Supporting Information

Machine Learning-Driven Global Optimization Reveals Nanometre-Scale Mixed Phases of Borophene on Ag(100)

Yunlei Wang, Haifeng Lv, Xiaojun Wu**

State Key Laboratory of Precision and Intelligent Chemistry, School of Chemistry and Materials Sciences, Key Laboratory of Materials Sciences for Energy Conversion, Collaborative Innovation Center of Chemistry for Energy Materials (iChEM), and CAS Center for Excellence in Nanoscience, University of Science and Technology of China, Hefei, Anhui 230026, China.

KEYWORDS: borophene, mixed phase, large-scale global optimization, metal substrate, neutral network potential.

Corresponding Author

*Email: hflv@ustc.edu.cn, xjwu@ustc.edu.cn

Materials and method

First-principles calculations

All first-principles calculations were performed using the Vienna Ab initio Simulation Package (VASP)[1, 2] with the projector augmented wave (PAW)[3, 4] method to describe ionic cores and valence electrons. The exchange-correlation functional is treated within the generalized gradient approximation (GGA), in the Perdew-Burke-Ernzerhof (PBE) functional[5, 6]. The D3 correction[7] was employed to describe the van der Waals interaction between substrates and borophenes. The kinetic energy cutoff for plane wave basis is 400 eV and the first Brillouin zone is sampled using the Γ-centered Monkhorst-Pack scheme with a *k*-space density of $2\pi \times 0.04$ Å^{-1}.

Neural network (NN) potential

The high dimensional neural network (HDNN) scheme in LASP software is employed to train the global data set. 331 power-type structural descriptors (PTSDs) are used to discriminate atomic structures during exploring the PES. The neural network architecture for Ag and B element is 177-80-80-80-6 in which 80-80-80 are three hidden layers, and there are 55452 learnable parameters in total. To represent the complex atomic environment of boron, the cut-off radius in NN potential is set from 1.20 Å to 7.60 Å.

Performance of NN potential

The energy convergence criterion is $\Delta E \leq 5$ meV/atom, where ΔE is defined as

$$\Delta E = \left|E_{NN} - E_{DFT}\right|/N \#(1)$$

Here, E_{NN} and E_{DFT} denote the energies calculated by the NN potential and DFT method, respectively, and N is the total number of atoms. If convergence is not achieved in the first iteration for a system, the next iteration starts from the minima obtained in the previous iteration, not from random configurations.

Figure S3a shows the convergence trend of structures with respect to the iteration number, indicating that all the structures, regardless of *M* values, achieve energy convergence criterion within 7 iterative steps. In the first iterative step, the ratio of converged structures increases with *M* from 6 to 30. Additionally, the initial attempt with the NN potential on the system of $M = 6$ converges relatively slower, but the convergence rate improves as *M* increases. This demonstrates the growing predictive power of NN potential for large systems as the global dataset continues to expand iteratively.

To further investigate the performance of this NN potential, we apply the model to a test dataset composed of 30,500 data. 61 representative systems are selected from supercells with $140° \leq \gamma \leq 160°$ and $20° \leq \gamma \leq 40°$ under *BC* ranging from 2.3 to 7.3, as listed in **Table S2**. The *M* values of supercells in test dataset are set to 10, 15, 20, and 30. We randomly collect 500 structures from SSW-NN trajectories for each system. The energy and force error ($\Delta E'$ and ΔF) distribution for the validation dataset is shown in **Figure S3b-c**. ΔE' is defined as

$$\Delta E^{'} = \left|E_{NN} - E_{DFT}\right|/N \#(2)$$

where E_{NN} and E_{DFT} denote energies calculated by the NN potential and DFT method, respectively, and N is the total number of atoms. ΔF denotes the absolute error of force between results obtained by DFT methods and the NN potential, defined as

$$\Delta F_j = \frac{\sum_{i}^{N_j^{'}} \left(\left|F_{x,i}^{DFT} - F_{x,i}^{NN}\right| + \left|F_{y,i}^{NN} - F_{y,i}^{NN}\right| + \left|F_{z,i}^{DFT} - F_{z,i}^{NN}\right|\right)}{3N_j^{'}} \#(3)$$

where $F_{x,i}$, $F_{y,i}$, $F_{z,i}$ represent the x, y, and z components of the force acting on atom i for structure j, containing $N_j^{'}$ relaxed atoms [8]. The energy and force error ranges are -0.05 to 0.05 eV/atom and 0.1 to 0.6 eV/Å, respectively, and are mainly distributed at -0.02 to 0.02 eV/atom and 0.2 to 0.4 eV/Å, respectively. The mean absolute error (MAE) for energy and force is 4.060 meV/atom and 0.323 eV/Å, respectively.

Stochastic surface walking method

The stochastic surface walking (SSW) global optimization[9, 10] implemented in the Large-scale Atomistic Simulation with neural network Potential (LASP) software[11] was used to sample and search the global potential energy surface (PES). During the SSW sampling with first-principles calculations, the SSW steps is set to be 100. During the SSW search based on the potential, the geometry convergence criteria is satisfied when the maximal force component is less than 0.05 eV/Å and the SSW steps is set to be 150. The distance-weighted Steinhardt-type order parameter[12, 13] is

applied to indicate the averaged geometrical environment of bonded atoms in the SSW search, which is defined as

$$OP_L = \left(\frac{4\pi}{2L+1} \sum_{M=-L}^{L} \left| \frac{1}{N_{bonds}} \sum_{i \neq j} e^{-\frac{1}{2}\frac{r_{ij}-r_c}{r_c}} Y_{Lm}(r_{ij}) \right|^2 \right)^{\frac{1}{2}} \#(4)$$

where Y_{Lm} is the spherical harmonic function of degree l and order m; i and j are atoms in the lattice, r_{ij} is the vector between atoms i and j, r_{ij} is the distance between them, r_{c} is set as 60% of the typical single bond length between i and j atoms (1.7 Å here for boron-boron single bonds), and N_{bonds} is the number of bonds in the first bonding shell (2.21 Å) which is set as 130% of the typical single bond length.

High Dimensional Neural Network (HDNN) architecture

In this work, we utilized the high dimensional neural network (HDNN) scheme to construct the NN. The total energy E_{tot} can be decomposed and written as a linear combination of atomic energy E_{i}, which is the output of the standard neural network. The input nodes are a set of geometry based structural descriptors $\{S_{\mathrm{i}}\}$.

$$E^{tot} = \sum_i E_i \#(5)$$

The atomic force can be analytically derived according to Eq. 2, where the force component $F_{\mathrm{k},\alpha}$=x, y or z, acting on the atom k is the derivative of the total energy with respect to its coordinate $R_{\mathrm{k},\alpha}$. The force component can be further related to the derivatives of the atomic energy with respect to j^{th} structural descriptors of atom i, $S_{\mathrm{j,i}}$:

$$F_{k,\alpha} = -\frac{\partial E^{tot}}{\partial R_{k,\alpha}} = -\sum_{i,j} \frac{\partial E_i}{\partial G_{j,i}} \frac{\partial S_{j,i}}{\partial R_{k,\alpha}} \#(6)$$

Similarly, the static stress tensor matrix element $\sigma_{\alpha\beta}$ can be analytically derived as:

$$\sigma_{\alpha\beta} = -\frac{1}{V} \sum_{i,j,d} \frac{(r_d)_\alpha (r_d)_\beta}{r_d} \frac{\partial E_i}{\partial G_{j,i}} \frac{\partial S_{j,i}}{\partial r_d} \#(7)$$

where $\mathbf{r}_d$ and r_d are the distance vector constituting of $S_{j,i}$ and its module, respectively, and V is the volume of the structure.

Structural descriptors

We use power-type structural descriptors (PTSDs) for structure discrimination, as shown in Eq. 8-13. In the PTSDs, S1 and S2 are two-body functions, S3, S4 and S5 are three-body functions, and S6 is a four-body function.

$$S_i^1 = \sum_{j \neq i} R^n(r_{ij}) \#(8)$$

$$S_i^2 = \left[\sum_{m=-L}^{L} \left| \sum_{j \neq i} R^n(r_{ij}) Y_{Lm}(r_{ij}) \right|^2 \right]^{\frac{1}{2}} \#(9)$$

$$S_i^3 = 2^{1-\zeta} \sum_{j,k \neq i} (1 + \lambda \cos\theta_{ijk})^\zeta \cdot R^n(r_{ij}) \cdot R^m(r_{ik}) \cdot R^p(r_{jk}) \#(10)$$

$$S_i^4 = 2^{1-\zeta} \sum_{j,k \neq i} (1 + \lambda \cos\theta_{ijk})^\zeta \cdot R^n(r_{ij}) \cdot R^m(r_{ik}) \#(11)$$

$$S_i^5 = \left[\sum_{m=-L}^{L} \left| \sum_{j,k \neq i} R^n(r_{ij}) \cdot R^m(r_{ik}) \cdot R^p(r_{jk}) \cdot \left(Y_{Lm}(r_{ij}) + Y_{Lm}(r_{ik}) \right) \right|^2 \right]^{\frac{1}{2}} \#(12)$$

$$S_i^6 = 2^{1-\zeta} \sum_{j,k,l \neq i} (1 + \lambda \cos\delta_{ijkl})^\zeta \cdot R^n(r_{ij}) R^m(r_{ik}) R^p(r_{il}) \#(13)$$

Structural models

The Ag(100) substrate model contains two layers with the bottom layer fixed during the SSW optimization. In the transfer learning, the bottom two layers are fixed for Ag(110) and Ag(111) substrates containing three metal layers. The vacuum thickness is larger than 15 Å to prevent any spurious interactions. Note that the Ag(100) primitive cell includes two Ag atoms and thus the total atoms of a system N can be described as

$$N = (2 + BC) \times M \#(14)$$

in which BC and M denote the boron coverage and the multiple of supercell relative to primitive cell, respectively. The largest target system contains up to 279 atoms, given BC of 7.3 and M of 30. To evaluate the stability of borophene on Ag(100), formation energy (E_{form}) is used, which is defined as

$$E_{form} = \frac{E_{tot} - E_{sub}}{N_B} \#(15)$$

E_{tot} and E_{sub} denote the energy of total system and Ag(100) substrate, respectively, and N_{B} is the number of boron atoms.

Expansion of dataset and enhanced sampling

When the NN potential was applied to predict minima, some configurations were collected from the SSW trajectory and added to the dataset to retrain the NN potential for the poorly predicted minima. During SSW + NN simulations, several candidate minima structures were generated. Firstly, the highest and lowest energy values of all candidate minima structures were identified. The energy range between these two

values were divided into parts with an interval of 1 eV, and 0.25% of structures from the SSW trajectory in each interval were then selected (if the number is less than 1, take 1). This will ensure the enhanced sampling in the PES near the poorly predicted minima. Second, these selected data were cleaned before DFT-computed forces and energies to achieve that the energy difference between any two configurations must exceed 5 meV/atom. This will avoid over-representation of atomic environments, concurrent with accurate description of PES.

The SSW optimization algorithm is most efficient for systems with no more than atoms/cell. Note that largest systems in this work contains up to 249 relaxed atoms. Therefore, to ensure the efficiency of exploring configurational space, the minima structures obtained from small systems were sampled and then added to the dataset during sequential expanding on systems with *M* values ranging from 6 to 30.

