Cells with High Faradaic Efficiency. *Advanced Energy Materials* n/a (n/a), 2500852. DOI: <https://doi.org/10.1002/aenm.202500852>.

(223) Yang, L.; Wang, S.; Blinn, K.; Liu, M.; Liu, Z.; Cheng, Z.; Liu, M. Enhanced Sulfur and Coking Tolerance of a Mixed Ion Conductor for SOFCs: BaZr_{0.1}Ce_{0.7}Y_{0.2-x}O_{3-δ}. *Science* **2009**, 326 (5949), 126-129. DOI: doi:10.1126/science.1174811.

(224) Tang, W.; Bian, W.; Ding, H.; Ding, Y.; Zhao, Z.; Sun, Q.; Koomson, S.; Wang, Y.; Xu, B.; Dong, P.; et al. Sintering protonic zirconate cells with enhanced electrolysis stability and Faradaic efficiency. *Nature Synthesis* **2025**, 4 (5), 592-602. DOI: 10.1038/s44160-025-00765-z.

(225) Chen, T.; Jing, Y.; Anderson, L. O.; Leonard, K.; Matsumoto, H.; Aluru, N.; Perry, N. H. Toward Durable Protonic Ceramic Cells: Hydration-Induced Chemical Expansion Correlates with Symmetry in the Y-Doped BaZrO₃–BaCeO₃ Solid Solution. *The Journal of Physical Chemistry C* **2021**, 125 (47), 26216-26228. DOI: 10.1021/acs.jpcc.1c08334.

(226) Sereda, V. V.; Tsvetkov, D. S.; Malyshkin, D. A.; Ivanov, I. L.; Sednev-Lugovets, A. L.; Zuev, A. Y. Hydration-induced chemical expansion of BaCa_{(1+y)/3}Nb_{(2-y)/3}O_{3-δ}·xH₂O (BCN) and other proton-conducting perovskite oxides. *Solid State Ionics* **2020**, 358, 115516. DOI: <https://doi.org/10.1016/j.ssi.2020.115516>.

(227) Kasyanova, A. V.; Zvonareva, I. A.; Tarasova, N. A.; Bi, L.; Medvedev, D. A.; Shao, Z. Electrolyte materials for protonic ceramic electrochemical cells: Main limitations and potential solutions. *Materials Reports: Energy* **2022**, 2 (4), 100158. DOI: <https://doi.org/10.1016/j.matre.2022.100158>.

(228) Marrocchelli, D.; Perry, N. H.; Bishop, S. R. Understanding chemical expansion in perovskite-structured oxides. *Physical Chemistry Chemical Physics* **2015**, 17 (15), 10028-10039, 10.1039/C4CP05885B. DOI: 10.1039/C4CP05885B.

- (229) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; et al. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications* **2022**, *271*, 108171. DOI: <https://doi.org/10.1016/j.cpc.2021.108171>.
- (230) He, L.; Zhang, F.; Xuan, Y.; Zhang, L.; Gao, H.; Pan, H.; Lian, S.; Wang, M.; Yin, J.; Chen, X.; et al. Unveiling The Effect of Dopants on the Hydration Reaction and Proton Conduction of Nd and Y Co-Doped BaZrO₃ in Solid Oxide Fuel Cells. *Journal of The Electrochemical Society* **2021**, *168* (3), 034517. DOI: 10.1149/1945-7111/abef86.
- (231) Rodriguez, A.; Lin, C.; Yang, H.; Al-Fahdi, M.; Shen, C.; Choudhary, K.; Zhao, Y.; Hu, J.; Cao, B.; Zhang, H.; et al. Million-scale data integrated deep neural network for phonon properties of heuslers spanning the periodic table. *npj Computational Materials* **2023**, *9* (1), 20. DOI: 10.1038/s41524-023-00974-0.
- (232) Isayev, O.; Anstine, D.; Zubaiuk, R. AIMNet2: a neural network potential to meet your neutral, charged, organic, and elemental-organic needs. *Chemical Science* **2025**.
- (233) Bonati, L.; Polino, D.; Pizzolitto, C.; Biasi, P.; Eckert, R.; Reitmeier, S.; Schlögl, R.; Parrinello, M. The role of dynamics in heterogeneous catalysis: Surface diffusivity and N₂ decomposition on Fe(111). *Proceedings of the National Academy of Sciences* **2023**, *120* (50), e2313023120. DOI: doi:10.1073/pnas.2313023120.
- (234) Blöchl, P. E.; Jepsen, O.; Andersen, O. K. Improved tetrahedron method for Brillouin-zone integrations. *Physical Review B* **1994**, *49* (23), 16223-16233. DOI: 10.1103/PhysRevB.49.16223.

(235) Liu, B.; Evans, D.; Deng, H.; Edgar, J. Structure, Stability, and Electronic Properties of Boron Suboxide: A Density Functional Theory Study. *The Journal of Physical Chemistry C* **2022**, *126* (37), 16050-16060. DOI: 10.1021/acs.jpcc.2c05590.

Appendix A - A high-dimensional neural network potential for molecular dynamics simulations of condensed phase nickel and phase transitions

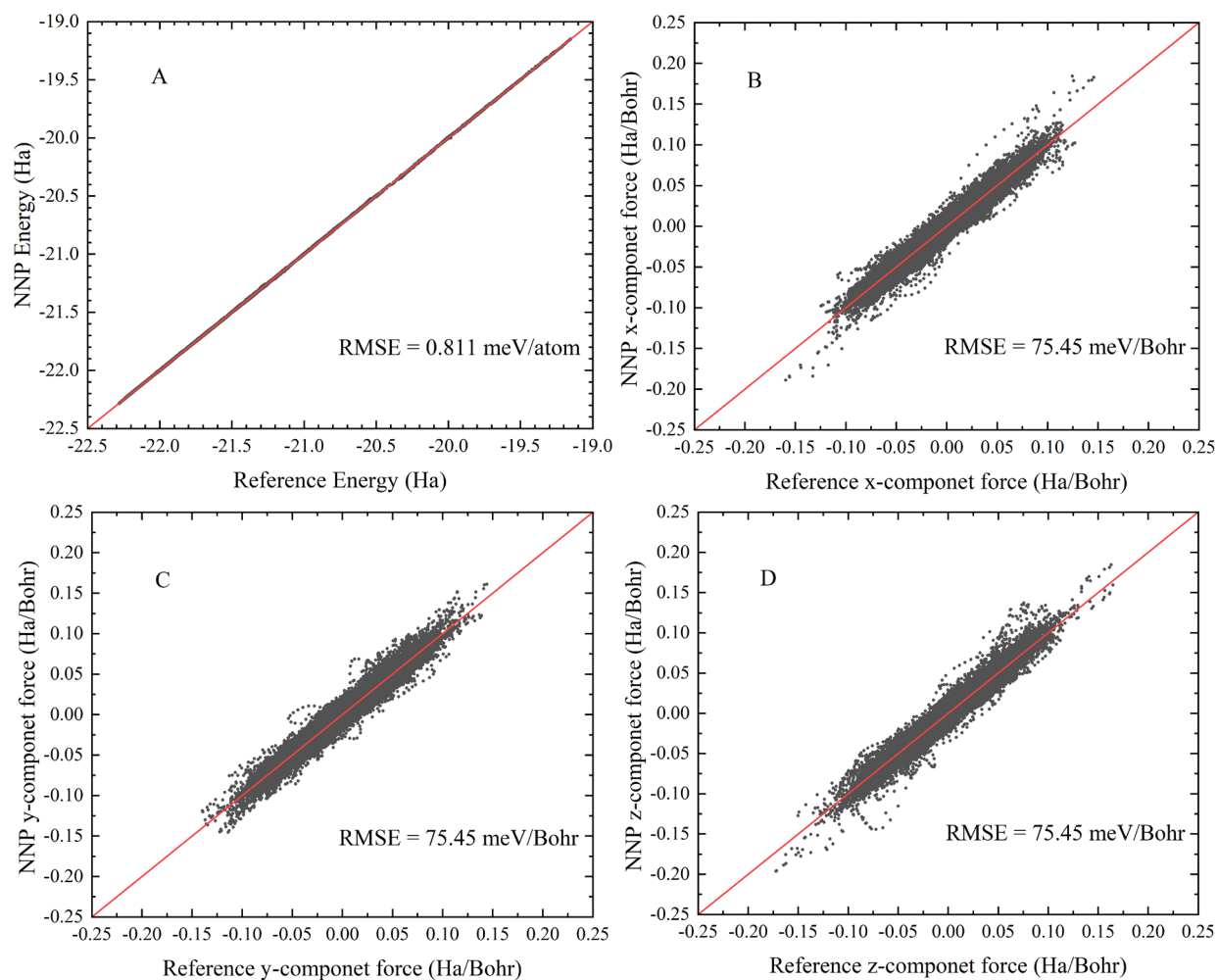
Detailed process of generating database for Ni NNP training

The database for Ni NNP training was all generated by AIMD. The strategy is exploring a wide range of pressure-temperature space around atmospheric pressure. The procedure could be summarized as following:

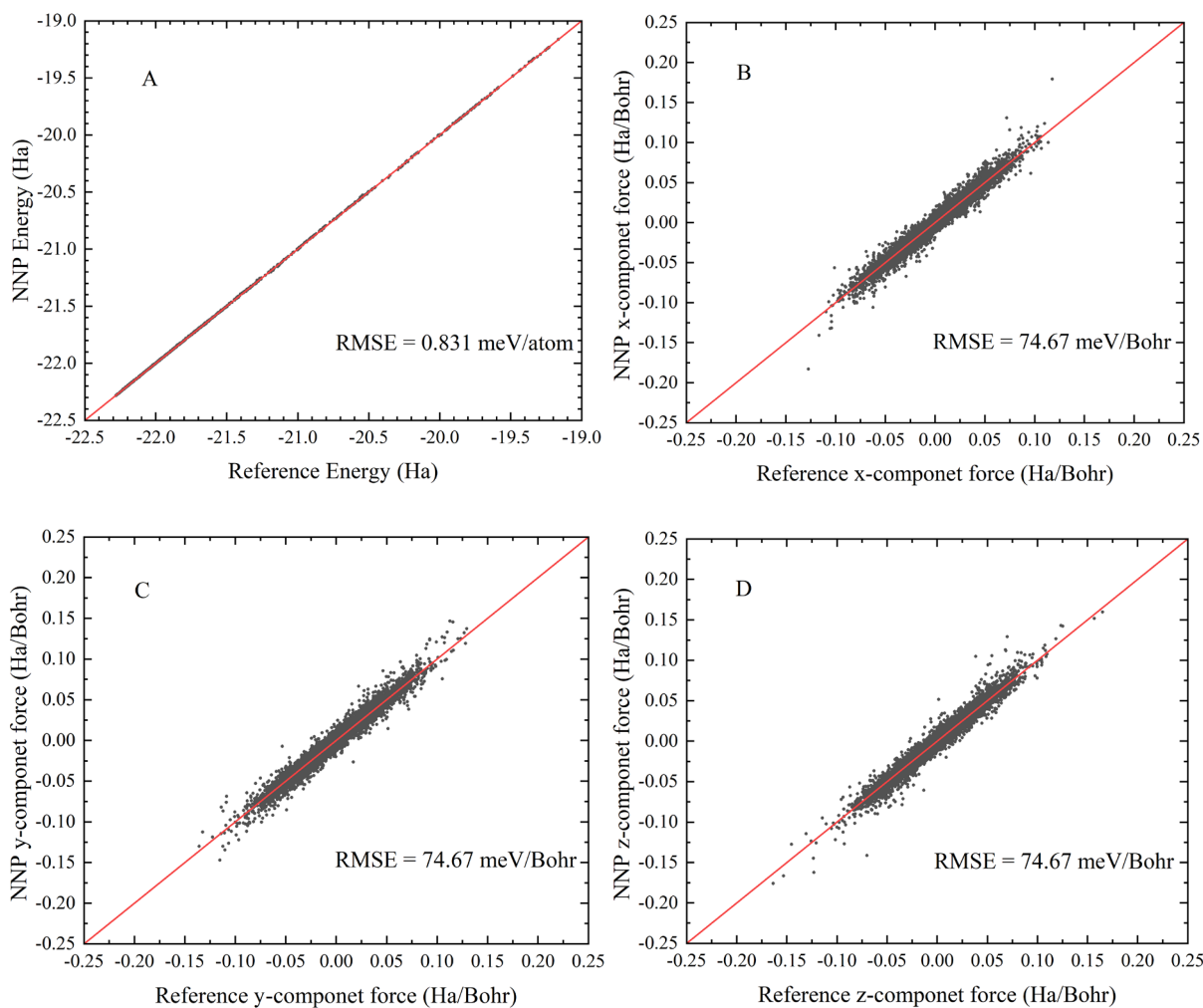
1. Generating the initial FCC structure by constructing $3 \times 3 \times 3$ of the conventional cell;
2. Selecting the temperature range based on our research interest. Since we hope to use the NNP to simulate behaviors of both liquid and solid phases and the phase transition, the range starting from 100K to 2589K was selected which included 12 temperature points. Some values are not neat just because they come from our former works and we included them directly;
3. Manipulating the scaling factors to make sure that at each temperature we could get three sets of data which are at around atmospheric pressure and a little lower and higher than atmospheric pressure. This is a kind of trial-and-error process which means people need to check the pressure after running AIMD for several steps to determine if the selected scaling factor is decent. The selected temperatures and corresponding scaling factors are shown in Table A.1.
4. Running AIMD for 200 steps to reach equilibrium (time step 2 fs). Then run another 200 steps to collect the trajectory file used for the training process.

Appendix Table A.1 Selected temperatures and corresponding scaling factors

Temperature (K)	100	300	490	680	863	1093
	1.01	1.01	1.019	1.023	1.02	1.033
Scaling Factors	1	1	1.008	1.01	1.01	1.018
	0.99	0.99	0.997	0.997	1	1.003
Temperature (K)	1323	1553	1726	2000	2295	2589
	1.041	1.05	1.039	1.052	1.073	1.09
Scaling Factors	1.025	1.03	1.029	1.043	1.053	1.0625
	1.009	1	1.019	1.03	1.033	1.045



Appendix Figure A.1 Comparison between energies and forces predicted by NNP and AIMD on training set: (A) energy, (B) x-component force, (C) y-component force, (D) z-component force.



Appendix Figure A.2 Comparison between energies and forces predicted by NNP and AIMD on testing set: (A) energy, (B) x-component force, (C) y-component force, (D) z-component force.

Appendix B - Predictions of Boron Phase Stability Using an Efficient Bayesian Machine Learning Interatomic Potential

Computational methods

Density functional theory (DFT): The projector augmented wave (PAW) pseudopotentials⁸¹, and the GGA-PBE functional¹³⁴ were used to account for the core electrons and the electron exchange-correlation effects. In OTF MD simulations, DFT calculations employed a cutoff energy of 400 eV, Gaussian smearing ($\sigma = 0.05$ eV), and a (4×4×4) mesh k -point mesh automatically generated from the Monkhorst-Pack scheme²¹³. For validations, the accuracy of DFT calculations is higher, i.e., cutoff energy of 520 eV, tetrahedron method with Blöchl corrections²³⁴, (6×6×6) k -point mesh.

A cutoff energy of 350 eV and gamma point were employed for AIMD simulations. All simulations used the Langevin thermostat for the NPT ensemble. The time step is 2 fs. The atomic and lattice degrees-of-freedom friction coefficients are 1.0 ps⁻¹. All simulations proceeded for 5000 steps (10 ps), and the volume was calculated based on the average of the last 4000 steps.

LAMMPS: The SGP-based energy minimization and MD simulations were employed with LAMMPS. The minimization style is the steepest descent algorithm. The stopping tolerances for energy and force during minimization are 1×10^{-4} and 1×10^{-6} eV/ Å, respectively. The maximum distance for the line search to move is 0.05 Å. The MD simulations employed the NPT ensemble with a time step of 0.5 fs. Each MD run lasts 1,000,000 steps (i.e., 0.5 ns of simulation time). The time-averaged values were calculated every 200,000 steps, and results from the last interval were collected for analysis.