Table S1. The elements of redefine matrix (r_{11}, r_{21}, r_{21}, r_{22}), lattice parameters (γ, a, b), Supercell type (referred as $r_{11}r_{21}r_{21}r_{22}$), and the determinant of redefine matrix (*M*) for supercells.

r_{11}	r_{12}	r_{21}	r_{22}	Supercell	a	b	γ	*M*
2	0	1	1	2011	5.778	4.086	45.0	2
1	1	-1	1	11-11	4.086	4.086	90.0	2
2	1	-1	1	21-11	6.460	4.086	108.4	3
2	2	0	2	2202	8.171	5.778	45.0	4
2	1	-2	1	21-21	6.460	6.460	126.9	4
2	1	0	2	2102	6.460	5.778	63.4	4
2	0	0	2	2002	5.778	5.778	90.0	4
3	1	1	2	3112	9.136	6.460	45.0	5
2	1	-1	2	21-12	6.460	6.460	90.0	5
4	1	-2	1	41-21	11.912	6.460	139.4	6
3	0	1	2	3012	8.667	6.460	63.4	6
3	0	2	2	3022	8.667	8.171	45.0	6
3	1	0	2	3102	9.136	5.778	71.6	6
3	2	0	2	3202	10.416	5.778	56.3	6
2	2	-1	2	22-12	8.171	6.460	71.6	6
3	0	0	2	3002	8.667	5.778	90.0	6
4	1	1	2	4112	11.912	6.460	49.4	7
3	2	-2	1	32-21	10.416	6.460	119.7	7
3	1	-1	2	31-12	9.136	6.460	98.1	7
4	2	-2	1	42-21	12.920	6.460	126.9	8
4	0	1	2	4012	11.556	6.460	63.4	8
4	0	2	2	4022	11.556	8.171	45.0	8
4	1	0	2	4102	11.912	5.778	76.0	8
4	2	0	2	4202	12.920	5.778	63.4	8
3	2	-1	2	32-12	10.416	6.460	82.9	8
3	1	-2	2	31-22	9.136	8.171	116.6	8
3	1	1	3	3113	9.136	9.136	53.1	8
2	2	-2	2	22-22	8.171	8.171	90.0	8
4	0	0	2	4002	11.556	5.778	90.0	8
5	1	1	2	5112	14.731	6.460	52.1	9
4	1	-1	2	41-12	11.912	6.460	102.5	9
3	3	-1	2	33-12	12.257	6.460	71.6	9
3	3	0	3	3303	12.257	8.667	45.0	9
3	2	-3	1	32-31	10.416	9.136	127.9	9
3	1	0	3	3103	9.136	8.667	71.6	9
3	2	0	3	3203	10.416	8.667	56.3	9
3	0	0	3	3003	8.667	8.667	90.0	9
5	0	1	2	5012	14.445	6.460	63.4	10
5	0	2	2	5022	14.445	8.171	45.0	10

5	1	0	2	5102	14.731	5.778	78.7	10
5	2	0	2	5202	15.558	5.778	68.2	10
4	2	-1	2	42-12	12.920	6.460	90.0	10
4	2	1	3	4213	12.920	9.136	45.0	10
4	1	-2	2	41-22	11.912	8.171	121.0	10
4	1	2	3	4123	11.912	10.416	42.3	10
3	2	-2	2	32-22	10.416	8.171	101.3	10
3	1	-1	3	31-13	9.136	9.136	90.0	10
5	0	0	2	5002	14.445	5.778	90.0	10
5	3	-2	1	53-21	16.846	6.460	122.5	11
5	2	-3	1	52-31	15.558	9.136	139.8	11
5	1	-1	2	51-12	14.731	6.460	105.3	11
4	3	-1	2	43-12	14.445	6.460	79.7	11
4	1	-3	2	41-32	11.912	10.416	132.3	11
4	1	1	3	4113	11.912	9.136	57.5	11
3	2	-1	3	32-13	10.416	9.136	74.7	11
6	0	1	2	6012	17.334	6.460	63.4	12
6	0	2	2	6022	17.334	8.171	45.0	12
6	1	0	2	6102	17.573	5.778	80.5	12
5	2	-1	2	52-12	15.558	6.460	94.8	12
5	1	-2	2	51-22	14.731	8.171	123.7	12
5	3	1	3	5313	16.846	9.136	40.6	12
4	4	-1	2	44-12	16.343	6.460	71.6	12
4	4	0	3	4403	16.343	8.667	45.0	12
4	2	-2	2	42-22	12.920	8.171	108.4	12
4	2	-4	1	42-41	12.920	11.912	139.4	12
4	0	1	3	4013	11.556	9.136	71.6	12
4	0	2	3	4023	11.556	10.416	56.3	12
4	0	3	3	4033	11.556	12.257	45.0	12
4	1	0	3	4103	11.912	8.667	76.0	12
4	2	0	3	4203	12.920	8.667	63.4	12
6	0	0	2	6002	17.334	5.778	90.0	12
3	3	-1	3	33-13	12.257	9.136	63.4	12
3	3	-2	2	33-22	12.257	8.171	90.0	12
3	2	-3	2	32-32	10.416	10.416	112.6	12
4	0	0	3	4003	11.556	8.667	90.0	12
6	1	-1	2	61-12	17.573	6.460	107.1	13
5	4	-2	1	54-21	18.499	6.460	114.8	13
5	3	-1	2	53-12	16.846	6.460	85.6	13
5	1	2	3	5123	14.731	10.416	45.0	13
5	2	1	3	5213	15.558	9.136	49.8	13
4	3	-3	1	43-31	14.445	9.136	124.7	13
4	1	-1	3	41-13	11.912	9.136	94.4	13
3	2	-2	3	32-23	10.416	10.416	90.0	13

7	0	1	2	7012	20.223	6.460	63.4	14
7	0	2	2	7022	20.223	8.171	45.0	14
6	2	-1	2	62-12	18.272	6.460	98.1	14
6	1	-2	2	61-22	17.573	8.171	125.5	14
5	4	-1	2	54-12	18.499	6.460	77.9	14
5	3	-3	1	53-31	16.846	9.136	130.6	14
5	2	-2	2	52-22	15.558	8.171	113.2	14
5	1	1	3	5113	14.731	9.136	60.3	14
4	3	-2	2	43-22	14.445	8.171	98.1	14
4	2	-1	3	42-13	12.920	9.136	81.9	14
4	2	-3	2	42-32	12.920	10.416	119.7	14
4	1	-2	3	41-23	11.912	10.416	109.7	14
4	2	1	4	4214	12.920	11.912	49.4	14
7	1	-1	2	71-12	20.428	6.460	108.4	15
6	3	-1	2	63-12	19.380	6.460	90.0	15
6	3	1	3	6313	19.380	9.136	45.0	15
6	1	-3	2	61-32	17.573	10.416	136.9	15
5	5	0	3	5503	20.428	8.667	45.0	15
5	0	1	3	5013	14.445	9.136	71.6	15
5	0	2	3	5023	14.445	10.416	56.3	15
5	0	3	3	5033	14.445	12.257	45.0	15
5	1	0	3	5103	14.731	8.667	78.7	15
5	2	0	3	5203	15.558	8.667	68.2	15
5	3	0	3	5303	16.846	8.667	59.0	15
5	4	0	3	5403	18.499	8.667	51.3	15
4	1	-3	3	41-33	11.912	12.257	121.0	15
4	1	1	4	4114	11.912	11.912	61.9	15
3	3	-2	3	33-23	12.257	10.416	78.7	15
5	0	0	3	5003	14.445	8.667	90.0	15
8	0	2	2	8022	23.112	8.171	45.0	16
7	3	-3	1	73-31	22.002	9.136	138.4	16
7	2	-1	2	72-12	21.032	6.460	100.6	16
7	1	-2	2	71-22	20.428	8.171	126.9	16
6	4	-1	2	64-12	20.833	6.460	82.9	16
6	2	-2	2	62-22	18.272	8.171	116.6	16
6	1	2	3	6123	17.573	10.416	46.9	16
6	2	1	3	6213	18.272	9.136	53.1	16
5	3	-2	2	53-22	16.846	8.171	104.0	16
5	2	-3	2	52-32	15.558	10.416	124.5	16
5	1	-1	3	51-13	14.731	9.136	97.1	16
5	2	2	4	5224	15.558	12.920	41.6	16
4	4	-1	3	44-13	16.343	9.136	63.4	16
4	4	-2	2	44-22	16.343	8.171	90.0	16
4	-4	0	4	4-404	16.343	11.556	135.0	16

4	3	-4	1	43-41	14.445	11.912	129.1	16
4	2	-2	3	42-23	12.920	10.416	97.1	16
4	2	-4	2	42-42	12.920	12.920	126.9	16
4	1	0	4	4104	11.912	11.556	76.0	16
4	2	0	4	4204	12.920	11.556	63.4	16
4	3	0	4	4304	14.445	11.556	53.1	16
4	0	0	4	4004	11.556	11.556	90.0	16
8	1	-1	2	81-12	23.292	6.460	109.4	17
7	3	-1	2	73-12	22.002	6.460	93.4	17
7	1	-3	2	71-32	20.428	10.416	138.2	17
7	2	2	3	7223	21.032	10.416	40.4	17
7	4	1	3	7413	23.292	9.136	41.8	17
6	5	-1	2	65-12	22.564	6.460	76.8	17
6	1	1	3	6113	17.573	9.136	62.1	17
5	4	-3	1	54-31	18.499	9.136	122.9	17
5	-3	-1	4	5-3-14	16.846	11.912	135.0	17
5	2	-1	3	52-13	15.558	9.136	86.6	17
5	1	-2	3	51-23	14.731	10.416	112.4	17
5	1	3	4	5134	14.731	14.445	41.8	17
4	3	-3	2	43-32	14.445	10.416	109.4	17
4	1	-1	4	41-14	11.912	11.912	90.0	17
9	0	2	2	9022	26.001	8.171	45.0	18
8	2	-1	2	82-12	23.823	6.460	102.5	18
8	1	-2	2	81-22	23.292	8.171	127.9	18
7	6	-3	0	76-30	26.635	8.667	139.4	18
7	4	-1	2	74-12	23.292	6.460	86.8	18
7	2	-2	2	72-22	21.032	8.171	119.1	18
7	3	1	3	7313	22.002	9.136	48.4	18
6	6	0	3	6603	24.514	8.667	45.0	18
6	4	-3	1	64-31	20.833	9.136	127.9	18
6	3	-2	2	63-22	19.380	8.171	108.4	18
6	3	-4	1	63-41	19.380	11.912	139.4	18
6	2	-3	2	62-32	18.272	10.416	127.9	18
6	0	1	3	6013	17.334	9.136	71.6	18
6	0	2	3	6023	17.334	10.416	56.3	18
6	0	3	3	6033	17.334	12.257	45.0	18
6	1	0	3	6103	17.573	8.667	80.5	18
6	2	0	3	6203	18.272	8.667	71.6	18
6	3	0	3	6303	19.380	8.667	63.4	18
6	4	0	3	6403	20.833	8.667	56.3	18
6	5	0	3	6503	22.564	8.667	50.2	18
5	4	-2	2	54-22	18.499	8.171	96.3	18
5	3	-1	3	53-13	16.846	9.136	77.5	18
5	2	-4	2	52-42	15.558	12.920	131.6	18

5	1	-3	3	51-33	14.731	12.257	123.7	18
5	1	2	4	5124	14.731	12.920	52.1	18
5	2	1	4	5214	15.558	11.912	54.2	18
4	3	-2	3	43-23	14.445	10.416	86.8	18
4	2	-1	4	42-14	12.920	11.912	77.5	18
4	2	-3	3	42-33	12.920	12.257	108.4	18
3	3	-3	3	33-33	12.257	12.257	90.0	18
6	0	0	3	6003	17.334	8.667	90.0	18
8	3	-1	2	83-12	24.684	6.460	96.0	19
8	1	-3	2	81-32	23.292	10.416	139.2	19
7	5	-1	2	75-12	24.852	6.460	81.0	19
7	4	-3	1	74-31	23.292	9.136	131.8	19
7	1	2	3	7123	20.428	10.416	48.2	19
7	2	1	3	7213	21.032	9.136	55.6	19
6	1	-1	3	61-13	17.573	9.136	99.0	19
5	4	-1	3	54-13	18.499	9.136	69.8	19
5	3	-3	2	53-32	16.846	10.416	115.4	19
5	2	-2	3	52-23	15.558	10.416	101.9	19
5	1	-4	3	51-43	14.731	14.445	131.8	19
5	1	1	4	5114	14.731	11.912	64.7	19
4	-3	1	4	4-314	14.445	11.912	112.8	19
9	1	-2	2	91-22	26.161	8.171	128.7	20
8	4	-1	2	84-12	25.840	6.460	90.0	20
8	4	1	3	8413	25.840	9.136	45.0	20
8	2	-2	2	82-22	23.823	8.171	121.0	20
8	2	2	3	8223	23.823	10.416	42.3	20
7	3	-2	2	73-22	22.002	8.171	111.8	20
7	2	-3	2	72-32	21.032	10.416	130.4	20
7	1	1	3	7113	20.428	9.136	63.4	20
6	4	-2	2	64-22	20.833	8.171	101.3	20
6	2	-1	3	62-13	18.272	9.136	90.0	20
6	2	2	4	6224	18.272	12.920	45.0	20
6	-1	2	3	6-123	17.573	10.416	65.8	20
6	4	1	4	6414	20.833	11.912	42.3	20
5	5	-2	2	55-22	20.428	8.171	90.0	20
5	5	0	4	5504	20.428	11.556	45.0	20
5	3	-5	1	53-51	16.846	14.731	137.7	20
5	2	-5	2	52-52	15.558	15.558	136.4	20
5	0	1	4	5014	14.445	11.912	76.0	20
5	0	2	4	5024	14.445	12.920	63.4	20
5	0	3	4	5034	14.445	14.445	53.1	20
5	0	4	4	5044	14.445	16.343	45.0	20
5	-1	0	4	5-104	14.731	11.556	101.3	20
5	2	0	4	5204	15.558	11.556	68.2	20

5	3	0	4	5304	16.846	11.556	59.0	20
5	4	0	4	5404	18.499	11.556	51.3	20
4	4	-1	4	44-14	16.343	11.912	59.0	20
4	4	-2	3	44-23	16.343	10.416	78.7	20
4	2	-2	4	42-24	12.920	12.920	90.0	20
5	0	0	4	5004	14.445	11.556	90.0	20
9	4	-3	1	94-31	28.453	9.136	137.6	21
9	3	-1	2	93-12	27.408	6.460	98.1	21
9	1	-3	2	91-32	26.161	10.416	140.0	21
8	5	-1	2	85-12	27.255	6.460	84.6	21
8	3	1	3	8313	24.684	9.136	51.0	21
7	7	0	3	7703	28.600	8.667	45.0	21
7	0	1	3	7013	20.223	9.136	71.6	21
7	0	2	3	7023	20.223	10.416	56.3	21
7	0	3	3	7033	20.223	12.257	45.0	21
7	1	0	3	7103	20.428	8.667	81.9	21
7	2	0	3	7203	21.032	8.667	74.1	21
7	3	0	3	7303	22.002	8.667	66.8	21
7	4	0	3	7403	23.292	8.667	60.3	21
7	5	0	3	7503	24.852	8.667	54.5	21
7	6	0	3	7603	26.635	8.667	49.4	21
6	5	-3	1	65-31	22.564	9.136	121.8	21
6	3	-1	3	63-13	19.380	9.136	81.9	21
6	3	-3	2	63-32	19.380	10.416	119.7	21
6	1	-3	3	61-33	17.573	12.257	125.5	21
6	1	3	4	6134	17.573	14.445	43.7	21
6	3	1	4	6314	19.380	11.912	49.4	21
5	4	-4	1	54-41	18.499	11.912	127.3	21
5	3	-2	3	53-23	16.846	10.416	92.7	21
5	2	-3	3	52-33	15.558	12.257	113.2	21
5	1	-1	4	51-14	14.731	11.912	92.7	21
5	4	1	5	5415	18.499	14.731	40.0	21
5	2	2	5	5225	15.558	15.558	46.4	21
4	3	-3	3	43-33	14.445	12.257	98.1	21
7	0	0	3	7003	20.223	8.667	90.0	21
10	4	-3	1	104-31	31.116	9.136	139.8	22
10	1	-2	2	101-22	29.034	8.171	129.3	22
9	4	-1	2	94-12	28.453	6.460	92.6	22
9	2	-2	2	92-22	26.635	8.171	122.5	22
9	5	1	3	9513	29.744	9.136	42.5	22
8	3	-2	2	83-22	24.684	8.171	114.4	22
8	2	-3	2	82-32	23.823	10.416	132.3	22
8	1	2	3	8123	23.292	10.416	49.2	22
8	2	1	3	8213	23.823	9.136	57.5	22

7	5	-3	1	75-31	24.852	9.136	126.0	22
7	4	-2	2	74-22	23.292	8.171	105.3	22
7	2	-4	2	72-42	21.032	12.920	137.5	22
7	1	-1	3	71-13	20.428	9.136	100.3	22
7	3	2	4	7324	22.002	12.920	40.2	22
6	5	-2	2	65-22	22.564	8.171	95.2	22
6	4	-1	3	64-13	20.833	9.136	74.7	22
6	4	-4	1	64-41	20.833	11.912	132.3	22
6	2	-2	3	62-23	18.272	10.416	105.3	22
6	2	-5	2	62-52	18.272	15.558	139.8	22
6	1	-4	3	61-43	17.573	14.445	133.7	22
6	1	2	4	6124	17.573	12.920	54.0	22
6	2	1	4	6214	18.272	11.912	57.5	22
5	4	-3	2	54-32	18.499	10.416	107.7	22
5	3	-4	2	53-42	16.846	12.920	122.5	22
5	2	-1	4	52-14	15.558	11.912	82.2	22
5	1	-2	4	51-24	14.731	12.920	105.3	22
5	3	1	5	5315	16.846	14.731	47.7	22
4	3	-2	4	43-24	14.445	12.920	79.7	22
9	2	2	3	9223	26.635	10.416	43.8	23
9	4	1	3	9413	28.453	9.136	47.6	23
8	5	-3	1	85-31	27.255	9.136	129.6	23
8	1	1	3	8113	23.292	9.136	64.4	23
7	-4	4	1	7-441	23.292	11.912	43.8	23
7	3	-3	2	73-32	22.002	10.416	123.1	23
7	2	-1	3	72-13	21.032	9.136	92.5	23
7	1	-2	3	71-23	20.428	10.416	115.6	23
7	5	1	4	7514	24.852	11.912	40.4	23
6	5	-1	3	65-13	22.564	9.136	68.6	23
6	1	-5	3	61-53	17.573	16.846	139.6	23
6	1	1	4	6114	17.573	11.912	66.5	23
5	4	-2	3	54-23	18.499	10.416	85.0	23
5	3	-1	4	53-14	16.846	11.912	73.1	23
5	2	-4	3	52-43	15.558	14.445	121.3	23
5	1	-3	4	51-34	14.731	14.445	115.6	23
5	2	1	5	5215	15.558	14.731	56.9	23
11	1	-2	2	111-22	31.910	8.171	129.8	24
10	2	-2	2	102-22	29.462	8.171	123.7	24
9	5	-3	1	95-31	29.744	9.136	132.5	24
9	3	-2	2	93-22	27.408	8.171	116.6	24
9	2	-3	2	92-32	26.635	10.416	133.8	24
9	3	1	3	9313	27.408	9.136	53.1	24
8	4	-2	2	84-22	25.840	8.171	108.4	24
8	4	-4	1	84-41	25.840	11.912	139.4	24

8	2	-4	2	82-42	23.823	12.920	139.4	24
8	0	1	3	8013	23.112	9.136	71.6	24
8	0	2	3	8023	23.112	10.416	56.3	24
8	0	3	3	8033	23.112	12.257	45.0	24
8	1	0	3	8103	23.292	8.667	82.9	24
8	2	0	3	8203	23.823	8.667	76.0	24
8	3	0	3	8303	24.684	8.667	69.4	24
8	4	0	3	8403	25.840	8.667	63.4	24
8	5	0	3	8503	27.255	8.667	58.0	24
8	6	0	3	8603	28.890	8.667	53.1	24
8	7	0	3	8703	30.711	8.667	48.8	24
7	6	-4	0	76-40	26.635	11.556	139.4	24
7	5	-2	2	75-22	24.852	8.171	99.5	24
7	3	-1	3	73-13	22.002	9.136	85.2	24
7	1	-3	3	71-33	20.428	12.257	126.9	24
7	2	2	4	7224	21.032	12.920	47.5	24
7	4	1	4	7414	23.292	11.912	46.2	24
6	6	-1	3	66-13	24.514	9.136	63.4	24
6	6	0	4	6604	24.514	11.556	45.0	24
6	4	-3	2	64-32	20.833	10.416	112.6	24
6	3	-2	3	63-23	19.380	10.416	97.1	24
6	3	-4	2	63-42	19.380	12.920	126.9	24
6	2	-3	3	62-33	18.272	12.257	116.6	24
6	0	1	4	6014	17.334	11.912	76.0	24
6	0	2	4	6024	17.334	12.920	63.4	24
8	0	0	3	8003	23.112	8.667	90.0	24
6	0	4	4	6044	17.334	16.343	45.0	24
6	1	0	4	6104	17.573	11.556	80.5	24
6	2	0	4	6204	18.272	11.556	71.6	24
6	2	3	5	6235	18.272	16.846	40.6	24
6	3	0	4	6304	19.380	11.556	63.4	24
6	3	2	5	6325	19.380	15.558	41.6	24
6	4	0	4	6404	20.833	11.556	56.3	24
6	5	0	4	6504	22.564	11.556	50.2	24
5	4	-1	4	54-14	18.499	11.912	65.4	24
5	3	-3	3	53-33	16.846	12.257	104.0	24
5	2	-2	4	52-24	15.558	12.920	94.8	24
5	1	-4	4	51-44	14.731	16.343	123.7	24
5	1	1	5	5115	14.731	14.731	67.4	24
4	4	-2	4	44-24	16.343	12.920	71.6	24
4	4	-3	3	44-33	16.343	12.257	90.0	24
6	0	0	4	6004	17.334	11.556	90.0	24
10	5	1	3	10513	32.300	9.136	45.0	25
9	1	2	3	9123	26.161	10.416	50.0	25