Free energy calculations: To predict the stability of elemental boron, accurate and efficient calculations of free energy (G) are needed. As proved in our previous study²³⁵, the electronic and configurational entropy contributions in systems with high symmetry or low defect concentrations are relatively small compared to vibrational entropy. For pure boron compounds, it is reported that the zero-point energy (ZPE) can influence the energetic ordering¹⁵⁶. Therefore, the decent calculation of free energy can be shown as Eqn. 1:

$$G = H - TS + ZPE \quad (1)$$

where H is the enthalpy that can be calculated using:

$$H = E + PV \quad (2)$$

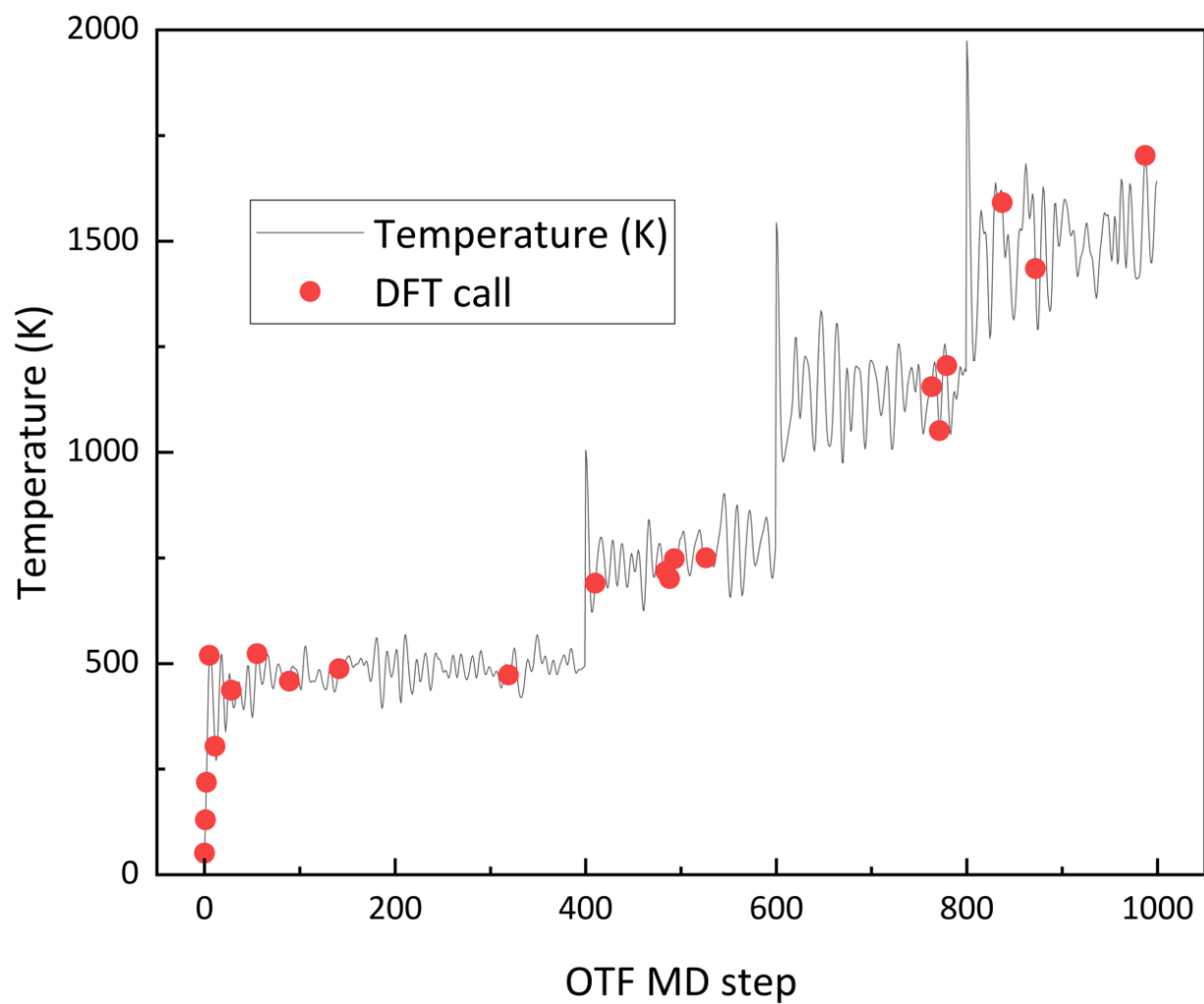
and E , P , V , T , and S are internal energy, pressure, volume, temperature, and vibrational entropy, respectively. H and T are collected from MD simulations, and T and ZPE are given by phonon calculations using Dump2phonon¹⁷¹.

Appendix Table B.1 Parameters used in OTF and offline training

	OTF training	Offline training
Kernel	Normalized Dot Product	Normalized Dot Product
Power of kernel	2	2
Descriptor	B2 ACE descriptor	B2 ACE descriptor
Cutoff radius	3.5 Å	5.5 Å
$n_{\max}$	8	14
$l_{\max}$	3	6
Uncertainty threshold	0.01 eV/Å	0.01 eV/Å

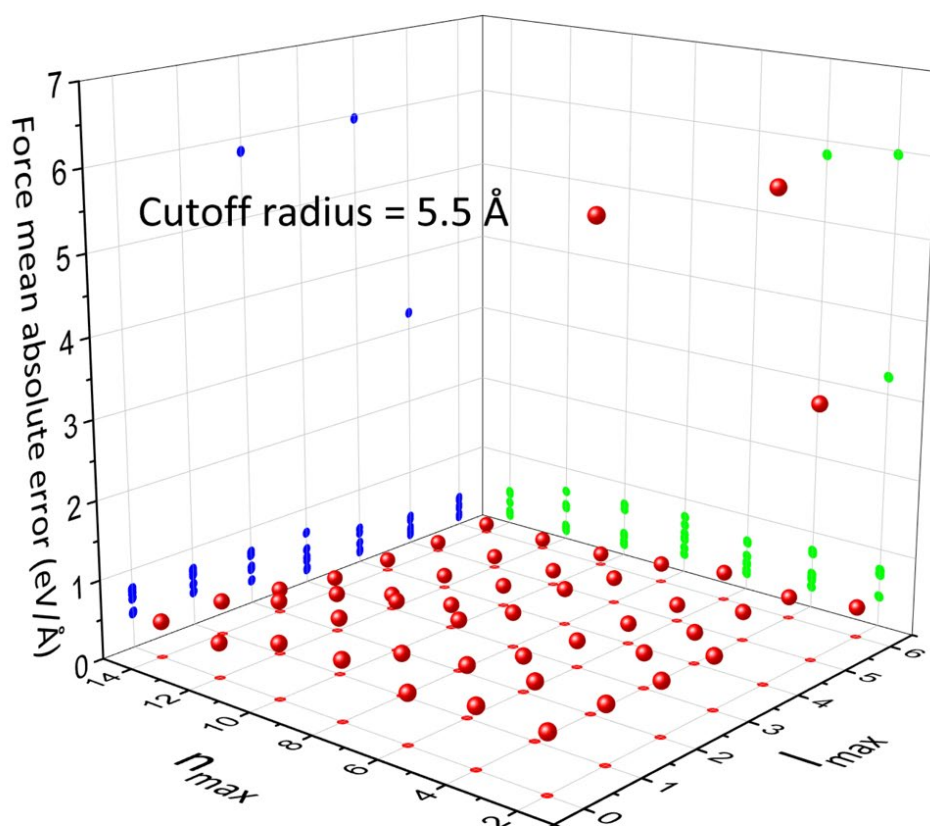
Appendix Table B.2 4th-order Birch-Murnaghan (BM) Eqn. of state (EOS) ¹⁶⁶ coefficients calculated from DFT and MLP results. And the mean absolute error (MAE) and mean absolute difference (MAD) of MLP predictions.

		4 th -order BM EOS: $E(V) = a + bV^{-2/3} + cV^{-4/3} + dV^{-2}$				MAE	MAD (eV)
		a	b	c	d		
α -B	DFT	4.75165	-88.57404	181.39894	-38.78579	0.07%	0.004353
	MLP	3.65424	-78.89218	155.54424	-19.77528		
β -B	DFT	3.16206	-74.44457	129.55622	43.32727	0.06%	0.003826
	MLP	5.75766	-104.53488	245.95259	-106.70463		
γ -B	DFT	5.95783	-98.39384	212.73029	-79.9814	0.03%	0.002287
	MLP	7.4671	-114.57088	270.18587	-147.48337		

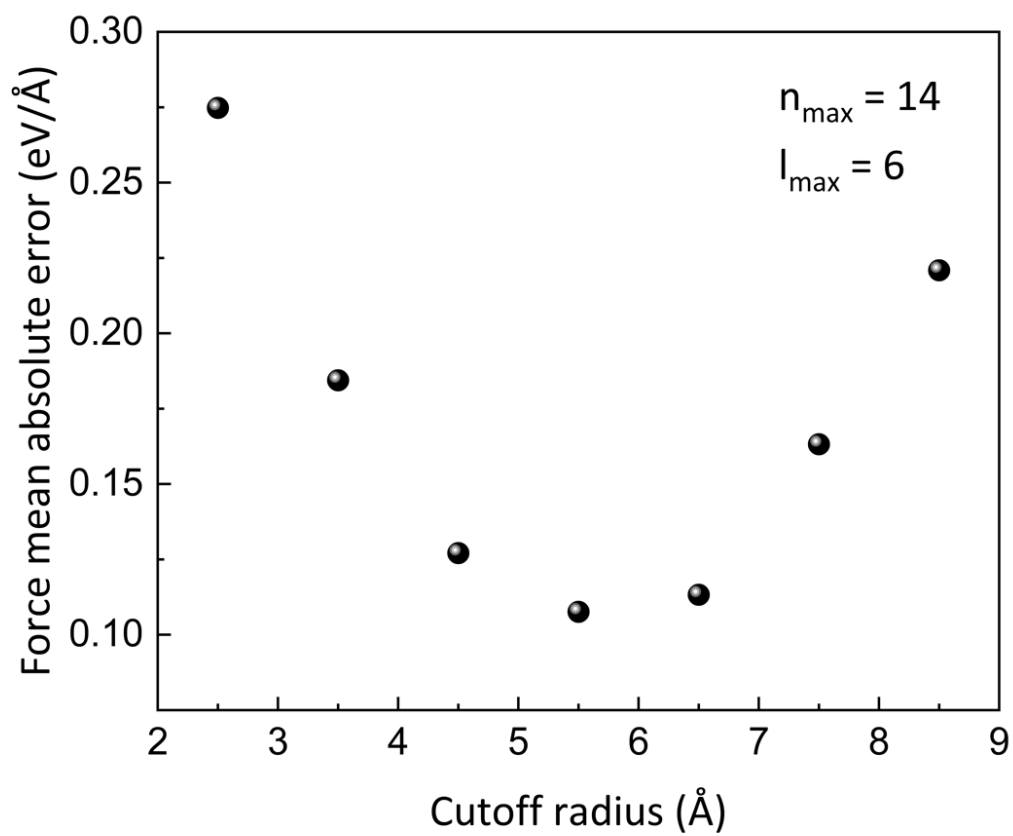


Appendix Figure B.1 System temperature during OTF MD of β -B with a scaling factor of 1.05.

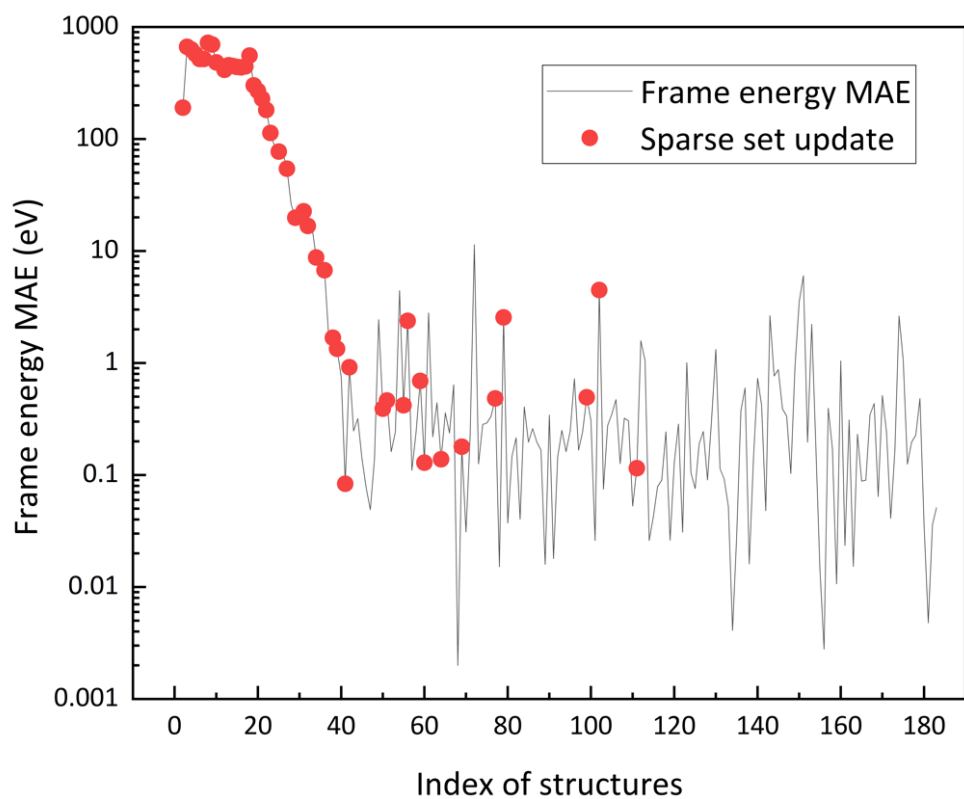
The steps associated with each DFT call are marked with red dots.



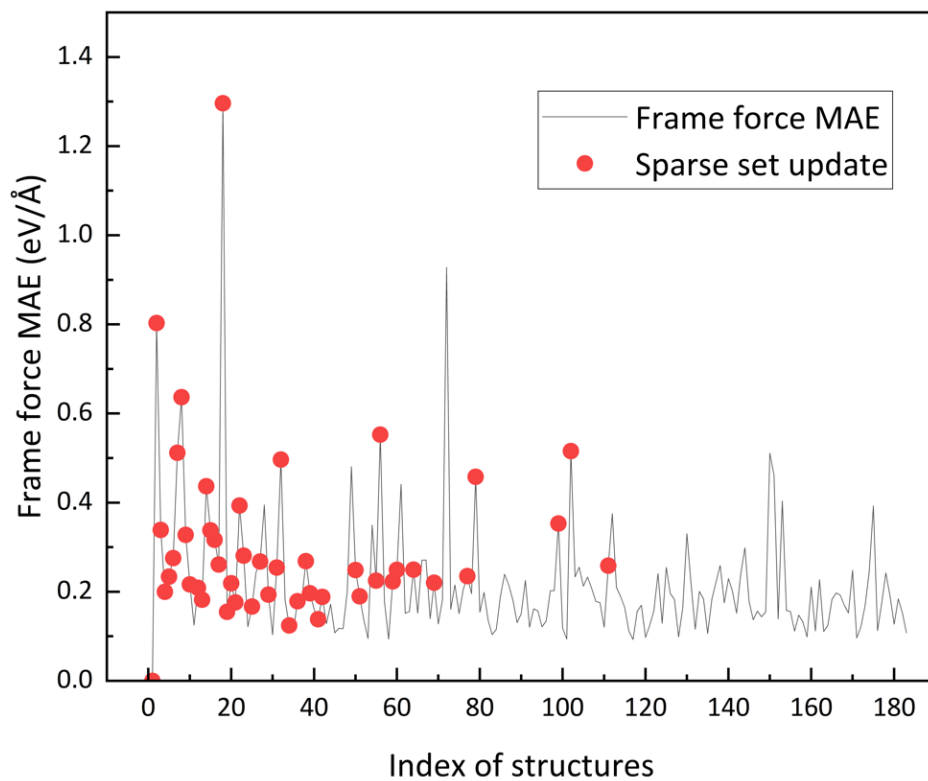
Appendix Figure B.2 Force mean absolute errors (MAEs) as a function of $n_{\max}$ and $l_{\max}$. Cutoff radius = 5.5 Å.



Appendix Figure B.3 Force mean absolute errors (MAEs) as a function of the cutoff radii. $n_{\text{max}} = 14$, $l_{\text{max}} = 6$.

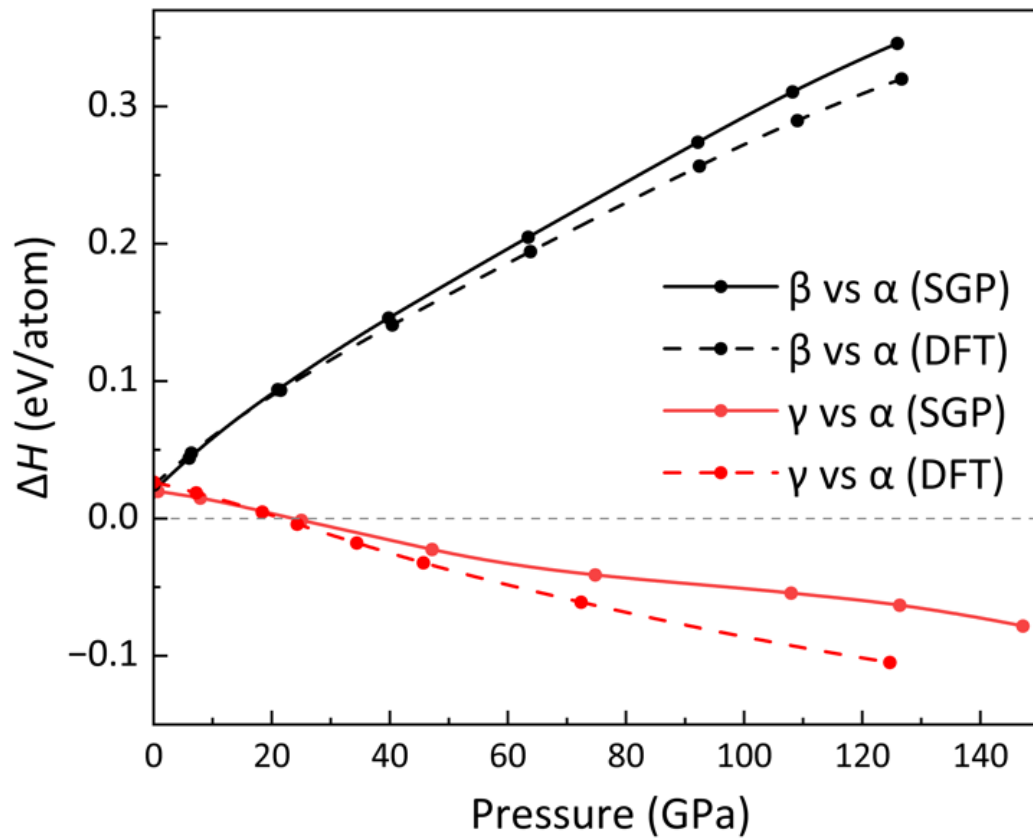


Appendix Figure B.4 Atomic energy mean absolute error (MAE) of every frame during offline training. Structures included in the sparse set are marked with red dots.