9	2	1	3	9213	26.635	9.136	59.0	25
8	3	-3	2	83-32	24.684	10.416	125.8	25
8	1	-1	3	81-13	23.292	9.136	101.3	25
7	2	-2	3	72-23	21.032	10.416	107.7	25
7	1	3	4	7134	20.428	14.445	45.0	25
7	3	1	4	7314	22.002	11.912	52.8	25
6	1	-1	4	61-14	17.573	11.912	94.6	25
5	5	-1	4	55-14	20.428	11.912	59.0	25
5	5	-2	3	55-23	20.428	10.416	78.7	25
5	4	-5	1	54-51	18.499	14.731	130.0	25
5	3	-5	2	53-52	16.846	15.558	127.2	25
5	1	0	5	5105	14.731	14.445	78.7	25
5	2	0	5	5205	15.558	14.445	68.2	25
5	3	0	5	5305	16.846	14.445	59.0	25
5	4	0	5	5405	18.499	14.445	51.3	25
5	0	0	5	5005	14.445	14.445	90.0	25
11	5	-3	1	115-31	34.908	9.136	137.1	26
11	2	-2	2	112-22	32.300	8.171	124.7	26
10	3	-2	2	103-22	30.162	8.171	118.3	26
10	-2	-2	3	10-2-23	29.462	10.416	135.0	26
10	4	1	3	10413	31.116	9.136	49.8	26
9	4	-2	2	94-22	28.453	8.171	111.0	26
9	1	1	3	9113	26.161	9.136	65.2	26
8	6	-3	1	86-31	28.890	9.136	124.7	26
8	5	-2	2	85-22	27.255	8.171	103.0	26
8	2	-1	3	82-13	23.823	9.136	94.4	26
8	1	-2	3	81-23	23.292	10.416	116.6	26
8	3	2	4	8324	24.684	12.920	42.9	26
7	6	-2	2	76-22	26.635	8.171	94.4	26
7	5	-1	3	75-13	24.852	9.136	72.9	26
7	3	-4	2	73-42	22.002	12.920	130.2	26
7	1	2	4	7124	20.428	12.920	55.3	26
7	2	1	4	7214	21.032	11.912	60.0	26
6	5	-4	1	65-41	22.564	11.912	126.2	26
6	4	-2	3	64-23	20.833	10.416	90.0	26
6	4	1	5	6415	20.833	14.731	45.0	26
6	2	-1	4	62-14	18.272	11.912	85.6	26
6	2	-4	3	62-43	18.272	14.445	124.7	26
6	1	-2	4	61-24	17.573	12.920	107.1	26
6	1	4	5	6145	17.573	18.499	41.9	26
6	2	2	5	6225	18.272	15.558	49.8	26
5	4	-4	2	54-42	18.499	12.920	114.8	26
5	3	-2	4	53-24	16.846	12.920	85.6	26
5	2	-3	4	52-34	15.558	14.445	105.1	26

5	1	-1	5	51-15	14.731	14.731	90.0	26
11	3	2	3	11323	32.940	10.416	41.1	27
10	3	1	3	10313	30.162	9.136	54.9	27
9	6	-3	1	96-31	31.249	9.136	127.9	27
9	3	-3	2	93-32	27.408	10.416	127.9	27
9	0	1	3	9013	26.001	9.136	71.6	27
9	0	2	3	9023	26.001	10.416	56.3	27
9	0	3	3	9033	26.001	12.257	45.0	27
9	1	0	3	9103	26.161	8.667	83.7	27
9	2	0	3	9203	26.635	8.667	77.5	27
9	3	0	3	9303	27.408	8.667	71.6	27
9	4	0	3	9403	28.453	8.667	66.0	27
9	5	0	3	9503	29.744	8.667	61.0	27
9	6	0	3	9603	31.249	8.667	56.3	27
9	7	0	3	9703	32.940	8.667	52.1	27
9	8	0	3	9803	34.788	8.667	48.4	27
8	3	-1	3	83-13	24.684	9.136	87.9	27
8	-1	3	3	8-133	23.292	12.257	52.1	27
8	5	1	4	8514	27.255	11.912	44.0	27
7	6	-1	3	76-13	26.635	9.136	67.8	27
7	5	-4	1	75-41	24.852	11.912	130.4	27
7	4	-5	1	74-51	23.292	14.731	139.0	27
7	3	-2	3	73-23	22.002	10.416	100.5	27
7	2	-3	3	72-33	21.032	12.257	119.1	27
7	1	1	4	7114	20.428	11.912	67.8	27
6	5	-3	2	65-32	22.564	10.416	106.5	27
6	3	-1	4	63-14	19.380	11.912	77.5	27
6	3	-3	3	63-33	19.380	12.257	108.4	27
6	3	-5	2	63-52	19.380	15.558	131.6	27
6	1	-3	4	61-34	17.573	14.445	117.4	27
6	1	3	5	6135	17.573	16.846	49.6	27
6	3	1	5	6315	19.380	14.731	52.1	27
5	4	-3	3	54-33	18.499	12.257	96.3	27
5	3	-4	3	53-43	16.846	14.445	112.2	27
5	2	-1	5	52-15	15.558	14.731	79.5	27
9	0	0	3	9003	26.001	8.667	90.0	27
12	2	-2	2	122-22	35.146	8.171	125.5	28
11	3	-2	2	113-22	32.940	8.171	119.7	28
11	2	-3	2	112-32	32.300	10.416	136.0	28
11	5	1	3	11513	34.908	9.136	47.1	28
10	6	-3	1	106-31	33.691	9.136	130.6	28
10	4	-2	2	104-22	31.116	8.171	113.2	28
10	1	2	3	10123	29.034	10.416	50.6	28
10	2	1	3	10213	29.462	9.136	60.3	28

9	5	-2	2	95-22	29.744	8.171	106.0	28
9	1	-1	3	91-13	26.161	9.136	102.1	28
8	7	-4	0	87-40	30.711	11.556	138.8	28
8	6	-2	2	86-22	28.890	8.171	98.1	28
8	5	-4	1	85-41	27.255	11.912	134.0	28
8	4	-1	3	84-13	25.840	9.136	81.9	28
8	4	-3	2	84-32	25.840	10.416	119.7	28
8	3	-4	2	83-42	24.684	12.920	132.9	28
8	2	-2	3	82-23	23.823	10.416	109.7	28
8	1	-4	3	81-43	23.292	14.445	136.0	28
8	2	2	4	8224	23.823	12.920	49.4	28
8	4	1	4	8414	25.840	11.912	49.4	28
7	7	-1	3	77-13	28.600	9.136	63.4	28
7	7	-2	2	77-22	28.600	8.171	90.0	28
7	7	0	4	7704	28.600	11.556	45.0	28
7	0	1	4	7014	20.223	11.912	76.0	28
7	0	2	4	7024	20.223	12.920	63.4	28
7	0	3	4	7034	20.223	14.445	53.1	28
7	0	4	4	7044	20.223	16.343	45.0	28
7	1	0	4	7104	20.428	11.556	81.9	28
7	2	0	4	7204	21.032	11.556	74.1	28
7	3	0	4	7304	22.002	11.556	66.8	28
7	4	0	4	7404	23.292	11.556	60.3	28
7	5	0	4	7504	24.852	11.556	54.5	28
7	6	0	4	7604	26.635	11.556	49.4	28
6	5	-2	3	65-23	22.564	10.416	83.9	28
6	4	-1	4	64-14	20.833	11.912	70.4	28
6	4	-4	2	64-42	20.833	12.920	119.7	28
6	2	-2	4	62-24	18.272	12.920	98.1	28
6	2	-5	3	62-53	18.272	16.846	130.6	28
6	1	-4	4	61-44	17.573	16.343	125.5	28
6	1	2	5	6125	17.573	15.558	58.7	28
6	2	1	5	6215	18.272	14.731	60.3	28
5	4	-2	4	54-24	18.499	12.920	77.9	28
5	3	-1	5	53-15	16.846	14.731	70.4	28
5	2	-4	4	52-44	15.558	16.343	113.2	28
4	4	-3	4	44-34	16.343	14.445	81.9	28
7	0	0	4	7004	20.223	11.556	90.0	28
11	6	-3	1	116-31	36.199	9.136	133.0	29
11	2	2	3	11223	32.300	10.416	46.0	29
11	4	1	3	11413	33.815	9.136	51.6	29
10	3	-3	2	103-32	30.162	10.416	129.6	29
10	1	1	3	10113	29.034	9.136	65.9	29
9	5	-4	1	95-41	29.744	11.912	136.9	29

9	2	-1	3	92-13	26.635	9.136	95.9	29
9	1	-2	3	91-23	26.161	10.416	117.4	29
8	7	-3	1	87-31	30.711	9.136	120.4	29
8	5	-1	3	85-13	27.255	9.136	76.4	29
8	1	3	4	8134	23.292	14.445	46.0	29
8	3	1	4	8314	24.684	11.912	55.4	29
7	5	-3	2	75-32	24.852	10.416	110.8	29
7	4	-2	3	74-23	23.292	10.416	94.0	29
7	3	2	5	7325	22.002	15.558	45.0	29
7	-2	4	3	7-243	21.032	14.445	52.8	29
7	1	-1	4	71-14	20.428	11.912	95.9	29
7	2	3	5	7235	21.032	16.846	43.1	29
6	5	-1	4	65-14	22.564	11.912	64.2	29
6	1	-5	4	61-54	17.573	18.499	131.9	29
6	1	1	5	6115	17.573	14.731	69.2	29
5	4	-1	5	54-15	18.499	14.731	62.7	29
5	3	-3	4	53-34	16.846	14.445	95.9	29
5	2	-2	5	52-25	15.558	15.558	90.0	29
13	2	-2	2	132-22	37.999	8.171	126.3	30
12	6	1	3	12613	38.760	9.136	45.0	30
12	3	-2	2	123-22	35.735	8.171	121.0	30
12	2	-3	2	122-32	35.146	10.416	136.9	30
12	3	2	3	12323	35.735	10.416	42.3	30
11	4	-2	2	114-22	33.815	8.171	115.0	30
11	3	1	3	11313	32.940	9.136	56.3	30
10	5	-2	2	105-22	32.300	8.171	108.4	30
10	5	-4	1	105-41	32.300	11.912	139.4	30
10	0	1	3	10013	28.890	9.136	71.6	30
10	0	2	3	10023	28.890	10.416	56.3	30
10	0	3	3	10033	28.890	12.257	45.0	30
10	1	0	3	10103	29.034	8.667	84.3	30
10	2	0	3	10203	29.462	8.667	78.7	30
10	3	0	3	10303	30.162	8.667	73.3	30
10	4	0	3	10403	31.116	8.667	68.2	30
10	5	0	3	10503	32.300	8.667	63.4	30
10	6	0	3	10603	33.691	8.667	59.0	30
10	7	0	3	10703	35.265	8.667	55.0	30
10	8	0	3	10803	36.997	8.667	51.3	30
9	6	-2	2	96-22	31.249	8.171	101.3	30
9	4	-3	2	94-32	28.453	10.416	122.4	30
9	3	-1	3	93-13	27.408	9.136	90.0	30
9	3	2	4	9324	27.408	12.920	45.0	30
9	1	-3	3	91-33	26.161	12.257	128.7	30
9	2	3	4	9234	26.635	14.445	40.6	30

9	6	1	4	9614	31.249	11.912	42.3	30
8	7	-2	2	87-22	30.711	8.171	93.8	30
8	3	-2	3	83-23	24.684	10.416	103.1	30
8	2	-3	3	82-33	23.823	12.257	121.0	30
8	1	2	4	8124	23.292	12.920	56.3	30
8	2	1	4	8214	23.823	11.912	61.9	30
7	3	-3	3	73-33	22.002	12.257	111.8	30
7	2	-1	4	72-14	21.032	11.912	88.1	30
7	1	-2	4	71-24	20.428	12.920	108.4	30
7	5	1	5	7515	24.852	14.731	43.2	30
6	6	-1	4	66-14	24.514	11.912	59.0	30
6	6	-2	3	66-23	24.514	10.416	78.7	30
6	6	0	5	6605	24.514	14.445	45.0	30
6	4	-3	3	64-33	20.833	12.257	101.3	30
6	4	-6	1	64-61	20.833	17.573	136.9	30
6	3	-2	4	63-24	19.380	12.920	90.0	30
6	3	0	5	6305	19.380	14.445	63.4	30
6	3	2	6	6326	19.380	18.272	45.0	30
6	2	0	5	6205	18.272	14.445	71.6	30
6	0	1	5	6015	17.334	14.731	78.7	30
6	0	2	5	6025	17.334	15.558	68.2	30
6	0	3	5	6035	17.334	16.846	59.0	30
6	0	4	5	6045	17.334	18.499	51.3	30
6	0	5	5	6055	17.334	20.428	45.0	30
6	1	0	5	6105	17.573	14.445	80.5	30
6	4	0	5	6405	20.833	14.445	56.3	30
6	5	0	5	6505	22.564	14.445	50.2	30
5	5	-1	5	55-15	20.428	14.731	56.3	30
5	5	-3	3	55-33	20.428	12.257	90.0	30
5	4	-5	2	54-52	18.499	15.558	119.5	30
5	3	-5	3	53-53	16.846	16.846	118.1	30
10	0	0	3	10003	28.890	8.667	90.0	30
6	0	0	5	6005	17.334	14.445	90.0	30

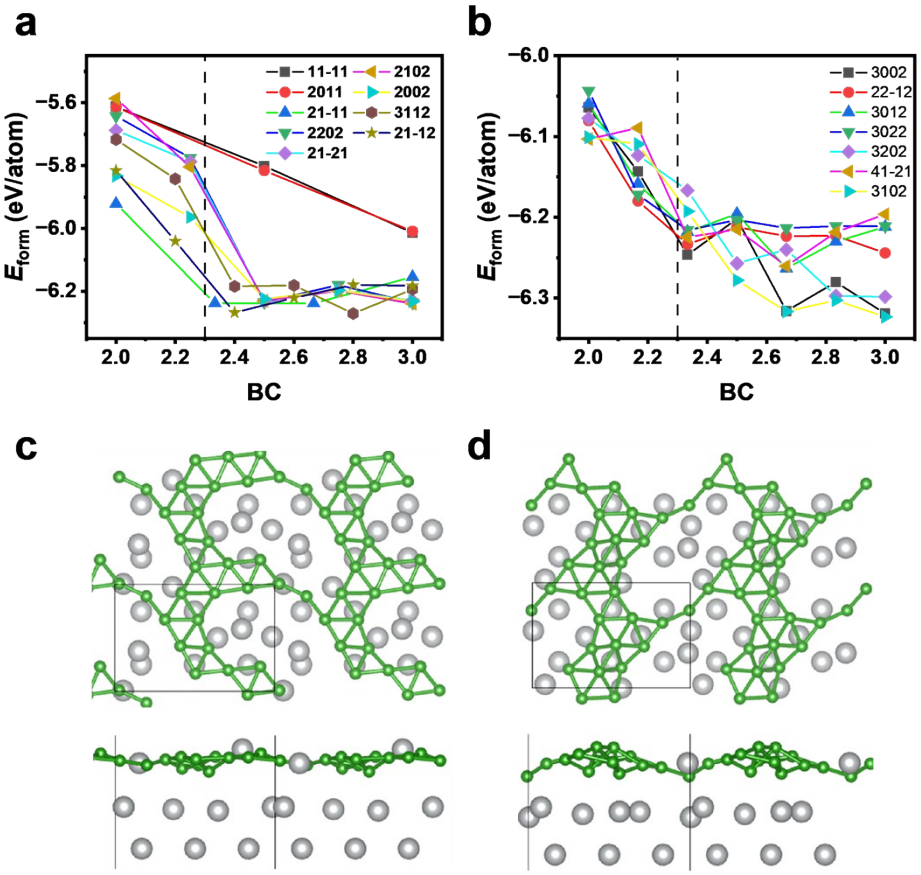

Figure S1. The variation of formation energy (E_{form}) with boron coverage (*BC*) for (a) sampling systems *M*=2~5 and (b) target systems *M*=6, respectively. The legend is different supercells referred as $r_{11}r_{12}r_{21}r_{22}$ in the redefine matrix. The minima structures obtained within 3002 at (c) *BC*=2 and (d) *BC*= 2.167 are displayed. It is found that high-energy mixed boron-metal phases are formed when *BC* is less than 2.3, so we do not consider the structures with *BC* value less than 2.3.

Table S2. The multiple (***M***) values, boron coverage (***BC***), and supercell types (referred as $r_{11}r_{21}r_{21}r_{22}$ in which r_{11}, r_{21}, r_{21} and r_{22} are elements of redefine matrix) for 61 test systems.

M	***BC***	**a**	**b**	**γ**	**Supercell**
10	6.900	17.573	11.912	156.5	61-41
10	2.600	14.445	10.416	146.3	50-32
10	3.300	15.558	14.445	158.2	52-50
10	4.600	17.573	8.171	144.5	6-1-22
10	5.500	14.731	14.731	157.4	51-51
10	3.700	14.731	14.731	22.6	5-151
10	5.500	14.445	15.558	21.8	5052
10	4.600	14.445	10.416	33.7	5032
10	6.800	14.445	12.92	26.6	5042
10	2.700	17.573	11.912	23.5	6-141
15	2.533	22.002	10.416	146.9	7-3-23
15	3.333	14.445	14.445	143.1	50-43
15	4.800	16.846	14.445	149.0	53-50
15	5.667	20.428	15.558	156.8	55-5-2
15	6.667	21.032	12.257	150.9	7-2-33
15	4.600	21.032	11.912	30.0	7-241
15	5.267	17.573	12.257	35.5	6133
15	6.667	20.428	15.558	23.2	5525
15	4.000	24.684	12.257	24.4	8333
15	3.067	21.032	12.257	29.1	7233
20	5.350	18.499	14.445	141.3	54-50
20	4.700	20.833	20.833	157.4	6-4-46
20	6.950	29.462	14.731	157.4	102-51
20	4.150	29.744	11.556	150.9	95-40
20	3.550	25.84	14.445	153.4	84-50
20	2.750	29.462	14.731	22.6	10-251
20	3.450	22.564	11.556	39.8	6-540
20	4.850	17.573	16.343	35.5	6144
20	5.950	28.89	12.92	26.6	10042
20	4.450	20.428	14.731	33.7	5515
20	6.700	25.84	14.445	153.4	84-50
20	2.550	17.573	16.343	144.5	6-1-44
25	2.760	24.852	15.558	147.3	7-5-25
25	3.000	27.255	16.846	153.0	8-5-35
25	3.520	22.002	17.573	147.3	73-61
25	3.880	28.453	20.428	159.0	9-4-55
25	4.760	23.292	23.292	157.4	8-1-74
25	5.520	30.162	14.731	152.0	103-51
25	6.440	24.684	20.428	155.6	8-3-55
25	2.760	33.815	10.416	36.3	11423

25	3.480	28.453	20.428	21.0	9455
25	4.080	24.852	14.445	35.5	7-550
25	4.920	22.564	14.731	38.9	6515
25	6.160	32.3	14.445	26.6	11243
25	7.200	30.71	11.912	34.8	8714
25	5.760	27.255	16.846	27.0	8535
30	4.967	34.788	18.272	156.8	121-62
30	7.233	22.564	17.334	140.2	65-60
30	2.633	28.89	16.846	149.0	100-53
30	5.567	36.199	14.445	151.4	116-50
30	4.600	40.135	17.573	159.2	127-6-1
30	4.033	36.997	14.731	152.7	108-5-1
30	3.233	20.428	20.428	143.1	7-1-55
30	3.767	24.852	17.334	144.5	75-60
30	2.833	31.115	20.428	23.2	10455
30	3.967	29.744	17.334	29.1	9-560
30	4.433	40.239	17.334	21.0	13-560
30	5.133	27.255	15.558	36.2	8525
30	5.700	29.744	16.846	30.0	9535
30	6.300	44.193	14.731	22.6	15-351
30	7.033	28.89	24.684	20.6	10083

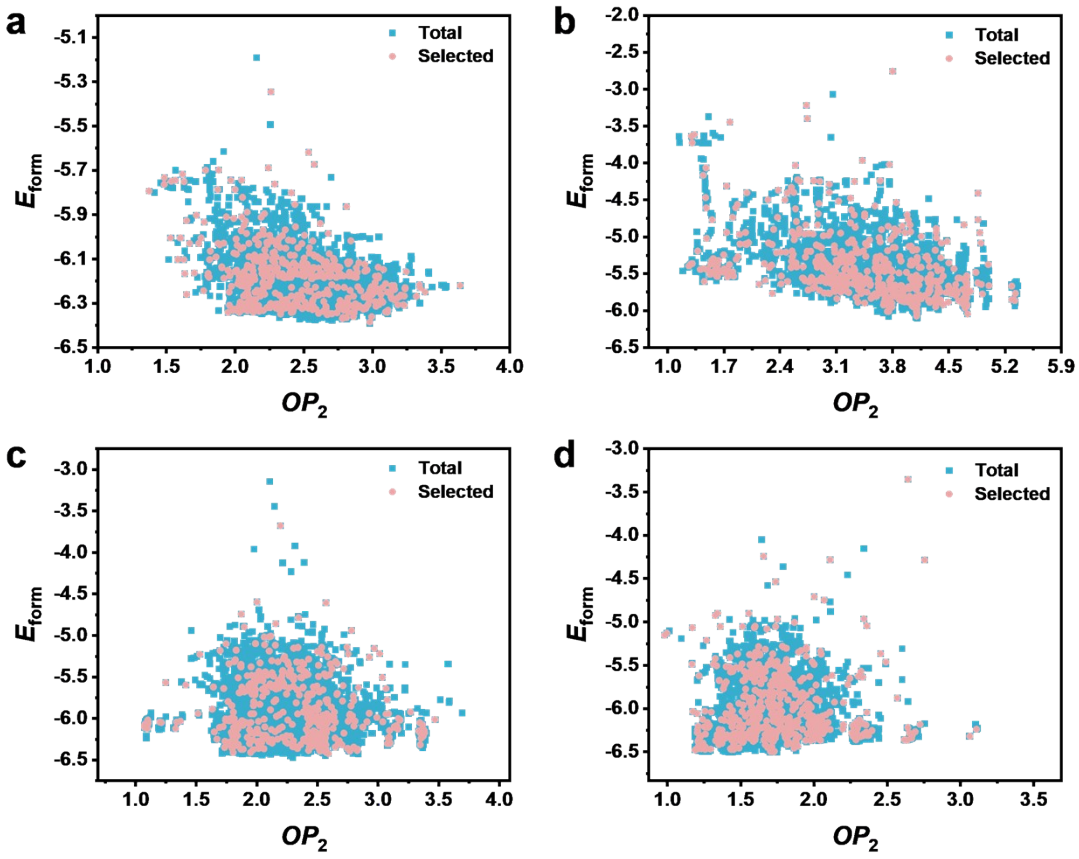

Figure S2. The selection of initial dataset in E_{form}-OP_2 distribution take four small systems as example: (a) supercell type 2011 at *BC* of 4, (b) supercell type 21-11 at *BC* of 3, (c) supercell type 3011 at *BC* of 2, and (d) supercell type 2102 at *BC* of 5. The total and selected dataset are denoted by the light blue square and pink sphere, respectively. Generally, the total dataset for a system includes about 14 thousand data, from which about 600 data are selected.