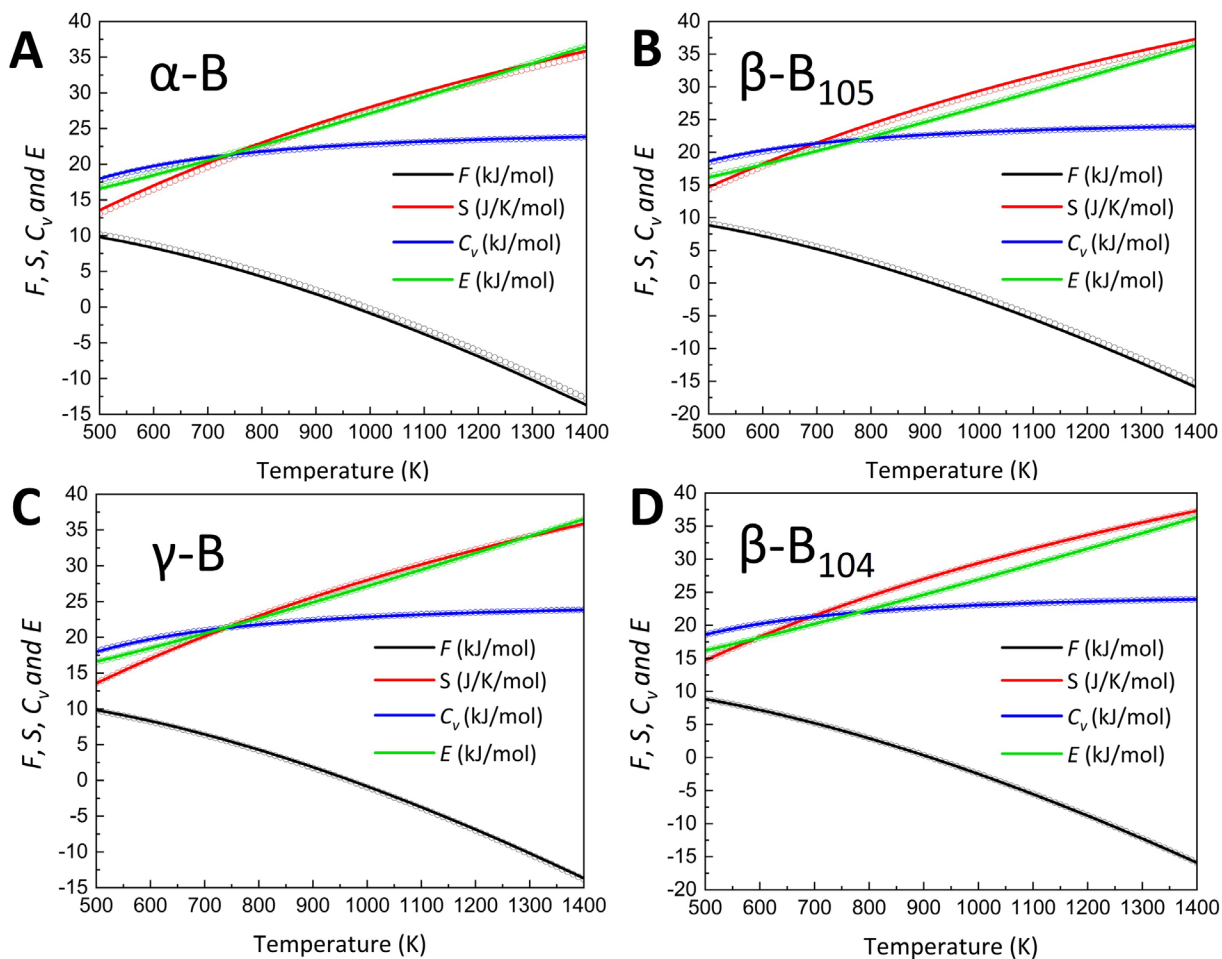


Appendix Figure B.5 Force mean absolute errors (MAEs) of every frame during offline training.

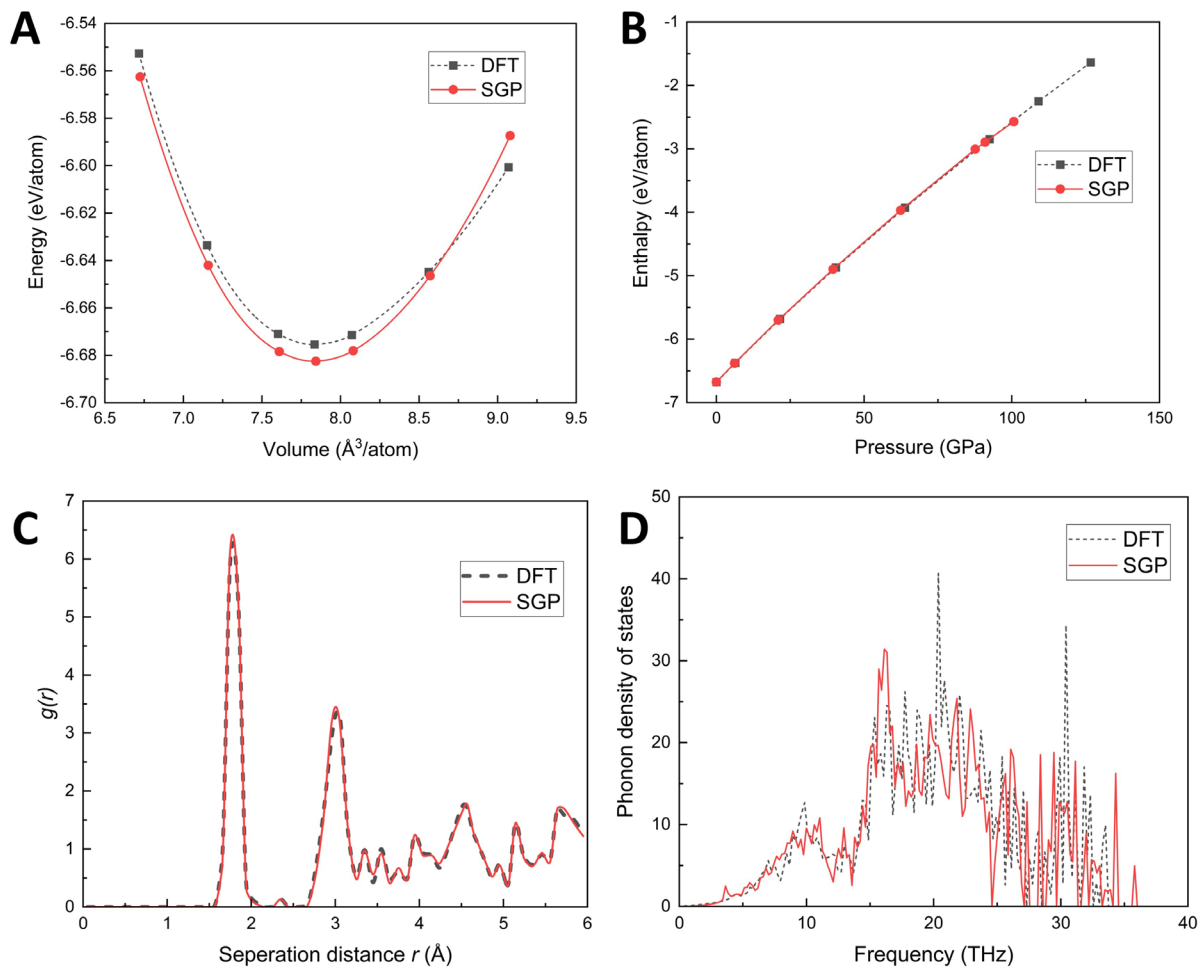
Structures included in the sparse set are marked with red dots.



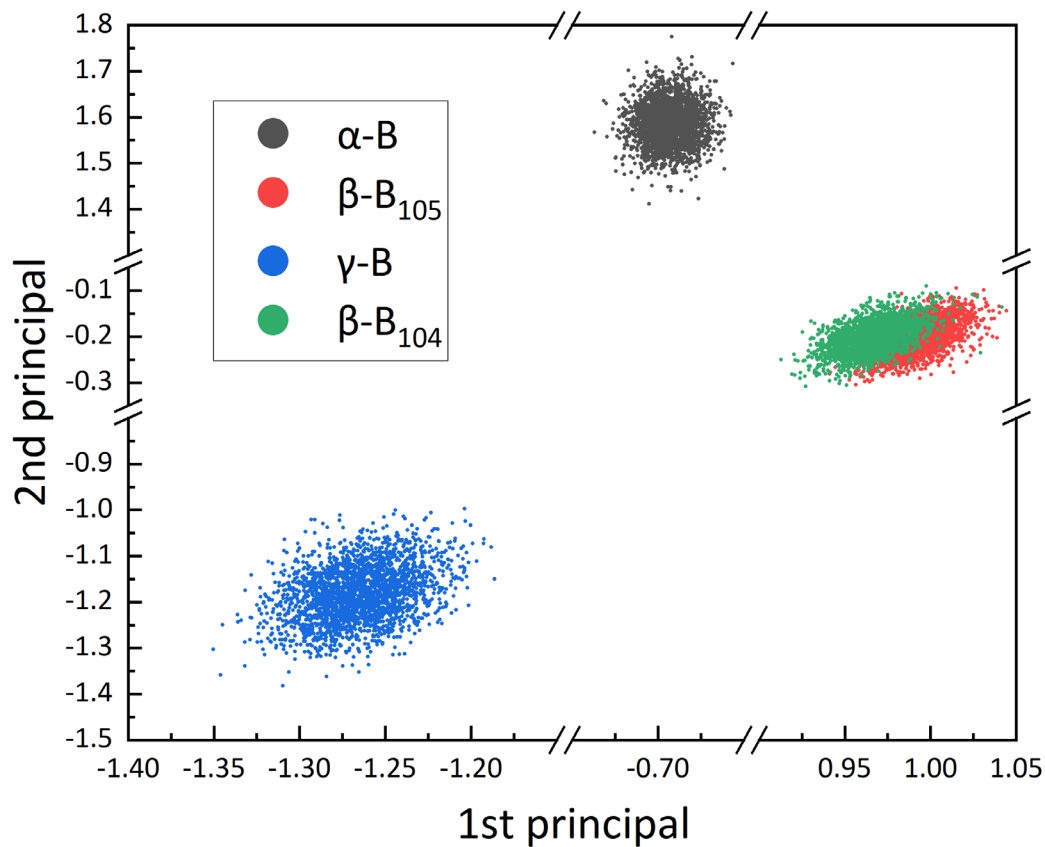
Appendix Figure B.6 Enthalpy differences (ΔH) versus pressure (0~150 GPa) for β -B (black) and γ -B (red) versus α -B. (SGP: solid lines; DFT: dashed lines)



Appendix Figure B.7 Free energy (F), entropy (S), heat capacity (C_v), and enthalpy (E) of (A) α -B, (B) β -B₁₀₅, (C) γ -B, and (D) β -B₁₀₄ based on the SGP potential simulations and DFT using the Phonopy code. DFT results are shown in hollow circles, and SGP results are shown in solid lines.



Appendix Figure B.8 (A) Equation of states, (B) atomic enthalpies (eV/atom), (C) radial distribution function $g(r)$, and (D) phonon density of states (PDOS) for β -B₁₀₄ based on DFT and SGP potential predictions.



Appendix Figure B.9 Principal component analysis (PCA) projection using SOAP descriptors. For each allotrope, 2500 configurations extracted from MD simulations every 0.2 ps under 500 K and 0 Pa were plotted. For building SOAP descriptors, cutoff radius, maximum radial and angular labels are 5.5 Å, 12, and 6, respectively.

Appendix C - Atomistic Simulations of Thermal and Chemical Expansions of $\text{PrNi}_x\text{Co}_{1-x}\text{O}_{3-\delta}$ Accelerated by Machine Learning Potentials

Appendix Table C.1 Lattice parameters ($\text{\AA}$) of optimized pristine and defective PNC unit cells.

	Pristine		$\text{V}_{\text{O}}^{\bullet\bullet}$ -defective		$\text{OH}_{\text{O}}^{\bullet}$ -defective	
PNC37	a	5.46	a	5.49	a	5.51
	b	5.50	b	5.53	b	5.55
	c	7.71	c	7.76	c	7.78
PNC55	a	5.44	a	5.47	a	5.48
	b	5.51	b	5.53	b	5.55
	c	7.70	c	7.74	c	7.76
PNC73	a	5.45	a	5.50	a	5.51
	b	5.47	b	5.52	b	5.54
	c	7.71	c	7.78	c	7.80

Appendix Table C.2 Bader charges (Q_B , e) of atoms in pristine and $V_O^{\bullet\bullet}$ -defective PNC37 and the charge change (ΔQ_B) due to the defect formation.

Atom index	Element	$Q_{B,pristine}$	$Q_{B, V_O^{\bullet\bullet}1}$	$\Delta Q_{B, V_O^{\bullet\bullet}1}$	$Q_{B, V_O^{\bullet\bullet}2}$	$\Delta Q_{B, V_O^{\bullet\bullet}2}$	$Q_{B, V_O^{\bullet\bullet}3}$	$\Delta Q_{B, V_O^{\bullet\bullet}3}$	$Q_{B, V_O^{\bullet\bullet}4}$	$\Delta Q_{B, V_O^{\bullet\bullet}4}$
1	Pr	8.843	8.907	+0.064	8.906	+0.064	8.927	+0.084	8.942	+0.100
2	Pr	8.846	8.894	+0.049	8.892	+0.046	8.924	+0.079	8.931	+0.086
3	Pr	8.846	8.861	+0.015	8.859	+0.013	8.842	-0.004	8.838	-0.008
4	Pr	8.843	8.898	+0.055	8.894	+0.052	8.840	-0.003	8.845	+0.002
5	Ni	8.679	8.864	+0.185	8.687	+0.008	8.681	+0.002	8.913	+0.234
6	Co	7.567	7.577	+0.009	7.767	+0.200	7.774	+0.207	7.588	+0.021
7	Co	7.432	7.721	+0.289	7.559	+0.127	7.710	+0.278	7.485	+0.053
8	Co	7.446	7.569	+0.123	7.724	+0.278	7.567	+0.121	7.872	+0.426
9	O	7.212	7.266	-	7.250	-	7.249	-	7.228	-
10	O	7.210	7.232	-	7.255	-	7.257	-	7.261	-
11	O	7.214	7.245	-	7.253	-	7.249	-	7.235	-
12	O	7.209	7.265	-	7.264	-	7.258	-	7.257	-
13	O	7.212	7.255	-	7.249	-	7.246	-	7.236	-
14	O	7.210	7.264	-	7.231	-	7.258	-	7.261	-
15	O	7.214	7.244	-	7.255	-	7.247	-	7.231	-
16	O	7.209	7.225	-	7.245	-	7.255	-	7.255	-
17	O	7.199	7.244	-	7.234	-	7.224	-	7.202	-
18	O	7.205	7.238	-	7.237	-	7.240	-	7.203	-
19	O	7.205	7.230	-	7.239	-	7.252	-	7.216	-
20	O	7.199								

Appendix Table C.3 Bader charge (Q_B , e) of atoms in pristine and OH_O -defective PNC37 and the charge change (ΔQ_B) due to the defect formation.

Atom index	Element	$Q_{B, \text{pristine}}$	$Q_{B, \text{OH}_\text{O}^1}$	$\Delta Q_{B, \text{OH}_\text{O}^1}$	$Q_{B, \text{OH}_\text{O}^2}$	$\Delta Q_{B, \text{OH}_\text{O}^2}$	$Q_{B, \text{OH}_\text{O}^3}$	$\Delta Q_{B, \text{OH}_\text{O}^3}$	$Q_{B, \text{OH}_\text{O}^4}$	$\Delta Q_{B, \text{OH}_\text{O}^4}$
1	Pr	8.843	8.841	-0.002	8.852	+0.010	8.864	+0.021	8.870	+0.027
2	Pr	8.846	8.848	+0.002	8.861	+0.015	8.851	+0.005	8.866	+0.020
3	Pr	8.846	8.873	+0.027	8.852	+0.006	8.854	+0.009	8.848	+0.002
4	Pr	8.843	8.854	+0.011	8.853	+0.010	8.858	+0.015	8.849	+0.006
5	Ni	8.679	8.715	+0.036	8.694	+0.015	8.696	+0.017	8.694	+0.015
6	Ni	7.567	7.586	+0.019	7.607	+0.040	7.641	+0.074	7.484	-0.083
7	Ni	7.432	7.477	+0.045	7.434	+0.002	7.442	+0.010	7.551	+0.119
8	Co	7.446	7.579	+0.133	7.610	+0.164	7.586	+0.140	7.576	+0.130
9	O	7.212	7.236	-	7.220	-	7.199	-	7.238	-
10	O	7.210	7.248	-	7.263	-	7.262	-	7.238	-
11	O	7.214	7.686	-	7.218	-	7.201	-	7.250	-
12	O	7.209	7.245	-	7.269	-	7.253	-	7.195	-
13	O	7.212	7.229	-	7.210	-	7.204	-	7.255	-
14	O	7.210	7.236	-	7.262	-	7.274	-	7.239	-
15	O	7.214	7.207	-	7.212	-	7.221	-	7.250	-
16	O	7.209	7.243	-	7.256	-	7.262	-	7.699	-
17	O	7.199	7.213	-	7.202	-	7.161	-	7.192	-
18	O	7.205	7.236	-	7.217	-	7.248	-	7.241	-
19	O	7.205	7.238	-	7.708	-	7.243	-	7.240	-
20	O	7.199	7.212	-	7.200	-	7.680	-	7.226	-
21	H		0.000	-	0.000	-	0.000	-	0.000	-

Appendix Table C.4 Bader charge (Q_B , e) of atoms in pristine and $V_O^{\bullet\bullet}$ -defective PNC55 and the charge change (ΔQ_B) due to the defect formation.