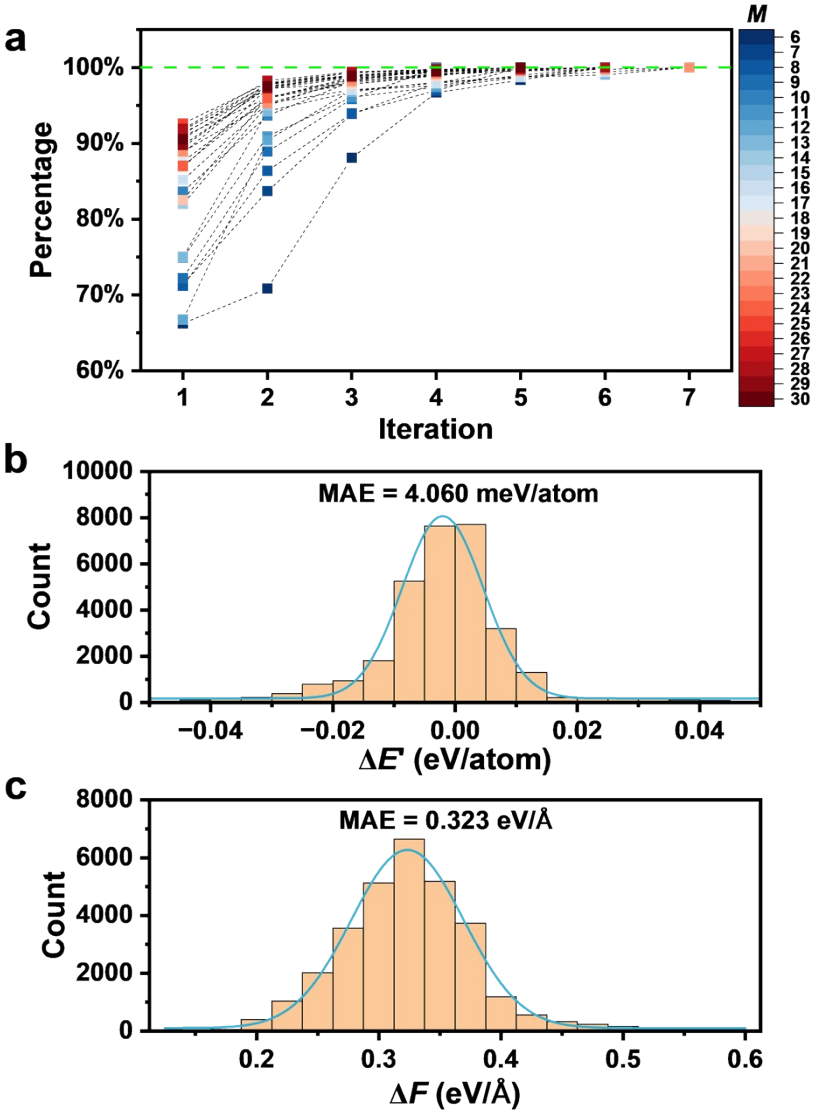

Figure S3. (a) The convergence trend of structures under different *M* values from 6 to 30 with respect to the number of iterative steps. The percentage denotes the number of structures meeting energy convergence criteria as a proportion of the total number of structures at each iteration. The green dash line denotes 100%. (b) Energy ($\Delta E'$) and (c) Force (ΔF) error distribution for the validation dataset

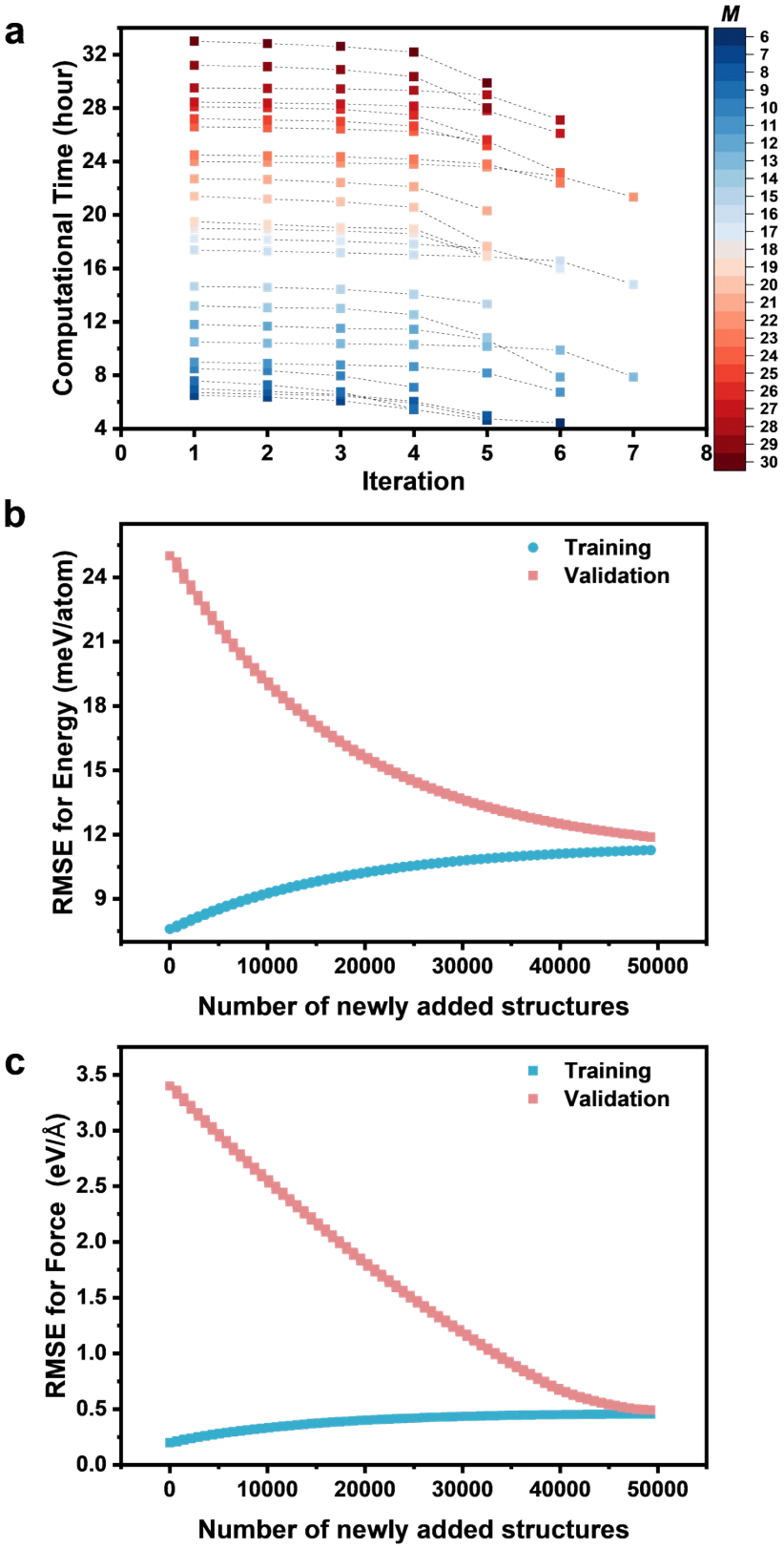

Figure S4. The evolution of root-mean-square errors (RMSE) for (a) energy (meV/atom) and (b) force (eV/Å) are plotted against the number of newly added structures. (c) The computational time per iteration using about 320 CPU cores.

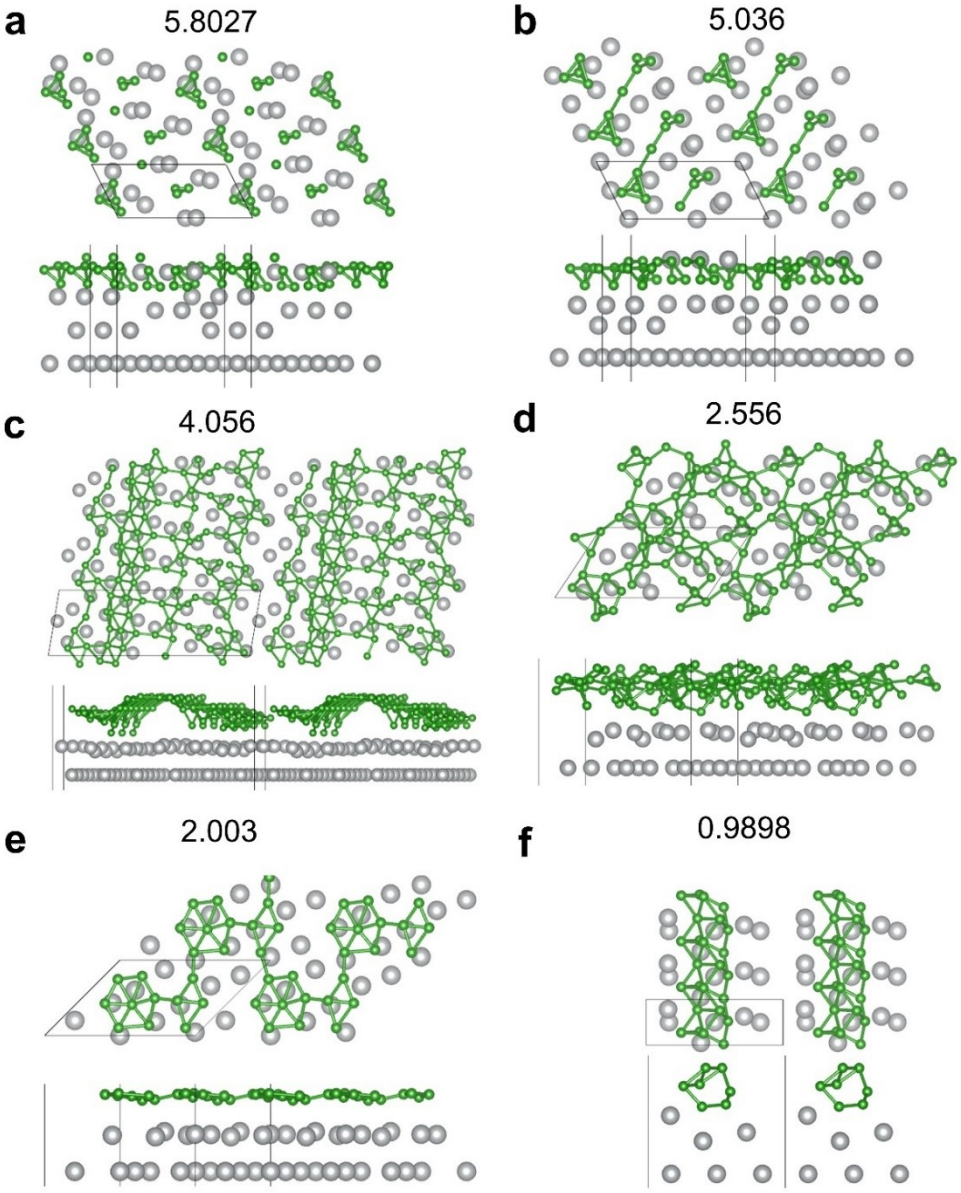

Figure S5. Several typical structures in the global dataset with different OP_2 values. The two values above structural panel are OP_2 values.