Atom index	Element	$Q_{B,pristine}$	$Q_B, V_O^{\bullet\bullet}1$	$\Delta Q_B, V_O^{\bullet\bullet}1$	$Q_B, V_O^{\bullet\bullet}2$	$\Delta Q_B, V_O^{\bullet\bullet}2$	$Q_B, V_O^{\bullet\bullet}3$	$\Delta Q_B, V_O^{\bullet\bullet}3$	$Q_B, V_O^{\bullet\bullet}4$	$\Delta Q_B, V_O^{\bullet\bullet}4$
1	Pr	8.843	8.909	+0.066	8.910	+0.067	8.928	+0.085	8.929	+0.086
2	Pr	8.843	8.892	+0.049	8.890	+0.047	8.929	+0.086	8.928	+0.085
3	Pr	8.843	8.858	+0.015	8.858	+0.015	8.843	+0.000	8.839	-0.004
4	Pr	8.843	8.900	+0.057	8.900	+0.057	8.839	-0.004	8.843	+0.000
5	Ni	8.679	8.854	+0.175	8.685	+0.006	8.683	+0.004	8.847	+0.168
6	Ni	8.679	8.683	+0.004	8.825	+0.146	8.847	+0.168	8.683	+0.004
7	Co	7.427	7.695	+0.268	7.559	+0.132	7.704	+0.277	7.556	+0.129
8	Co	7.427	7.556	+0.129	7.711	+0.284	7.556	+0.129	7.704	+0.277
9	O	7.202	7.254	-	7.245	-	7.245	-	7.245	-
10	O	7.202	7.224	-	7.240	-	7.245	-	7.245	-
11	O	7.206	7.245	-	7.252	-	7.239	-	7.252	-
12	O	7.206	7.250	-	7.261	-	7.252	-	7.239	-
13	O	7.202	7.250	-	7.245	-	7.246	-	7.247	-
14	O	7.202	7.257	-	7.227	-	7.247	-	7.246	-
15	O	7.206	7.243	-	7.251	-	7.239	-	7.246	-
16	O	7.206	7.225	-	7.240	-	7.246	-	7.239	-
17	O	7.196	7.238	-	7.233	-	7.229	-	7.233	-
18	O	7.196	7.233	-	7.235	-	7.233	-	7.229	-
19	O	7.196	7.233	-	7.231	-	7.249	-	7.248	-
20	O	7.196								

Appendix Table C.5 Bader charge (Q_B , e) of atoms in pristine and OH_O -defective PNC55 and the charge change (ΔQ_B) due to the defect formation.

Atom index	Element	$Q_{B,\text{pristine}}$	$Q_{B,\text{OH}_\text{O}1}$	$\Delta Q_{B,\text{OH}_\text{O}1}$	$Q_{B,\text{OH}_\text{O}2}$	$\Delta Q_{B,\text{OH}_\text{O}2}$	$Q_{B,\text{OH}_\text{O}3}$	$\Delta Q_{B,\text{OH}_\text{O}3}$	$Q_{B,\text{OH}_\text{O}4}$	$\Delta Q_{B,\text{OH}_\text{O}4}$
1	Pr	8.843	8.838	-0.005	8.852	+0.009	8.861	+0.018	8.869	+0.025
2	Pr	8.843	8.853	+0.010	8.864	+0.021	8.851	+0.008	8.863	+0.020
3	Pr	8.843	8.872	+0.029	8.854	+0.011	8.852	+0.009	8.841	-0.002
4	Pr	8.843	8.852	+0.009	8.852	+0.009	8.849	+0.006	8.845	+0.002
5	Ni	8.679	8.780	+0.101	8.706	+0.027	8.691	+0.011	8.689	+0.010
6	Ni	8.679	8.695	+0.016	8.699	+0.020	8.721	+0.042	8.713	+0.034
7	Ni	7.427	7.324	-0.103	7.299	-0.128	7.430	+0.003	7.422	-0.006
8	Co	7.427	7.435	+0.008	7.584	+0.157	7.577	+0.150	7.579	+0.152
9	O	7.202	7.284	-	7.250	-	7.191	-	7.209	-
10	O	7.202	7.220	-	7.254	-	7.260	-	7.254	-
11	O	7.206	7.727	-	7.209	-	7.191	-	7.208	-
12	O	7.206	7.216	-	7.250	-	7.254	-	7.243	-
13	O	7.202	7.278	-	7.268	-	7.196	-	7.232	-
14	O	7.202	7.205	-	7.253	-	7.269	-	7.249	-
15	O	7.206	7.261	-	7.208	-	7.209	-	7.206	-
16	O	7.206	7.215	-	7.260	-	7.262	-	7.698	-
17	O	7.196	7.280	-	7.206	-	7.178	-	7.194	-
18	O	7.196	7.194	-	7.226	-	7.248	-	7.240	-
19	O	7.196	7.195	-	7.699	-	7.237	-	7.242	-
20	O	7.196	7.273	-	7.206	-	7.675	-	7.207	-
21	H		0.000	-	0.000	-	0.000	-	0.000	-

Appendix Table C.6 Bader charge (Q_B , e) of atoms in pristine and $V_O^{\bullet}$ -defective PNC73 and the charge change (ΔQ_B) due to the defect formation.

Atom index	Element	$Q_{B,pristine}$	$Q_{B,V_O^{\bullet}1}$	$\Delta Q_{B,V_O^{\bullet}1}$	$Q_{B,V_O^{\bullet}2}$	$\Delta Q_{B,V_O^{\bullet}2}$	$Q_{B,V_O^{\bullet}3}$	$\Delta Q_{B,V_O^{\bullet}3}$	$Q_{B,V_O^{\bullet}4}$	$\Delta Q_{B,V_O^{\bullet}4}$
1	Pr	8.839	8.896	+0.057	8.902	+0.063	8.928	+0.090	8.920	+0.081
2	Pr	8.844	8.889	+0.044	8.888	+0.044	8.927	+0.082	8.927	+0.083
3	Pr	8.844	8.863	+0.018	8.860	+0.016	8.841	-0.003	8.842	-0.002
4	Pr	8.839	8.893	+0.054	8.892	+0.053	8.836	-0.003	8.838	+0.000
5	Ni	8.652	8.852	+0.201	8.659	+0.007	8.661	+0.009	8.840	+0.188
6	Ni	8.663	8.672	+0.010	8.848	+0.186	8.852	+0.189	8.669	+0.006
7	Ni	8.570	8.810	+0.240	8.600	+0.030	8.799	+0.229	8.602	+0.032
8	Co	7.422	7.430	+0.008	7.705	+0.283	7.534	+0.112	7.694	+0.272
9	O	7.195	7.298	-	7.235	-	7.244	-	7.250	-
10	O	7.194	7.195	-	7.243	-	7.232	-	7.234	-
11	O	7.197	7.208	-	7.238	-	7.247	-	7.256	-
12	O	7.196	7.304	-	7.259	-	7.240	-	7.237	-
13	O	7.195	7.214	-	7.236	-	7.245	-	7.252	-
14	O	7.194	7.280	-	7.219	-	7.240	-	7.234	-
15	O	7.197	7.206	-	7.245	-	7.244	-	7.256	-
16	O	7.196	7.283	-	7.260	-	7.231	-	7.236	-
17	O	7.197	7.214	-	7.223	-	7.259	-	7.238	-
18	O	7.184	7.210	-	7.228	-	7.209	-	7.219	-
19	O	7.184	7.284	-	7.260	-	7.231	-	7.256	-
20	O	7.197								

Appendix Table C.7 Bader charge (Q_B , e) of atoms in pristine and OH_O -defective PNC73 and the charge change (ΔQ_B) due to the vacancy formation.

Atom index	Element	$Q_{B, \text{pristine}}$	$Q_{B, \text{OH}_\text{O}^1}$	$\Delta Q_{B, \text{OH}_\text{O}^1}$	$Q_{B, \text{OH}_\text{O}^2}$	$\Delta Q_{B, \text{OH}_\text{O}^2}$	$Q_{B, \text{OH}_\text{O}^3}$	$\Delta Q_{B, \text{OH}_\text{O}^3}$	$Q_{B, \text{OH}_\text{O}^4}$	$\Delta Q_{B, \text{OH}_\text{O}^4}$
1	Pr	8.839	8.829	-0.009	8.858	+0.019	8.863	+0.024	8.864	+0.025
2	Pr	8.844	8.859	+0.015	8.857	+0.012	8.847	+0.003	8.865	+0.021
3	Pr	8.844	8.872	+0.027	8.852	+0.008	8.850	+0.006	8.840	-0.004
4	Pr	8.839	8.845	+0.007	8.858	+0.019	8.854	+0.015	8.846	+0.007
5	Ni	8.652	8.666	+0.014	8.659	+0.007	8.666	+0.014	8.635	-0.017
6	Ni	8.663	8.681	+0.018	8.691	+0.028	8.688	+0.025	8.710	+0.047
7	Ni	8.570	8.665	+0.095	8.681	+0.110	8.668	+0.098	8.676	+0.106
8	Co	7.422	7.428	+0.006	7.426	+0.004	7.424	+0.002	7.419	-0.003
9	O	7.195	7.241	-	7.238	-	7.231	-	7.240	-
10	O	7.194	7.216	-	7.179	-	7.213	-	7.203	-
11	O	7.197	7.714	-	7.237	-	7.266	-	7.241	-
12	O	7.196	7.208	-	7.192	-	7.208	-	7.192	-
13	O	7.195	7.244	-	7.259	-	7.235	-	7.252	-
14	O	7.194	7.192	-	7.198	-	7.193	-	7.178	-
15	O	7.197	7.242	-	7.232	-	7.237	-	7.242	-
16	O	7.196	7.209	-	7.208	-	7.197	-	7.670	-
17	O	7.197	7.253	-	7.271	-	7.269	-	7.277	-
18	O	7.184	7.185	-	7.172	-	7.191	-	7.184	-
19	O	7.184	7.191	-	7.663	-	7.176	-	7.187	-
20	O	7.197	7.260	-	7.270	-	7.724	-	7.280	-
21	H	-	0.000		0.000		0.000		0.000	

Appendix Table C.8 Parameters of two-dimensional polynomial fitting of the relationship between expansion (dL/d_0) and co-effect of temperature (T) and hydration (H), as shown in Figure 5. The R^2 values of each fitting are shown in the last row.

Formula: $dL/d_0 = z_0 + a_1 * T + a_2 * T^2 + a_3 * T^3 + a_4 * T^4 + a_5 * T^5$ $+ b_1 * H + b_2 * H^2 + b_3 * H^3 + b_4 * H^4 + b_5 * H^5$			
Material	PNC37	PNC55	PNC73
z₀	-3.272E+00	-2.529E+00	-2.299E+00
a₁	8.070E-03	4.070E-03	3.390E-03
a₂	1.080E-05	2.160E-05	1.790E-05
a₃	1.240E-10	-2.020E-08	-1.900E-08
a₄	-8.010E-12	1.030E-11	1.180E-11
a₅	5.390E-15	-1.560E-15	-2.680E-15
b₁	1.515E+01	7.316E+00	1.028E+01
b₂	-3.813E+01	-5.357E-02	-1.052E-01
b₃	7.821E+01	6.570E-04	1.790E-03
b₄	-7.913E+01	-9.720E-06	-3.660E-05
b₅	3.159E+01	1.510E-07	7.700E-07
R²	1.0000	0.9982	0.9993

Appendix Table C.9 Comparisons between GGA-PBE and r²SCAN+rVV10 functionals on energies (in eV) of PNC73 and corresponding V_O^{••} formation.

	GGA-PBE		r ² SCAN+rVV10	
	$\Delta E_{V_O^{\bullet\bullet}}$	$\Delta E_{V_{O1}^{\bullet\bullet}} - \Delta E_{V_{O1}^{\bullet\bullet}}$	$\Delta E_{V_O^{\bullet\bullet}}$	$\Delta E_{V_{O1}^{\bullet\bullet}} - \Delta E_{V_{O1}^{\bullet\bullet}}$
V_{O1}^{••}	1.639	0.000	1.828	0.000
V_{O2}^{••}	1.844	0.205	1.966	0.138
V_{O3}^{••}	1.922	0.284	2.084	0.256
V_{O4}^{••}	1.733	0.094	1.939	0.111

Appendix Table C.10 Parameters used in OTF-MD for dataset generation with FLARE

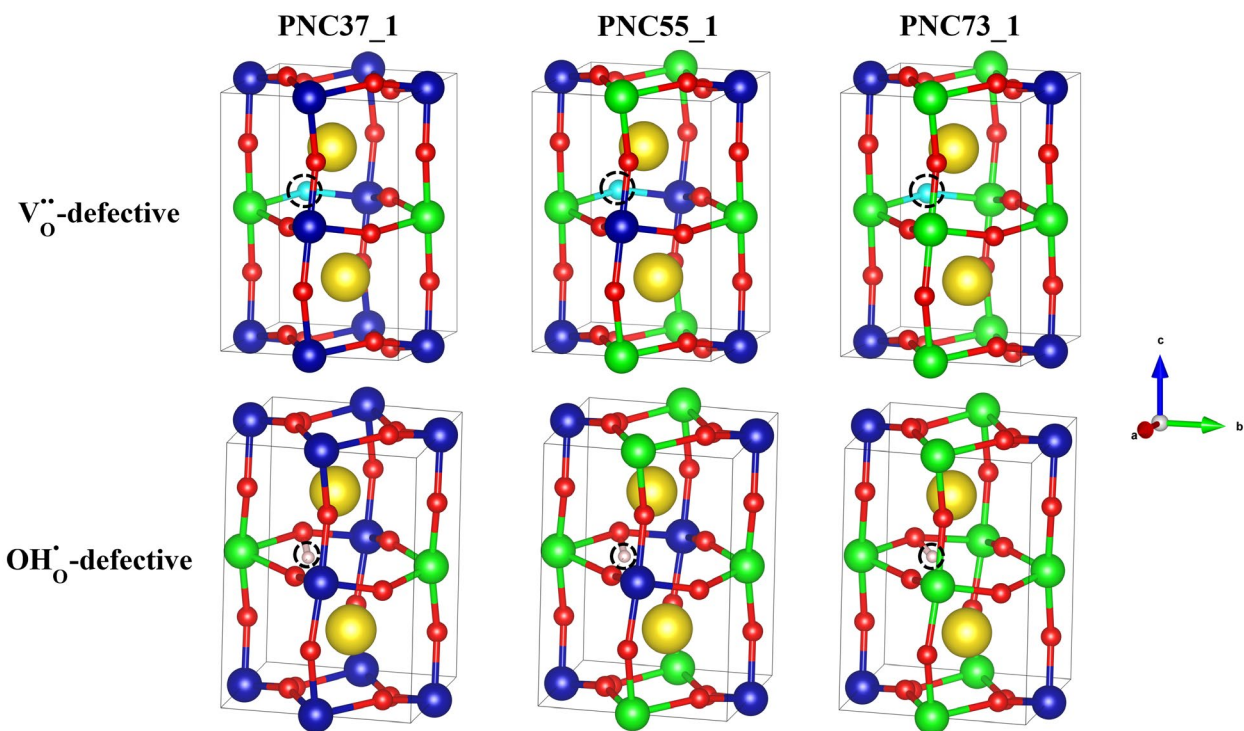
	Pristine lattice	Vo-defective	OH-defective
Uncertainty threshold	0.01 eV/Å	0.25 eV/Å	0.39 eV/Å
Time step	2 fs	2 fs	0.5 fs
Number of steps	10,000		
Ensemble	NVT		
Temperature	1300 K		
Kernel	Normalized Dot Product		
Power of kernel	2		
Descriptor	B2 ACE descriptor		
Cutoff radius	4.0 Å		
n_{max}	8		
l_{max}	3		