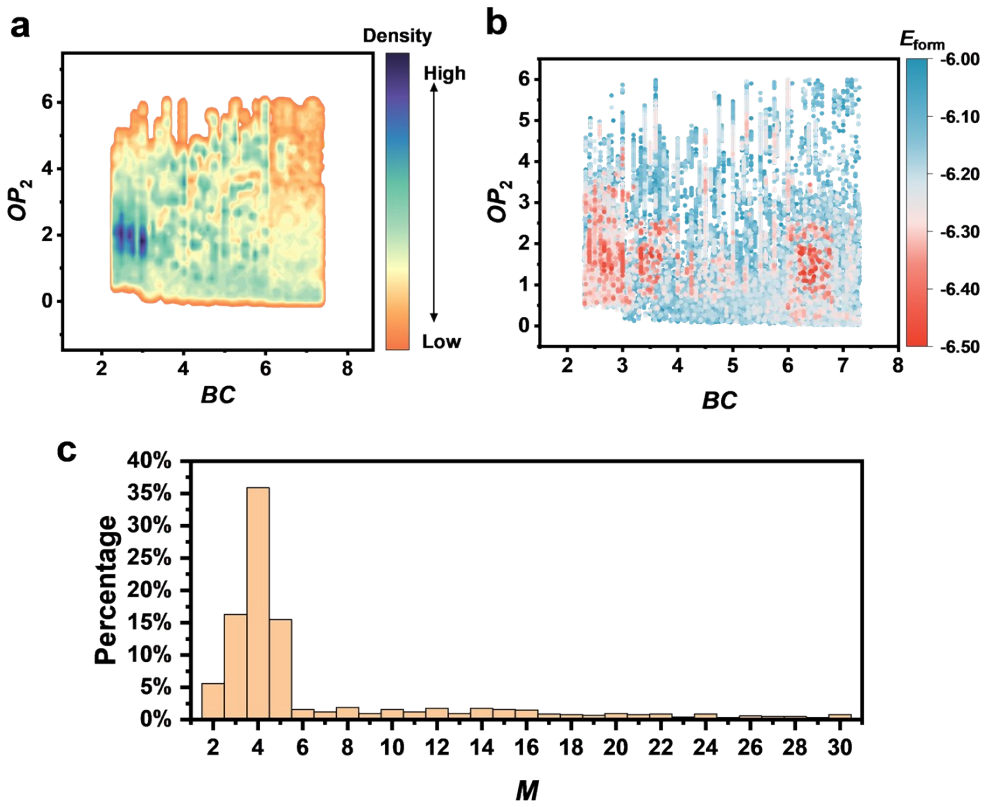

Figure S6. **(a)** Illustration of the global dataset from first principles for borophene on Ag(100) systems by OP_2 vs. BC contour plot. The density of states (DOS) is indicated by color. OP_2 parameter indicates the averaged geometrical environment of bonded boron atoms. **(b)** The OP_2 vs. BC scatter plot for low-energy data with E_{form} lower than -6.0 eV/atom. **(c)** The percentages of different M values from 2 to 30 of supercells in the global dataset.

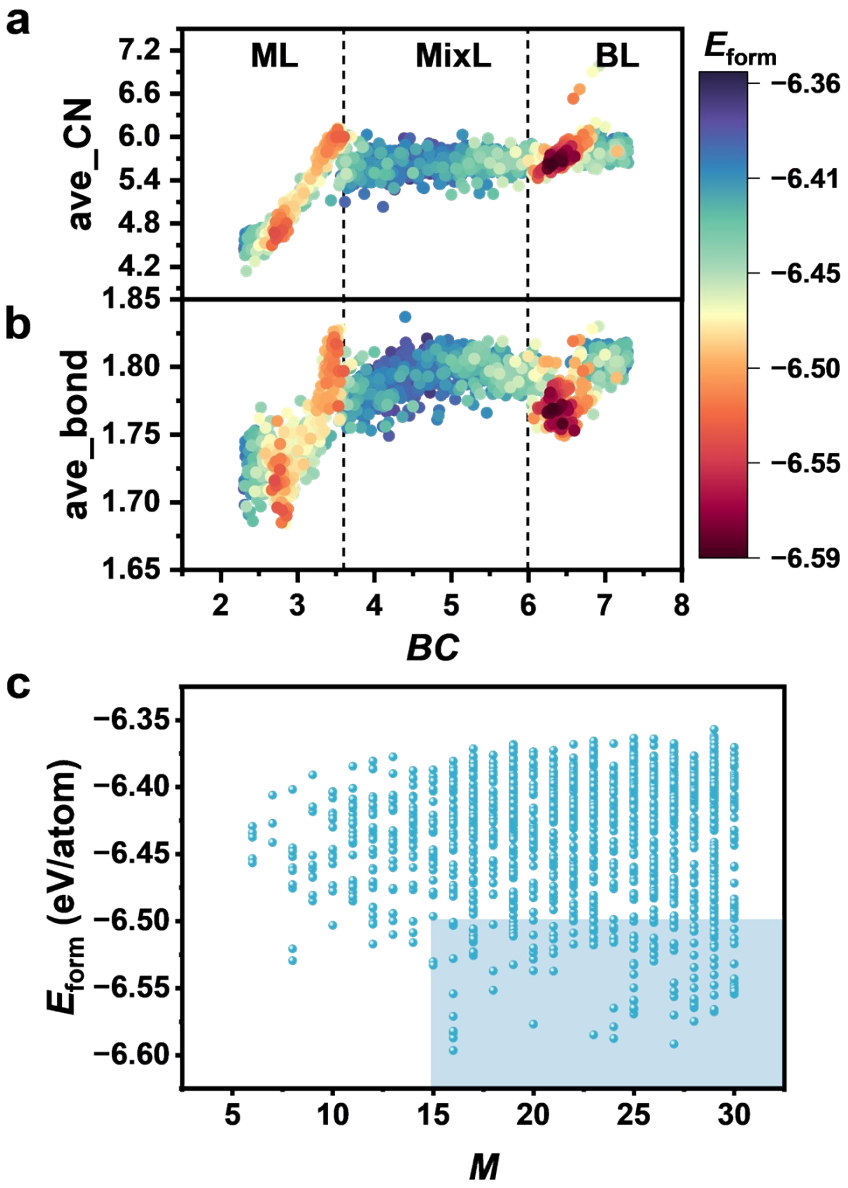

Figure S7. The variation of **(a)** averaged coordination number (ave_CN) and **(b)** averaged boron-boron bond length (ave_bond) with *BC* values colored by E_{form} values for low-energy minima. (c) The corresponding distribution of E_{form} values versus *M* values.

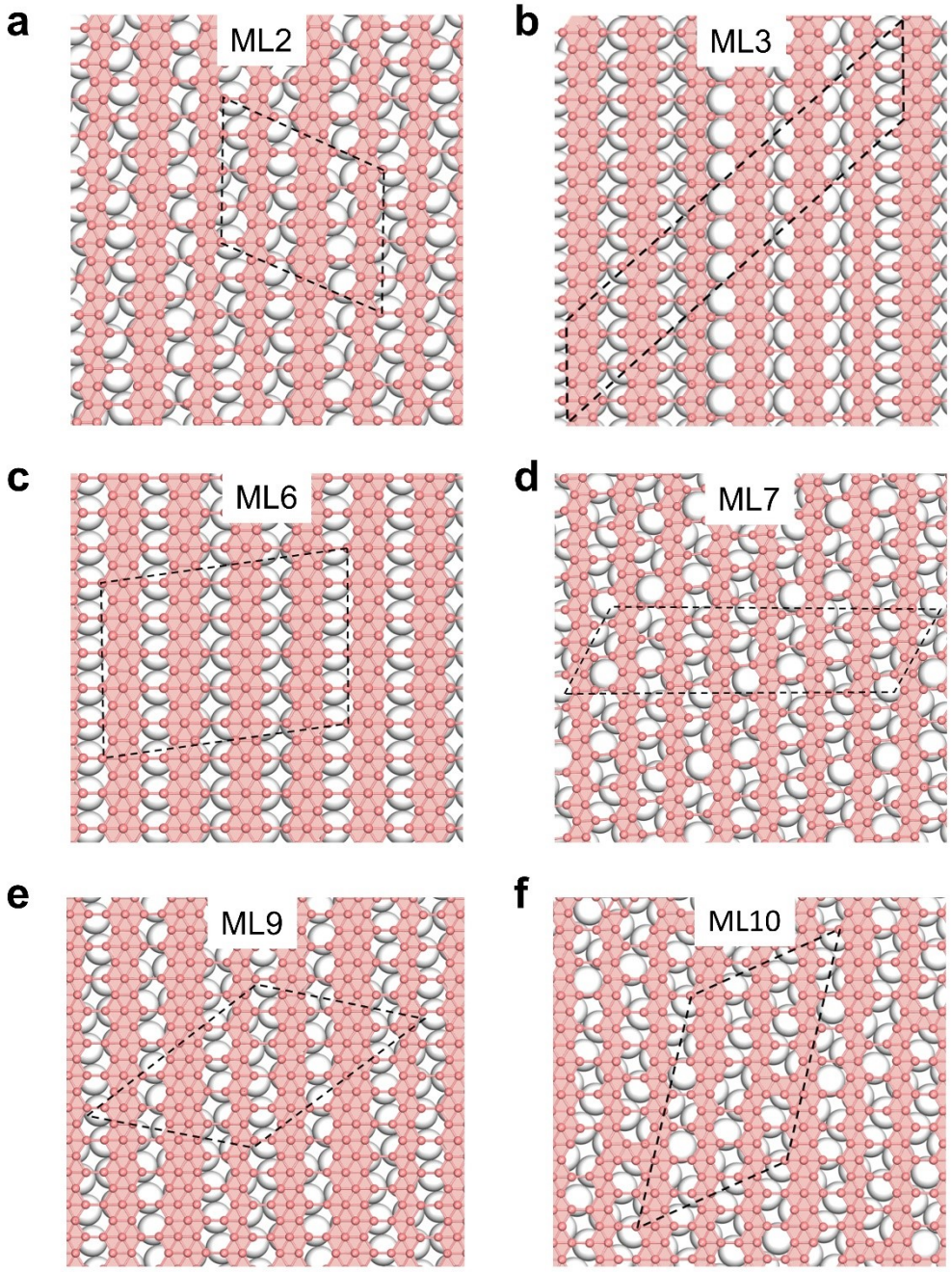

Figure S8. The atomic patterns of (a) ML2, (b) ML3, (c) ML6, (d) ML7, (e) ML9, and (f) ML10 in energy valley 1. The light gray spheres denote Ag atoms, and the pink colors denote boron atoms. The black dashed lines denote the unit cell.