Appendix Table C.11 Parameters used in final offline training using Allegro

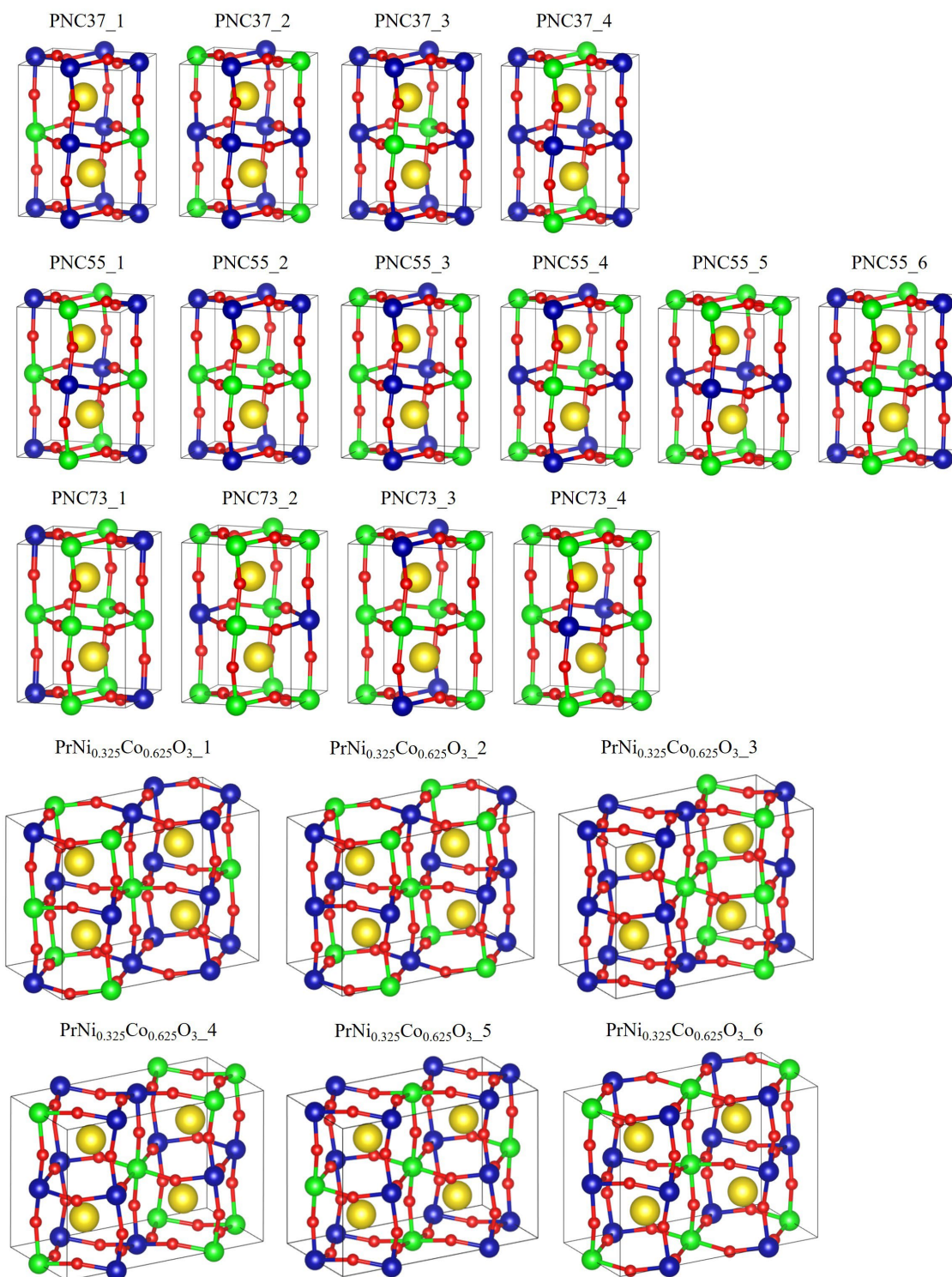
Cutoff radius	5.0 Å
Number of Bessel radial basis	8
Polynomial cutoff	48
Maximum order of spherical harmonic embedding	3
Parity	o3_full
Number of tensor product layers	2
Number of features	16
2-body latent multi-layer perceptron dimension	[64, 128, 256]
Later latent multi-layer perceptron dimension	[256, 256, 256]
Final edge energy multi-layer perceptron dimension	[64]
Learning rate	0.001
Batch size	1

Appendix Table C.12 Wall time (in second) comparison between DFT and MLP on single-point energy calculation and geometry optimization on PNC73 unit cell.

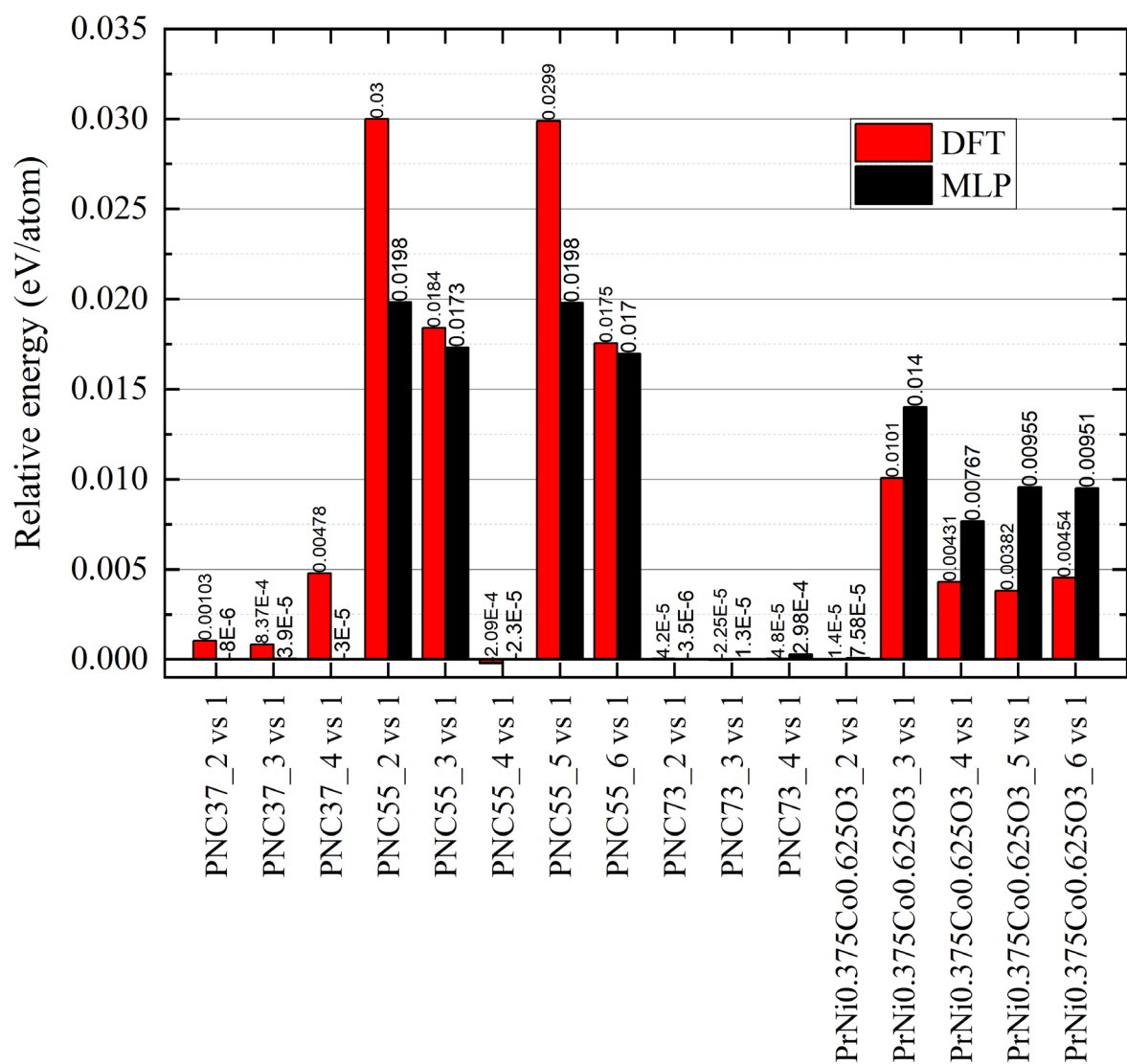
	DFT	MLP
Single-point energy	2765 (8 cores)	1 (1 core)
Geometry optimization	1957(16 cores)	10 (1 core)



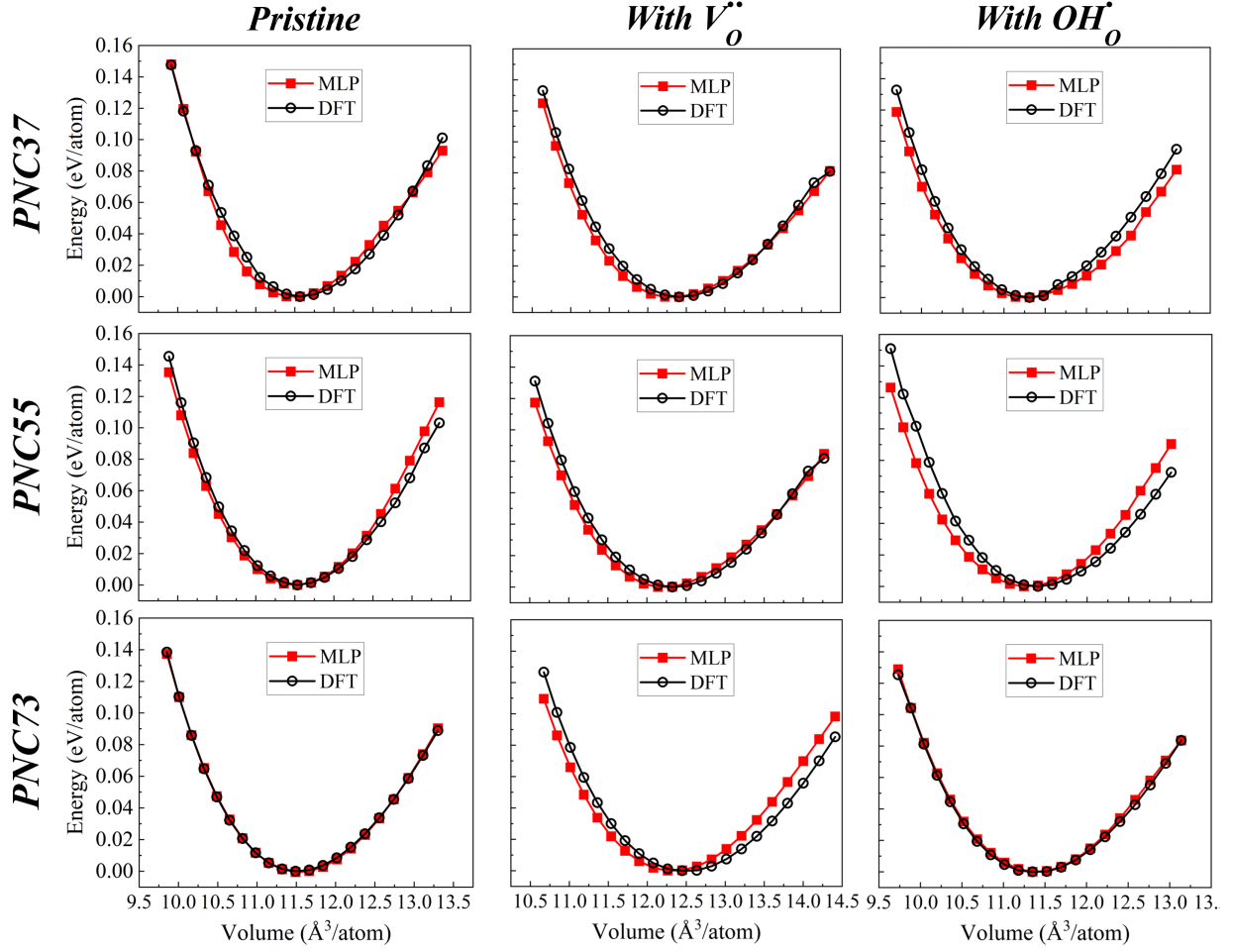
Appendix Figure C.1 Schematic illustrations of defective PNC structures for MLP-based phonon calculations. Color scheme: Pr (yellow), Ni (green), Co (dark blue), O (red), V_O'' (cyan), and H (pink). Defects are highlighted with black dashed circles.



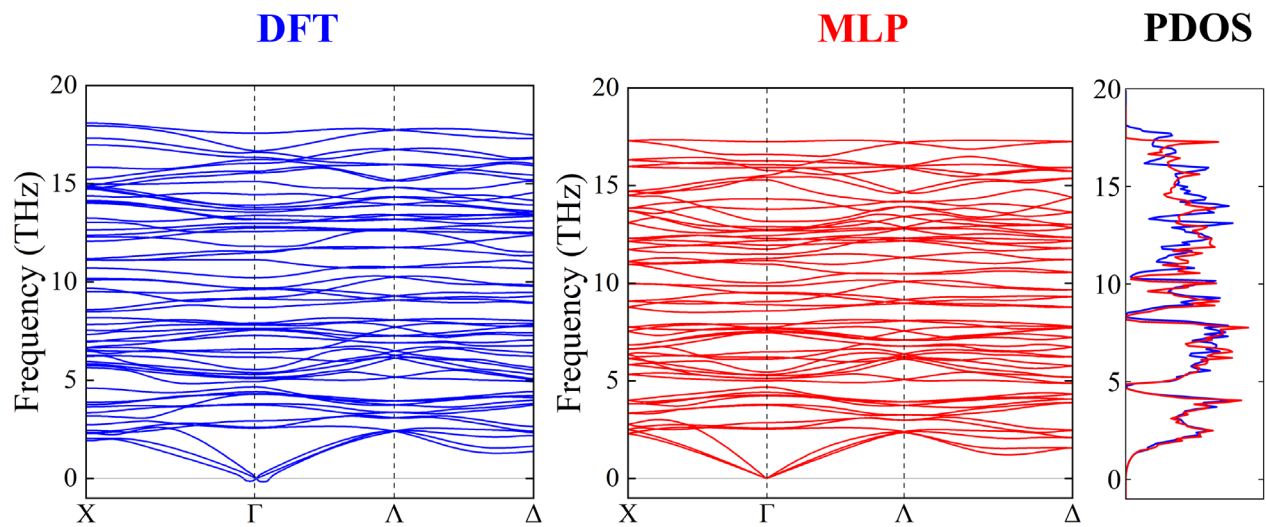
Appendix Figure C.2 Illustrations of PNCs with different Ni/Co ordering for calculations in Figure S9. Color scheme: Pr (yellow), Ni (green), Co (dark blue), and O (red).



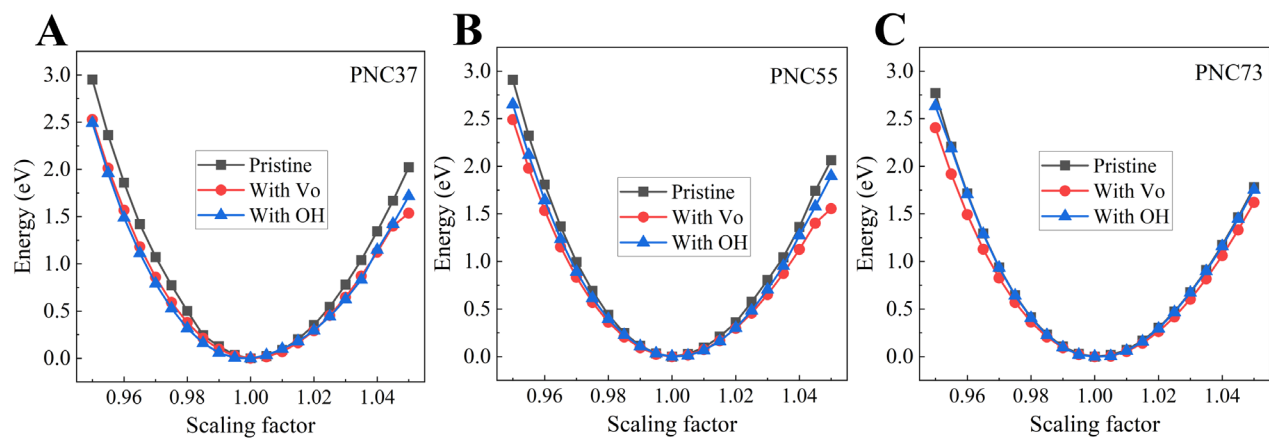
Appendix Figure C.3 Relative energies of PNCs with different Ni/Co ordering.



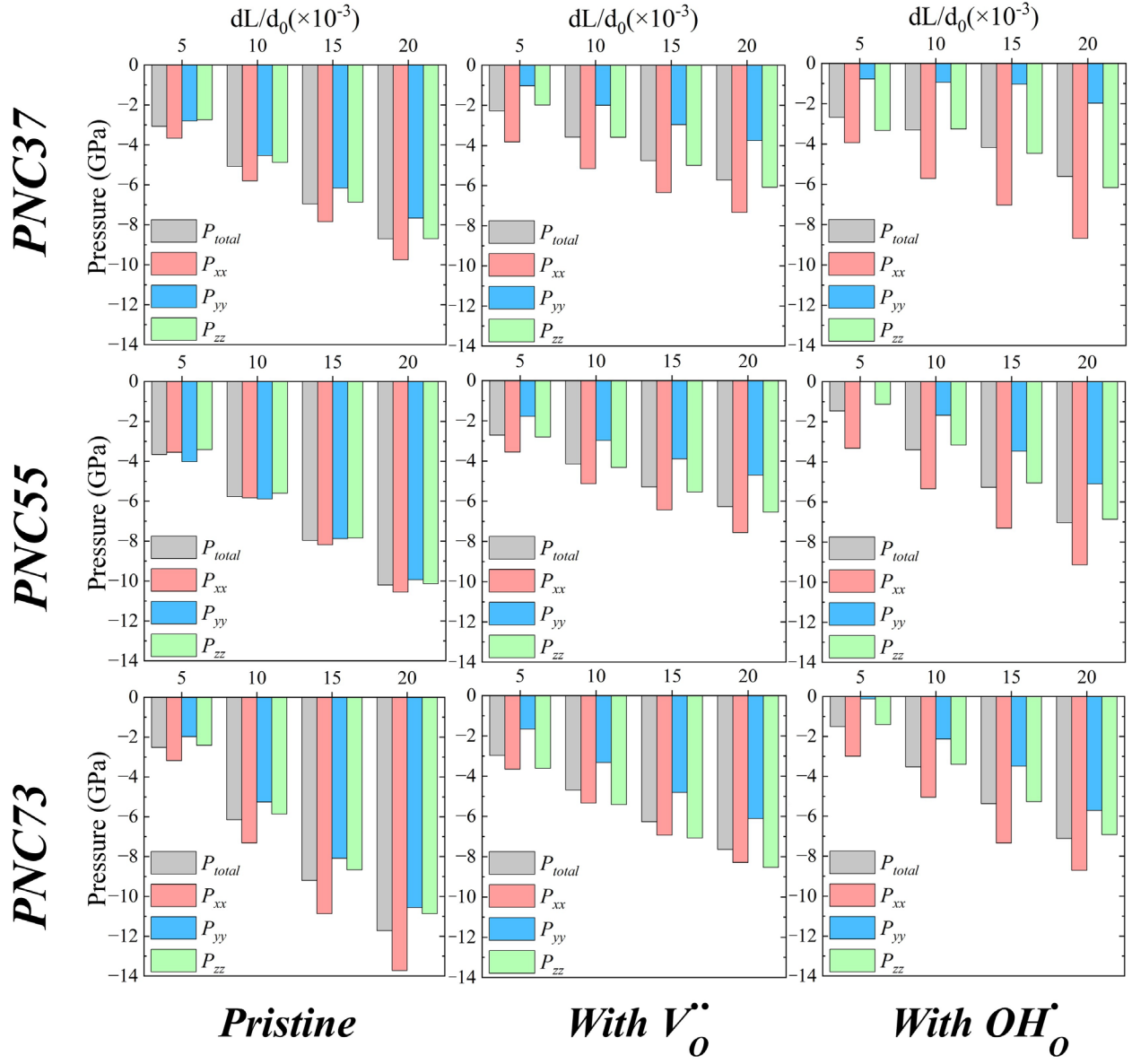
Appendix Figure C.4 Equations of state (EOS) of pristine and defective PNC unit cells, PNC37 (top), PNC55 (middle), and PNC73 (bottom), based on MLP (solid red squares) and DFT (hollow black dots).



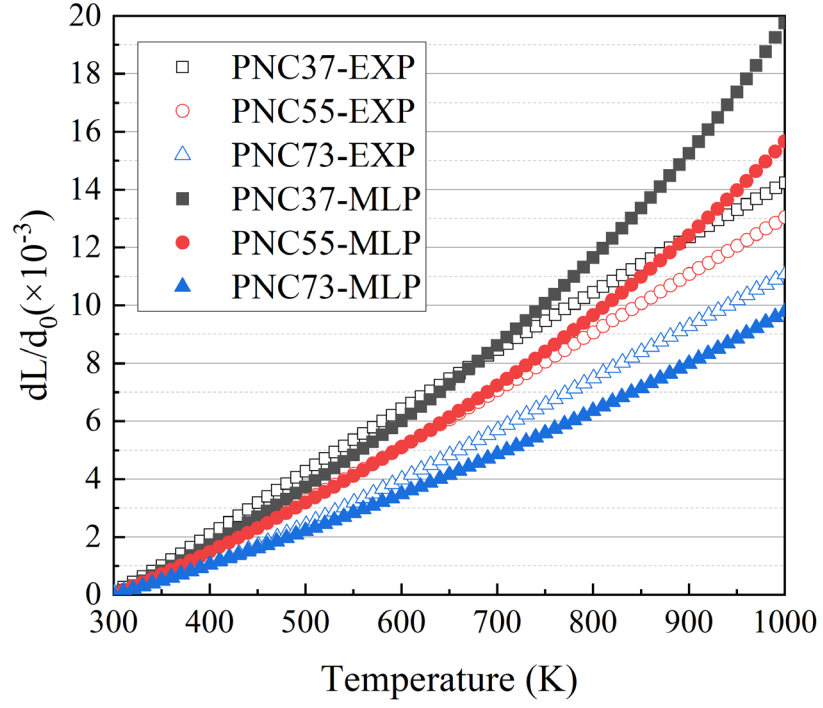
Appendix Figure C.5 Phonon dispersion and phonon density of states (PDOS) of PNC55 calculated from DFT (blue) and MLP (red) calculations.



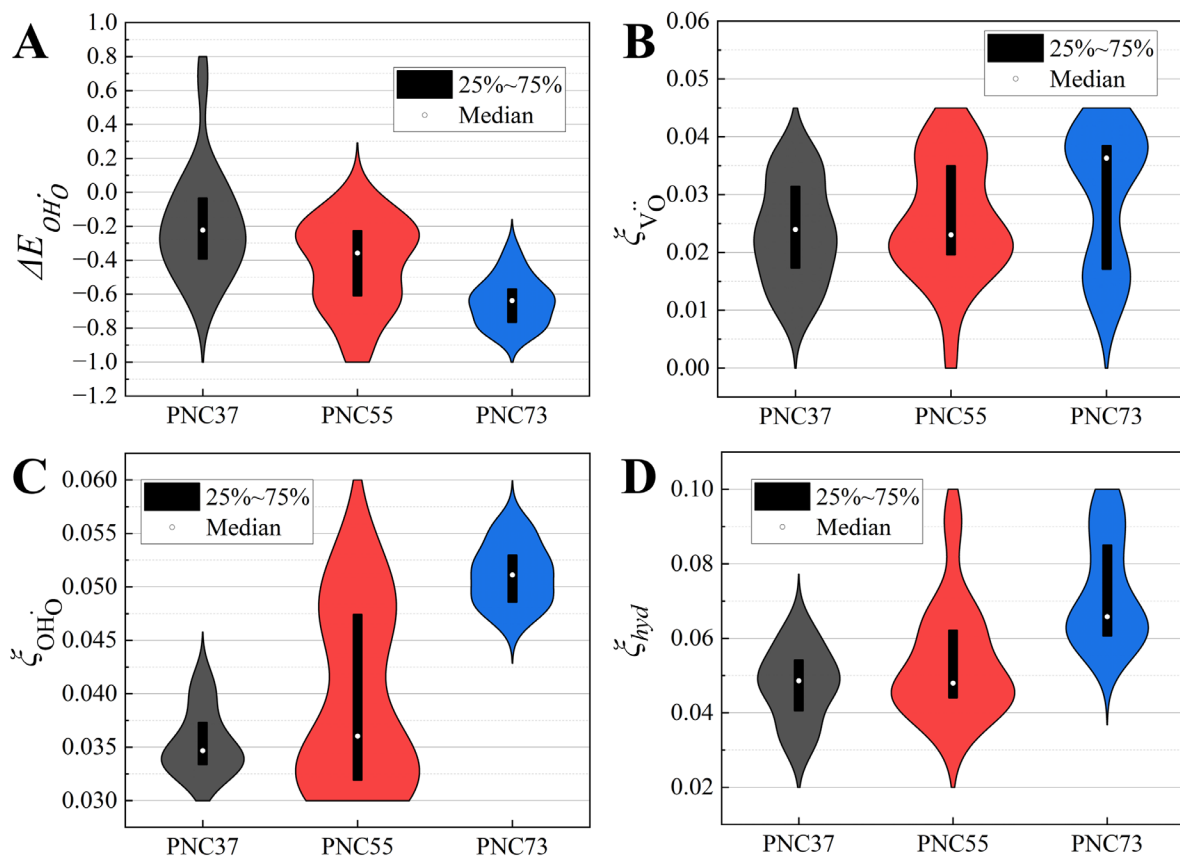
Appendix Figure C.6 Equation of state (EOS) of pristine and defective (A) PNC37, (B) PNC55, and (C) PNC73, adjusted with the same volume scaling factor.



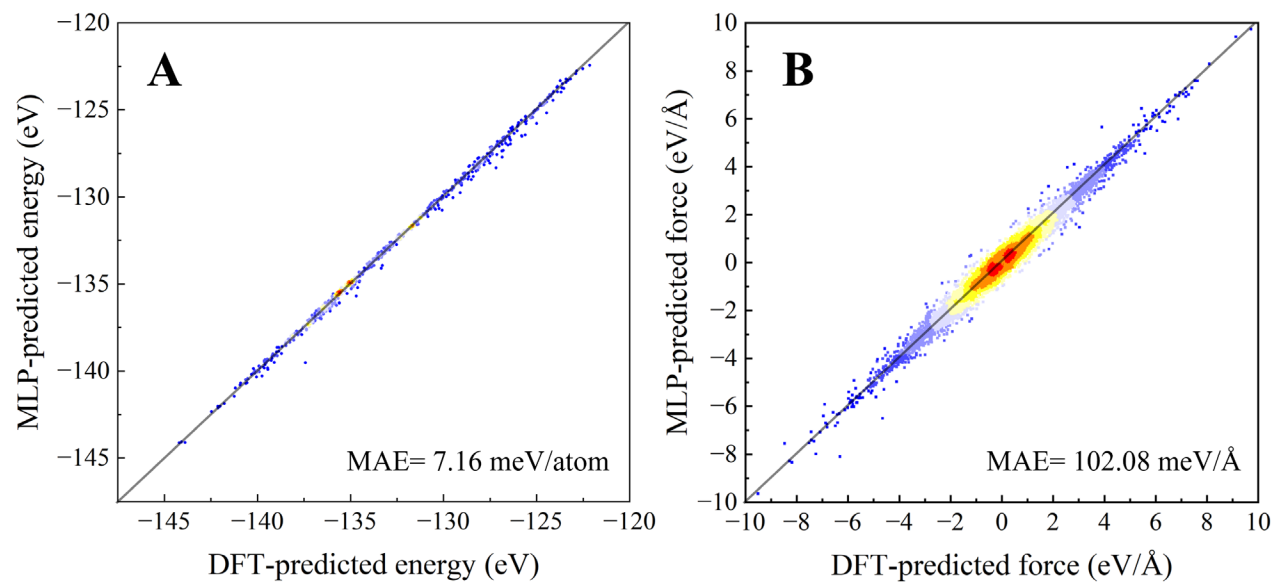
Appendix Figure C.7 Pressures (P_{total}) and their normal components (P_{xx} , P_{yy} , and P_{zz}) of the PNCs with and without defects under expansion (dL/d_0) of 5×10^{-3} to 20×10^{-3} .



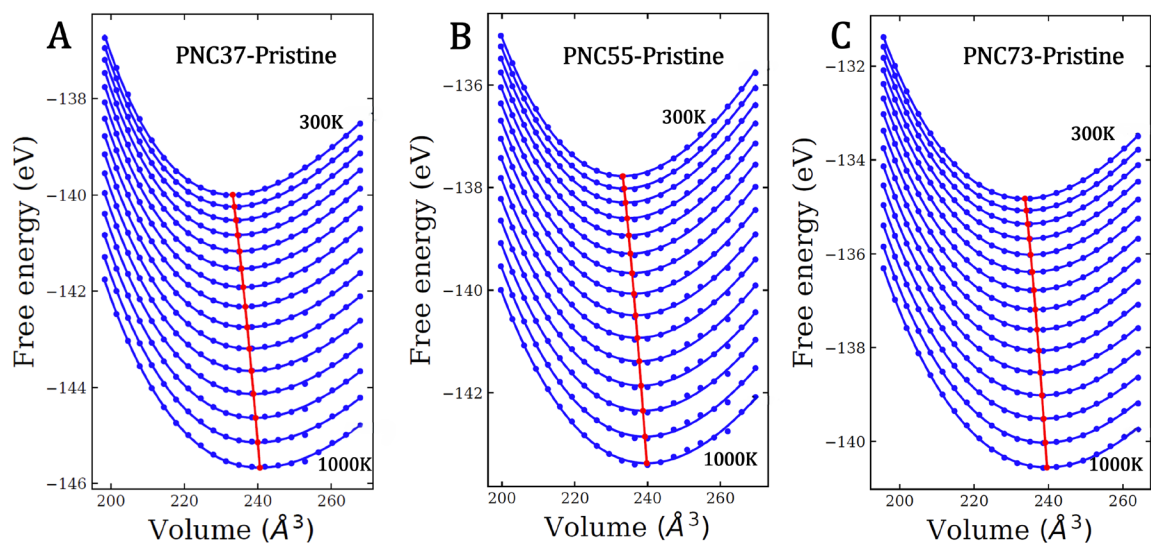
Appendix Figure C.8 Thermal expansions predicted by MLP (solid symbols, based on unit cell with one $V_{\text{O}}^{\bullet}$) and comparisons with experiments ⁷⁰ (hollow symbols).



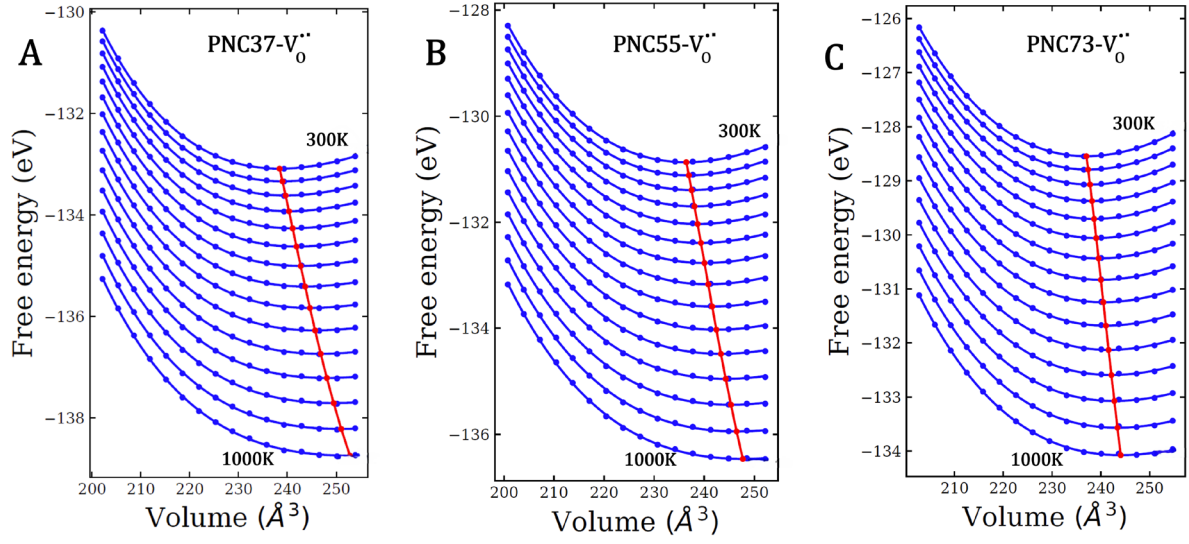
Appendix Figure C.9 (A) OH_O formation energy (ΔE_{OH_O}), and chemical expansion (ξ) induced by (B) V_O , (C) OH_O , and (D) hydration, for PNC37 (grey), PNC55 (red), and PNC73 (blue), from DFT calculations.



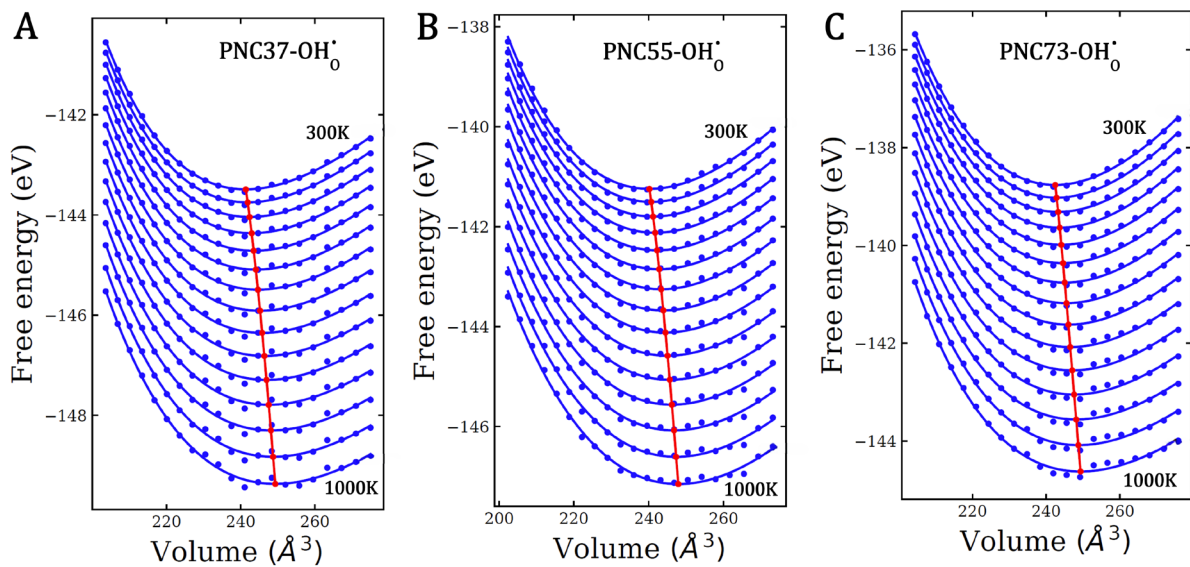
Appendix Figure C.10 Parity plots for the predicted energies and forces based on the testing set (MAE: mean absolute error). Red color indicates the highest density.



Appendix Figure C.11 Free energies of (A) pristine PNC37, (B) pristine PNC55, and (C) pristine PNC73, as a function of cell volume at 300-1000K (with a temperature interval of 50K). The red lines indicate the locations of the energy minima.



Appendix Figure C.12 Free energies of PNC consisting of a V_0'' , (A) PNC37, (B) PNC55, and (C) PNC73, as a function of the cell volume at 300-1000K (with a temperature interval of 50K). The red lines indicate the locations of the energy minima.



Appendix Figure C.13 Free energies of PNC consisting of a OH_0 , (A) PNC37, (B) PNC55, and (C) PNC73, as a function of the cell volume at 300-1000K (with a temperature interval of 50K). The red lines indicate the locations of the energy minima.

Appendix D - Uncovering the Atomic Origins of Lattice Expansion and Electrode Incompatibility in Y-doped BaZrO₃ Electrolytes

Materials synthesis methods

BaZr_{0.8}Y_{0.2}O_{3-δ} (BZY20) and BaZr_{0.1}Ce_{0.7}Y_{0.1}Yb_{0.1}O_{3-δ} (BZCYYb1711) were prepared by solid state reaction method. The stoichiometric amounts of barium carbonate (99%, Sigma Aldrich), zirconium oxide (99.5%, Alfa Aesar), cerium oxide (99.9%, Alfa Aesar), yttrium oxide (99.9%, Alfa Aesar), and ytterbium oxide (99.9%, Alfa Aesar) were mixed by ball-milling in ethanol for 24 hours with yttria-stabilized zirconia balls. After heating the mixture with a magnetic stirred hot plate to remove the ethanol, the powder was lightly grounded and then calcined at 1100 °C for 10 hours. The milling and drying steps were repeated a second time, and then the powders were calcined again at 1350 °C for 10 h to ensure the phase formation. The milling and drying steps were repeated a third time to get fine particles for tape casting.

Oxygen electrode material PrNi_{0.7}Co_{0.3}O_{3-δ} (PNC73) was synthesized by a glycine-citrate combustion method. Stoichiometric amounts of Pr(NO₃)₃·6H₂O (99.99% purity, Alfa Aesar), Ni(NO₃)₂·6H₂O (99.9985% purity, Alfa Aesar) and Co(NO₃)₂·6H₂O (99.999% purity, Alfa Aesar) were dissolved in deionized water to form a homogeneous solution. Glycine and citric acid were next added into the solution in a mole ratio of glycine: citric acid: metal ions of 2:1:1. The as-prepared solution was heated and stirred on a magnetic stirrer until a gel was formed, following an auto-ignition process to produce the solid ash. The final PrNi_{0.7}Co_{0.3}O_{3-δ} powders were obtained by calcining the ash at 1000 °C for 4 h.

PNC73-BZY20 cell fabrication

The hydrogen electrode-supported BZY20 half-cell was fabricated by a tape-casting method. The detailed experimental conditions can be found in our previous work ²²⁴. Green tapes of NiO-BZCYYb1711 (6:4) hydrogen electrode supporting layer, NiO-BZY20 (6:4) buffer layer, and BZY20 electrolyte layer were first fabricated by a laboratory tape casting machine, respectively. Next, hydrogen electrode tape, buffer layer tape, and electrolyte tape were pressed together by the hot lamination method at 70°C under an applied pressure of 3 ton for 5 h to obtain green half-cell tape. The raw half-cell was cut from the green half-cell tape with a diameter of 11 mm and pre-sintered at 920°C for 3 hours to remove all organic binder/solvent. Finally, the half-cell was co-sintered at about 1480°C for 15 hours to densify the electrolyte membrane.

To fabricate NiO+BZCYYb1711/NiO+BZY20/BZY20/PNC73 full cells, oxygen electrode PNC73 slurry was prepared using ethanol and Thinner (a texanol-based binder). Firstly, mix PNC73 powder with ethanol and then run high energy ball milling process for 10 min to get finer particles. Next, add a certain proportion of Thinner to the mixture and apply THINKY MIXER to obtain the final PNC73 slurry and brush-painting it to the BZY20 electrolyte surface of the co-sintered half-cell. The full cell was successfully fabricated after a heat treatment process at 1000°C for 2h.

Calculation of defect formation energy (ΔE) and chemical expansion (ε):

The defect formation energy (ΔE_i) due to the formation of defect i (V_O or OH) was calculated by:

$$\Delta E_i = E_i + \frac{1}{2}E_{gas} - E_{bulk} \quad (1)$$

where E_i , E_{gas} , and E_{bulk} are the total energies of the unit cell with a vacancy i , the corresponding gas molecule (O_2 or H_2), and the pristine unit cell, respectively.

The chemical expansion (ε_i) due to the formation of the defect i (V_O or OH) can be expressed as follows:

$$\varepsilon_i = \frac{1}{\delta_i} \left(\frac{l_i(T) - l_0(T)}{l_0(T)} \right) \quad (2)$$

where δ_i is the number of defects per formula unit (i.e., 1/8, due to the single defect formation in the $2 \times 2 \times 2$ supercell), $l_i(T)$ and $l_0(T)$ are the dimensions of respective defective and pristine unit cells¹⁹⁹. And the chemical expansion due to hydration (ε_{hyd}) can be estimated by combining the effect of filling V_O and forming OH :

$$\varepsilon_{hyd} = 2\varepsilon_{OH} - \varepsilon_{V_O} \quad (3)$$

Appendix Table D.1 OTF-MD parameters for dataset generation based on 2x2x2 supercell

	BZO			1Y-sub BZY			2Y-sub BZY		
	Pristine	With V_o	With OH	Pristine	With V_o	With OH	Pristine	With V_o	With OH
Uncertainty threshold (eV/Å)	0.01	0.03	0.3	0.01	0.12	0.36	0.01	0.12	0.35
Time step	2 fs								
Number of steps	10,000								
Ensemble	NVT								
Temperature	1300 K								
Kernel	Normalized Dot Product								
Power of kernel	2								
Descriptor	B2 ACE descriptor								
Cutoff radius	4.0 Å								
n_{max}	8								
l_{max}	3								

Appendix Table D.2 Parameters used in final offline training using Allegro

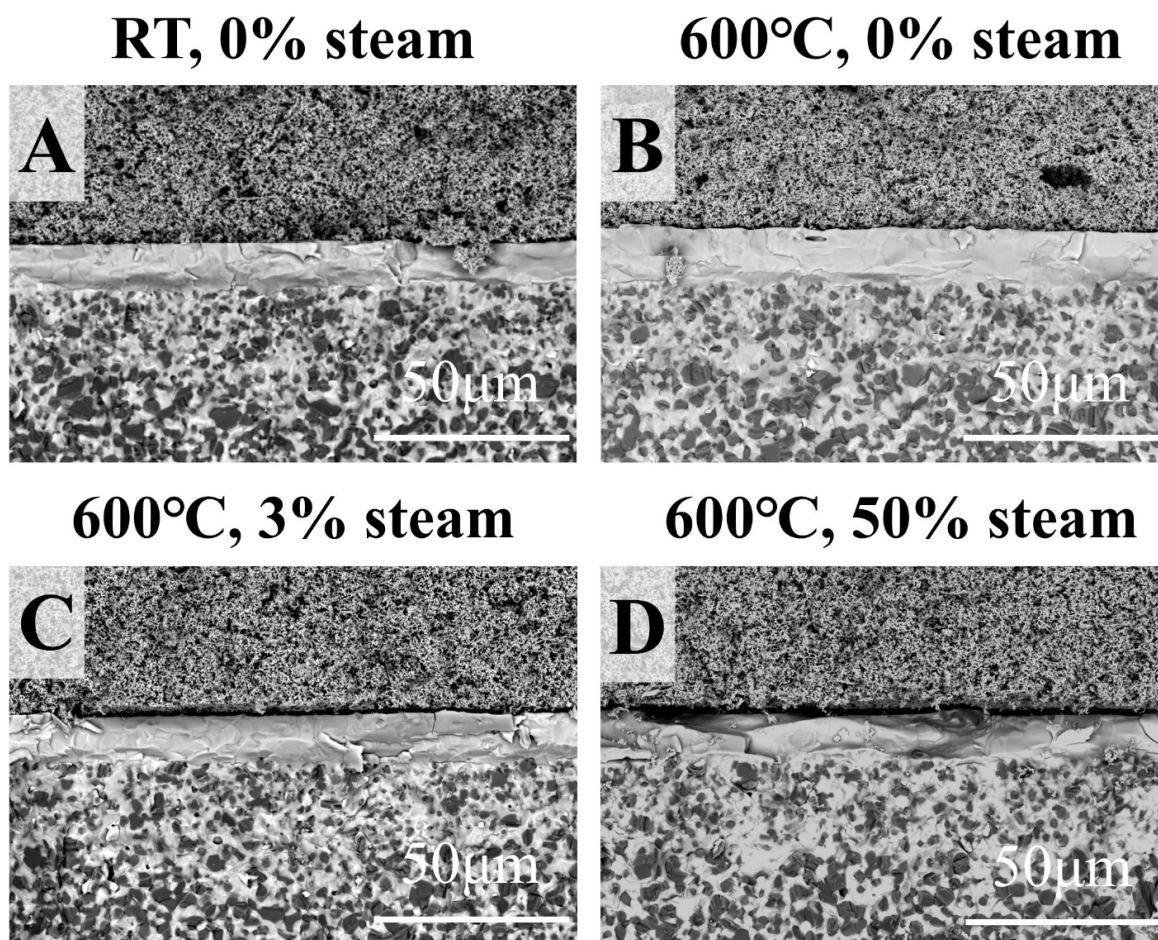
Cutoff radius	5.0 Å
Number of Bessel radial basis	8
Polynomial cutoff	48
Maximum order of spherical harmonic embedding	3
Parity	o3_full
Number of tensor product layers	2
Number of features	16
2-body latent multi-layer perceptron dimension	[64, 128, 256]
Later latent multi-layer perceptron dimension	[256, 256, 256]
Final edge energy multi-layer perceptron dimension	[64]
Learning rate	0.001
Batch size	1

Appendix Table D.3 Lattice constants of pristine and V_O -defective $2\times 2\times 2$ supercells, and formation energy (ΔE_i) or chemical expansion (ϵ_i) due to the defect i (V_O or OH) formation at distinct sites of 1Y-sub and 2Y-sub BZY, as shown in Figure 6.4.

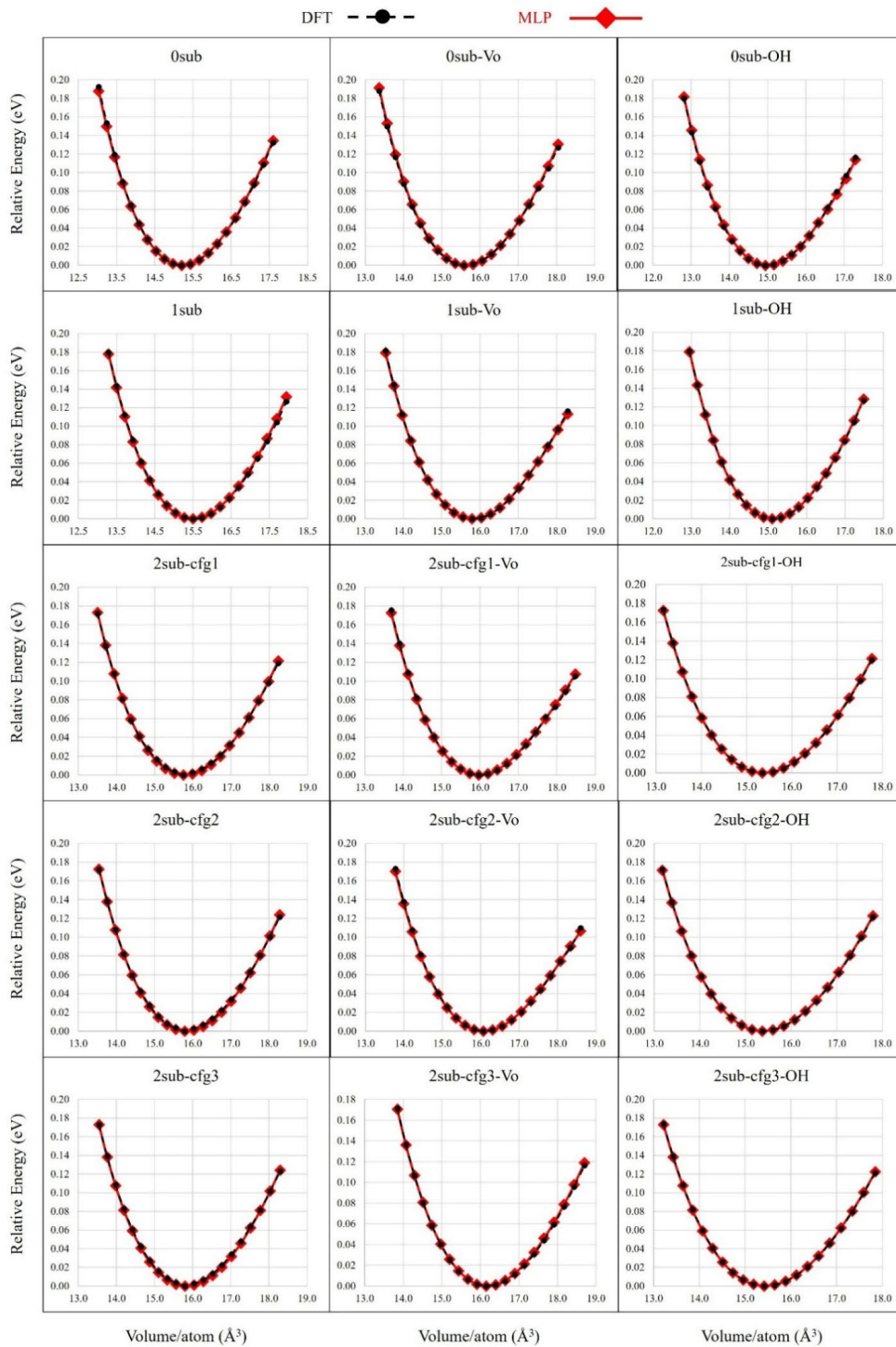
Structures	Defect position	$a(\text{\AA})$	$b(\text{\AA})$	$c(\text{\AA})$	$D_{def-Y}(\text{\AA})$	$\Delta E_{V_O}(\text{eV})$	$\Delta E_{OH}(\text{eV})$	ϵ_{V_O}	ϵ_{OH}
1Y-sub BZY	Pristine	8.528	8.527	8.527	N/A	N/A	N/A	N/A	N/A
	O ₁	8.520	8.520	8.485	2.21	3.37	-1.42	-1.77×10^{-2}	-3.52×10^{-3}
	O ₂	8.547	8.519	8.498	4.77	3.60	-1.33	-5.66×10^{-3}	2.12×10^{-3}
	O ₃	8.542	8.512	8.512	6.40	3.73	-1.20	-4.99×10^{-3}	4.46×10^{-3}
2Y-sub-1 BZY	Pristine	8.553	8.582	8.583	N/A	N/A	N/A	N/A	N/A
	O ₁	8.496	8.561	8.561	2.14	0.36	-1.72	-3.11×10^{-2}	-2.28×10^{-3}
	O ₂	8.505	8.579	8.589	3.49	0.76	-1.64	-1.43×10^{-2}	-3.31×10^{-3}
	O ₃	8.664	8.471	8.595	4.89	0.99	-1.58	2.99×10^{-3}	7.02×10^{-4}
	O ₄	8.569	8.495	8.669	5.63	1.19	-1.47	4.20×10^{-3}	2.63×10^{-3}
	O ₅	8.636	8.549	8.549	6.43	1.54	-1.32	4.85×10^{-3}	5.56×10^{-3}
2Y-sub-2 BZY	Pristine	8.580	8.579	8.583	N/A	N/A	N/A	N/A	N/A
	O ₁	8.578	8.523	8.571	3.49	0.61	-1.65	-2.19×10^{-2}	-4.77×10^{-3}
	O ₂	8.546	8.546	8.578	4.34	0.93	-1.53	-2.23×10^{-2}	-3.78×10^{-3}
	O ₃	8.540	8.540	8.641	4.81	0.92	-1.58	-6.87×10^{-3}	1.31×10^{-3}
	O ₄	8.574	8.646	8.503	5.62	1.08	-1.48	-5.91×10^{-3}	1.20×10^{-3}
2Y-sub-3 BZY	Pristine	8.582	8.582	8.582	N/A	N/A	N/A	N/A	N/A
	O ₁	8.584	8.545	8.545	4.32	0.87	-1.53	-2.25×10^{-2}	-3.91×10^{-3}
	O ₂	8.533	8.653	8.533	4.80	0.71	-1.57	-8.79×10^{-3}	3.46×10^{-4}

Appendix Table D.4 Linear fitting parameters of relationships between averaged oxygen-yttrium distance (D_{O-Y}) and defect formation energy (ΔE) or chemical expansion (ϵ) in $2\times 2\times 2$ supercells, as shown in Figure 6.4.

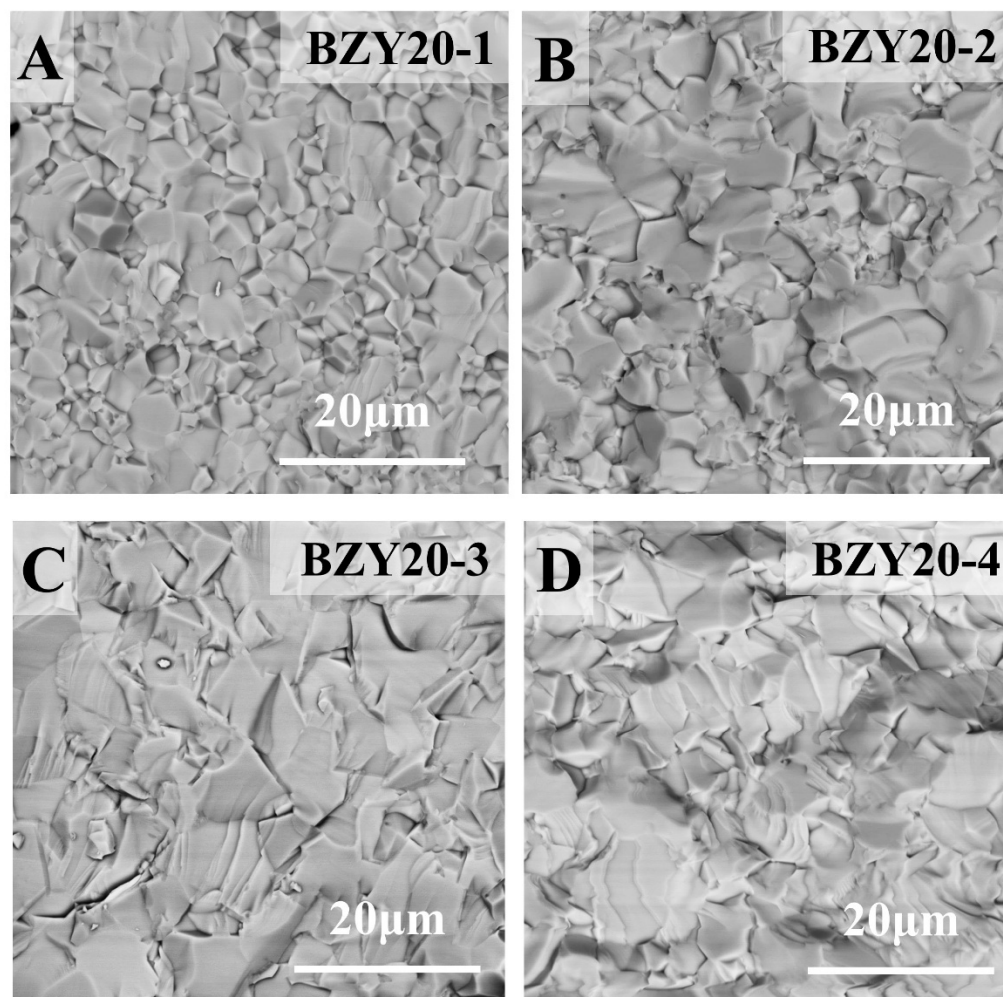
Linear relationship: $y = ax + b$	Slope (a)	Intercept (b)	R^2
1Y-sub ΔE_{Vo}	0.0867	3.1796	0.9990
1Y-sub ϵ_{Vo}	0.0032	-0.0237	0.8827
1Y-sub ΔE_{OH}	0.0525	-1.5498	0.9459
1Y-sub ϵ_{OH}	0.0019	-0.0076	0.9887
1Y-sub ϵ_{hyd}	0.0007	0.0085	0.4415
2Y-sub ΔE_{Vo}	0.2426	-0.1969	0.8670
2Y-sub ϵ_{Vo}	0.0089	-0.0514	0.7391
2Y-sub ΔE_{OH}	0.0814	-1.9230	0.8412
2Y-sub ϵ_{OH}	0.0022	-0.0104	0.6317
2Y-sub ϵ_{hyd}	-0.0045	0.0305	0.5230



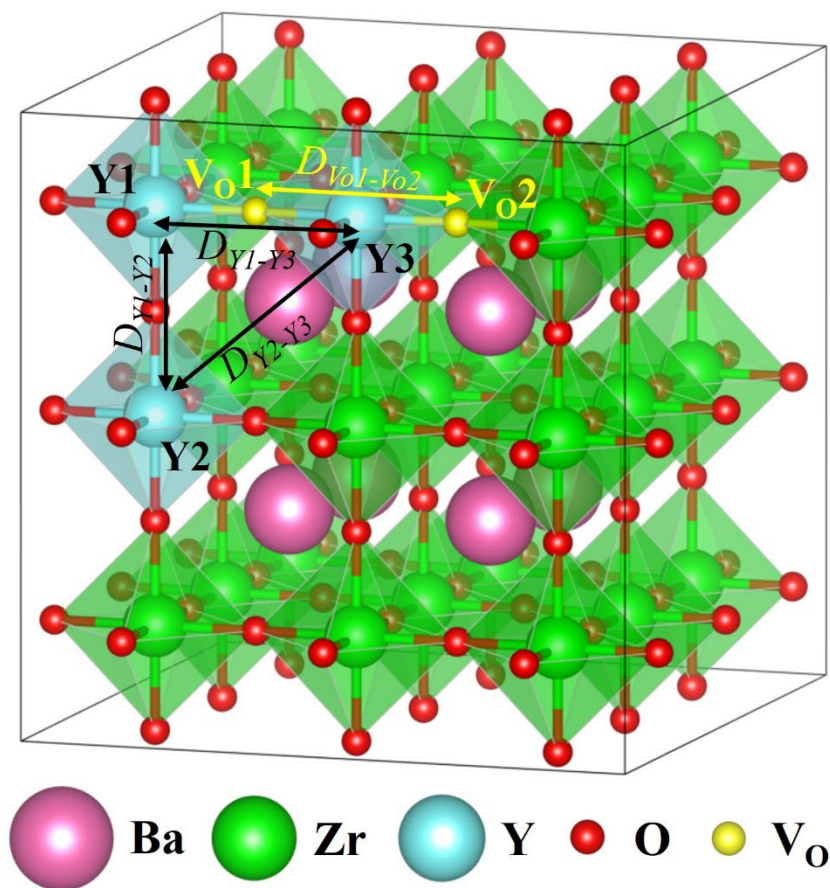
Appendix Figure D.1 Cross-sectional SEM images with amplify magnitude of 2000 of PNC73-BZY20 cells treated under various conditions, corresponding to Figure 1E to 1H.



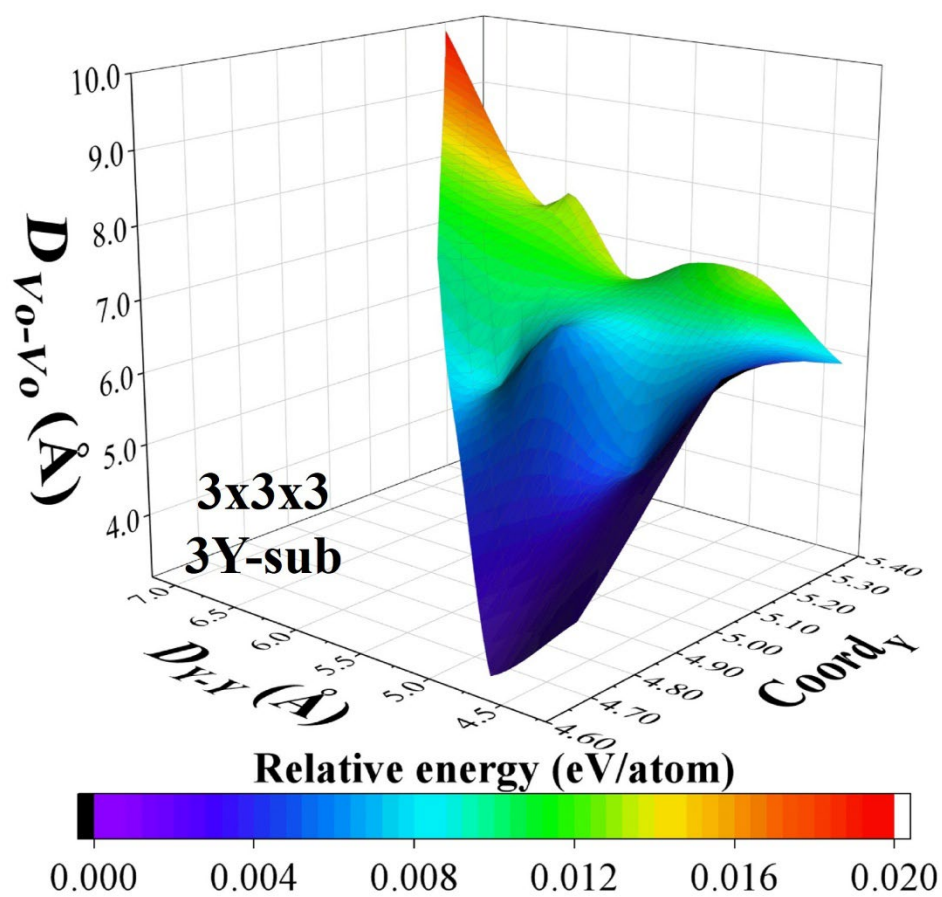
Appendix Figure D.2 Comparison of the equation of state (EOS) of 2x2x2 BZO and BZY supercells between DFT and MLP.



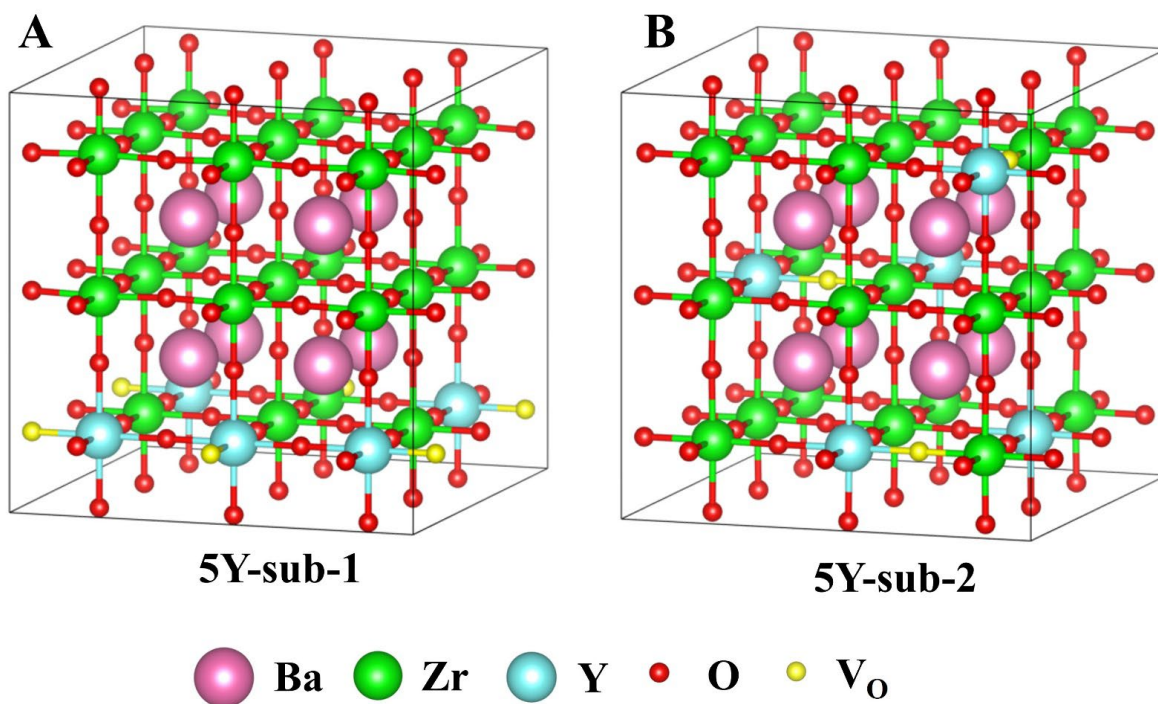
Appendix Figure D.3 SEM images of BZY samples.



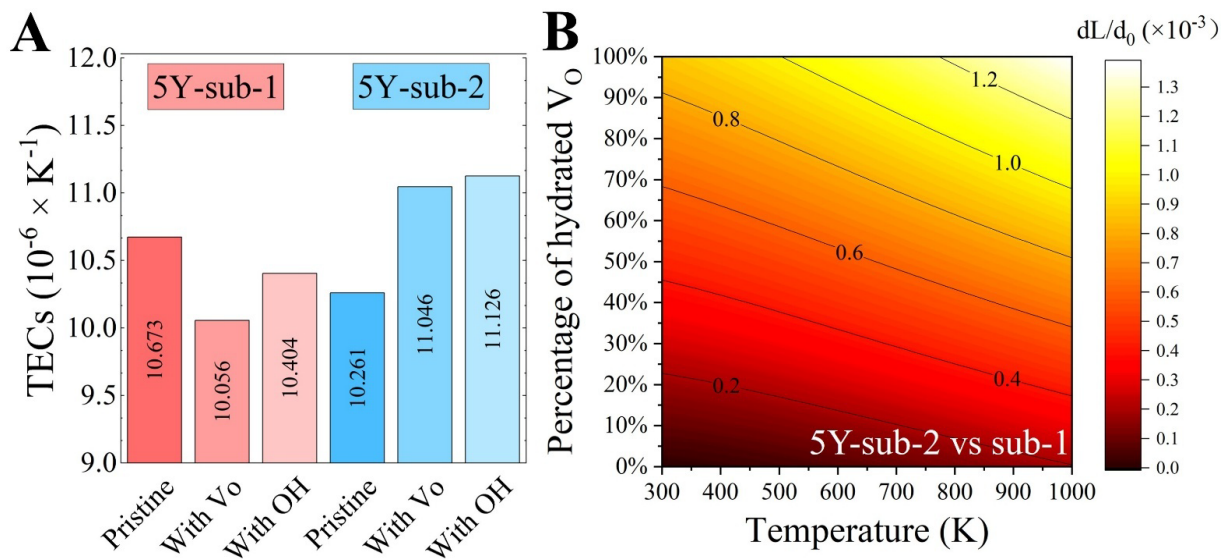
Appendix Figure D.4 Schematic illustration of determining descriptors including averaged V_O-V_O distance ($D_{V_{O}-V_{O}}$), averaged Y-Y distance (D_{Y-Y}), and averaged Y coordination ($Coord_Y$) number. as described in Eqn. 6.2-6.4.



Appendix Figure D.5 Three-dimensional colormapped surface illustrating energies of 3Y-sub $3 \times 3 \times 3$ BZY supercells



Appendix Figure D.6 Scheme of (A) 5Y-sub-1 and (B) 5Y-sub-2 $3\times3\times3$ supercells.



Appendix Figure D.7 (A) TEC comparison between 5Y-sub-1 and 5Y-sub-2, and (B) Lattice expansion difference between 5Y-sub-1 and 5Y-sub-2.

Appendix E - Reprint Permissions



Taylor & Francis
Taylor & Francis Group

A high-dimensional neural network potential for molecular dynamics simulations of condensed phase nickel and phase transitions

Author: Hao Deng, Jeffrey Comer, Bin Liu, et al

Publication: Molecular Simulation

Publisher: Taylor & Francis

Date: Dec 15, 2022

Rights managed by Taylor & Francis

Thesis/Dissertation Reuse Request

Taylor & Francis will be pleased to grant permission at no cost for use of content in a Master's Thesis and/or Doctoral Dissertation, subject to the following limitations.

You may use a single excerpt, up to 400 words and up to 3 figures/tables, on the condition that your acknowledgement must be included showing article title, author, full Journal title, and © copyright # [year], reprinted by permission of Informa UK Limited, trading as Taylor & Francis Group, <https://www.tandfonline.com>

If you use more than those limits, or intend to distribute or sell your Master's Thesis/Doctoral Dissertation to the general public through print or website publication, further permission will be required.

BACK

CLOSE



Predictions of Boron Phase Stability Using an Efficient Bayesian Machine Learning Interatomic Potential

Author: Hao Deng, Bin Liu

Publication: Journal of Physical Chemistry Letters

Publisher: American Chemical Society

Date: Mar 1, 2024

Copyright © 2024, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms and Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from {COMPLETE REFERENCE CITATION}. Copyright {YEAR} American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.

If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.

BACK

CLOSE WINDOW



Atomistic Simulations of Thermal and Chemical Expansions of PrNi₃Co_{1-x}O_{3-δ} Accelerated by Machine Learning Potentials

Author: Hao Deng, Quanwen Sun, Meng Li, et al

Publication: Small Methods

Publisher: John Wiley and Sons

Date: Jul 9, 2025

© 2025 Wiley-VCH GmbH

Order Completed

Thank you for your order.

This Agreement between Kansas State University ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

License Number 6064860880625

[Printable Details](#)

License date Jul 09, 2025

✓ Licensed Content

Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Small Methods
Licensed Content Title	Atomistic Simulations of Thermal and Chemical Expansions of PrNi ₃ Co _{1-x} O _{3-δ} Accelerated by Machine Learning Potentials
Licensed Content Author	Hao Deng, Quanwen Sun, Meng Li, et al
Licensed Content Date	Jul 9, 2025
Licensed Content Volume	0
Licensed Content Issue	0
Licensed Content Pages	9

📁 Order Details

Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Electronic
Portion	Full article
Will you be translating?	No

📁 About Your Work

Title of new work	Understanding Material Properties and Performance Enabled by Molecular Simulations and Machine Learning Potentials
Institution name	Kansas State University
Expected presentation date	Jul 2025

📁 Additional Data

The Requesting Person / Organization to Appear on the License	Kansas State University
---	-------------------------