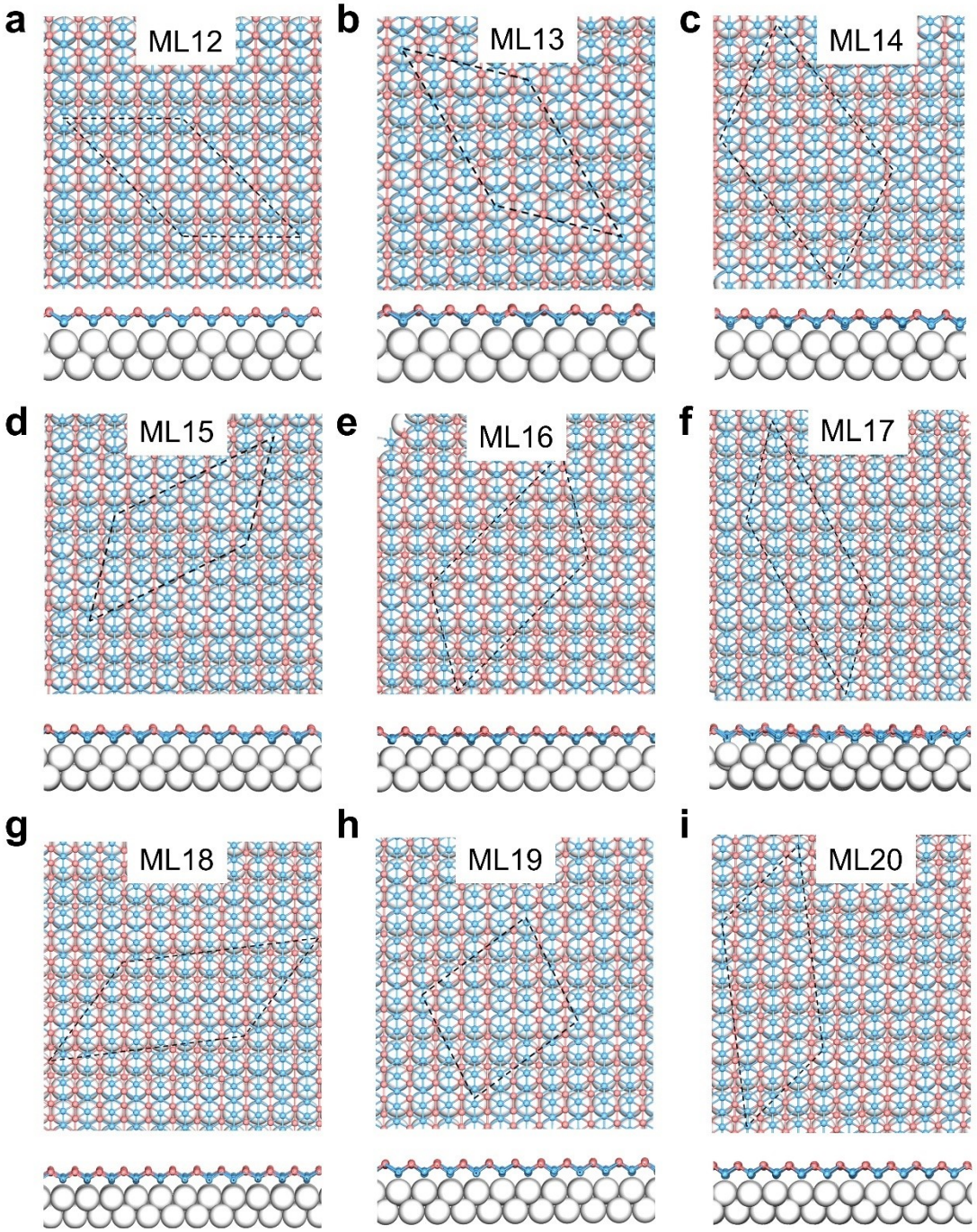

Figure S9. The atomic patterns of (a~i) ML12~ML20 in energy valley 2. The light gray spheres denote Ag atoms, and the pink colors and light cyan denote upper and lower boron atoms, respectively. The black dashed lines denote the unit cell.

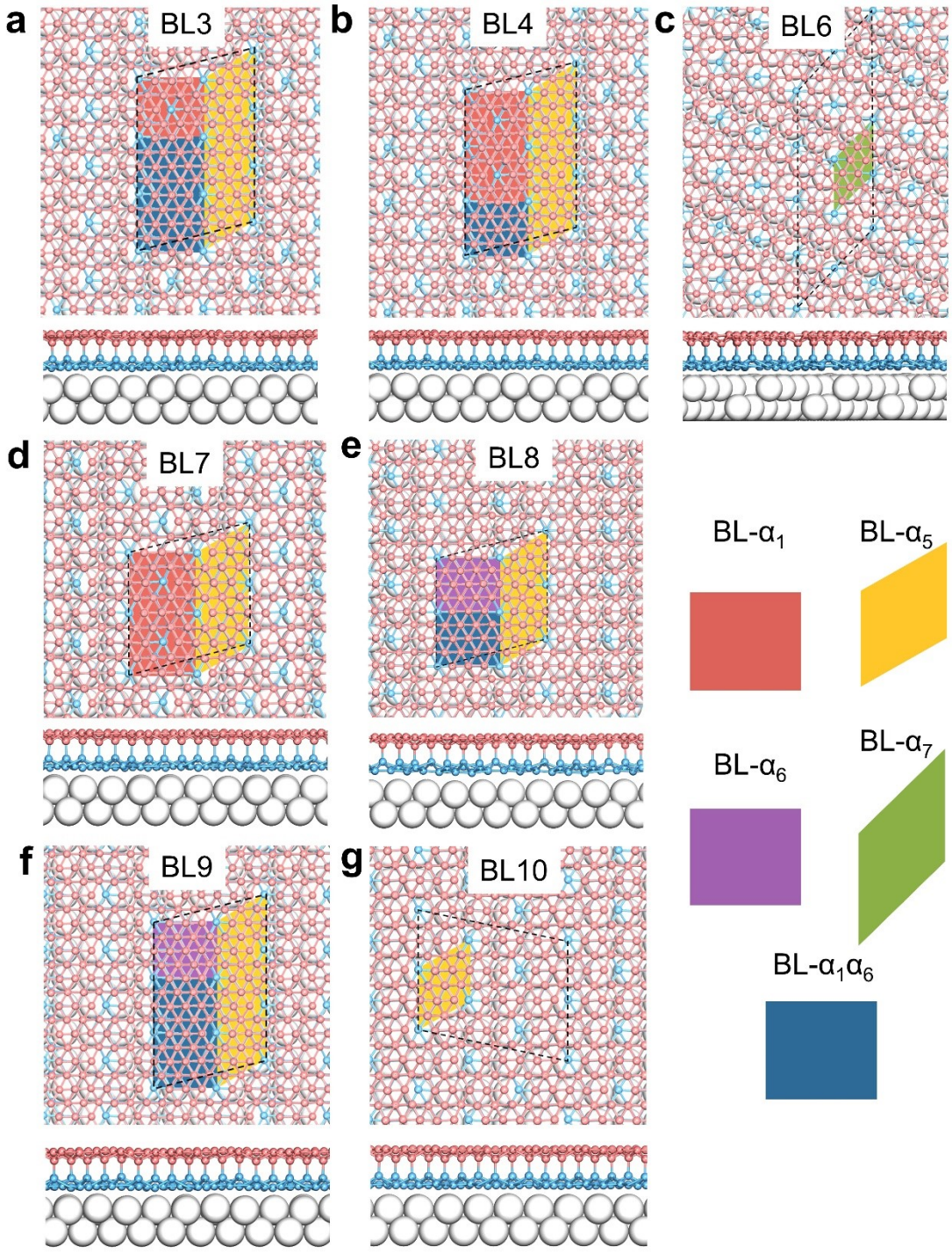

Figure S10. The atomic patterns of (a~b) BL3~BL4, and (c~g) BL6~BL10 in energy valley 3. The light gray spheres denote Ag atoms, and the pink colors and light cyan denote top and bottom boron atoms, respectively. The black dashed lines denote the unit cell. Different bilayer borophene fragments are depicted by different shapes with different colors. BL-$\alpha_1\alpha_6$ denotes the BL with α_1 as the bottom sheet and α_6 as the top sheet.

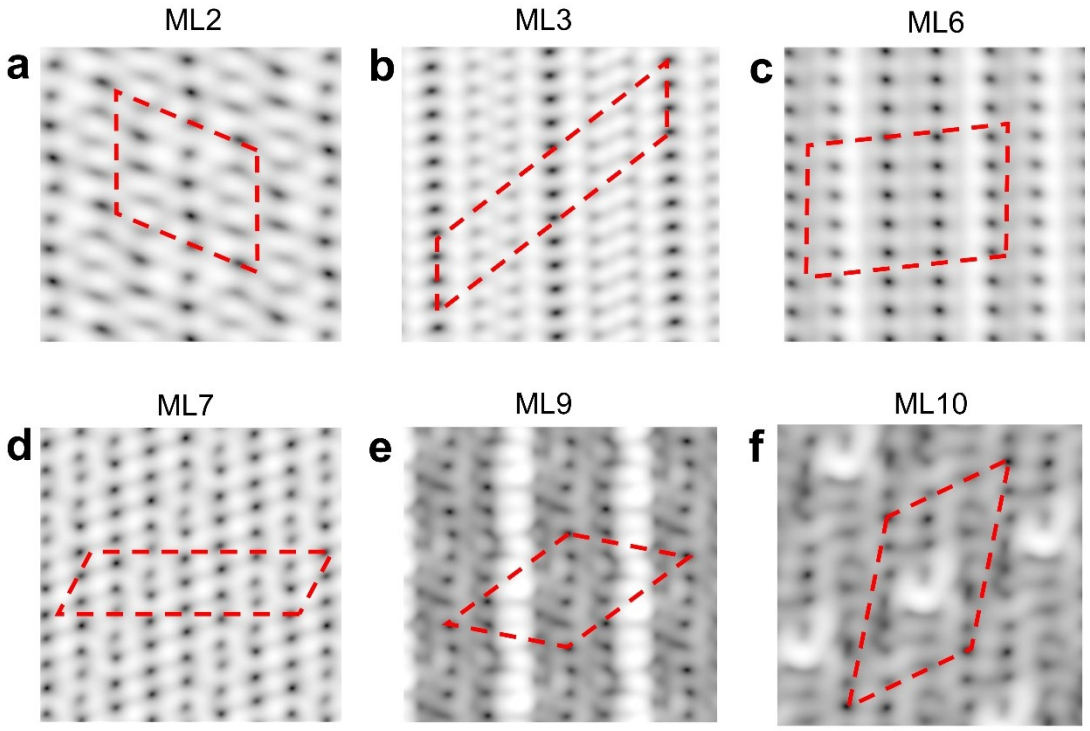

Figure S11. Simulated STM images of (a) ML2, (b) ML3, (c) ML6, (d) ML7, (e) ML9, and (f) ML10 on Ag(100), which was performed under the constant-current mode. The red dashed lines denote the primitive cell.

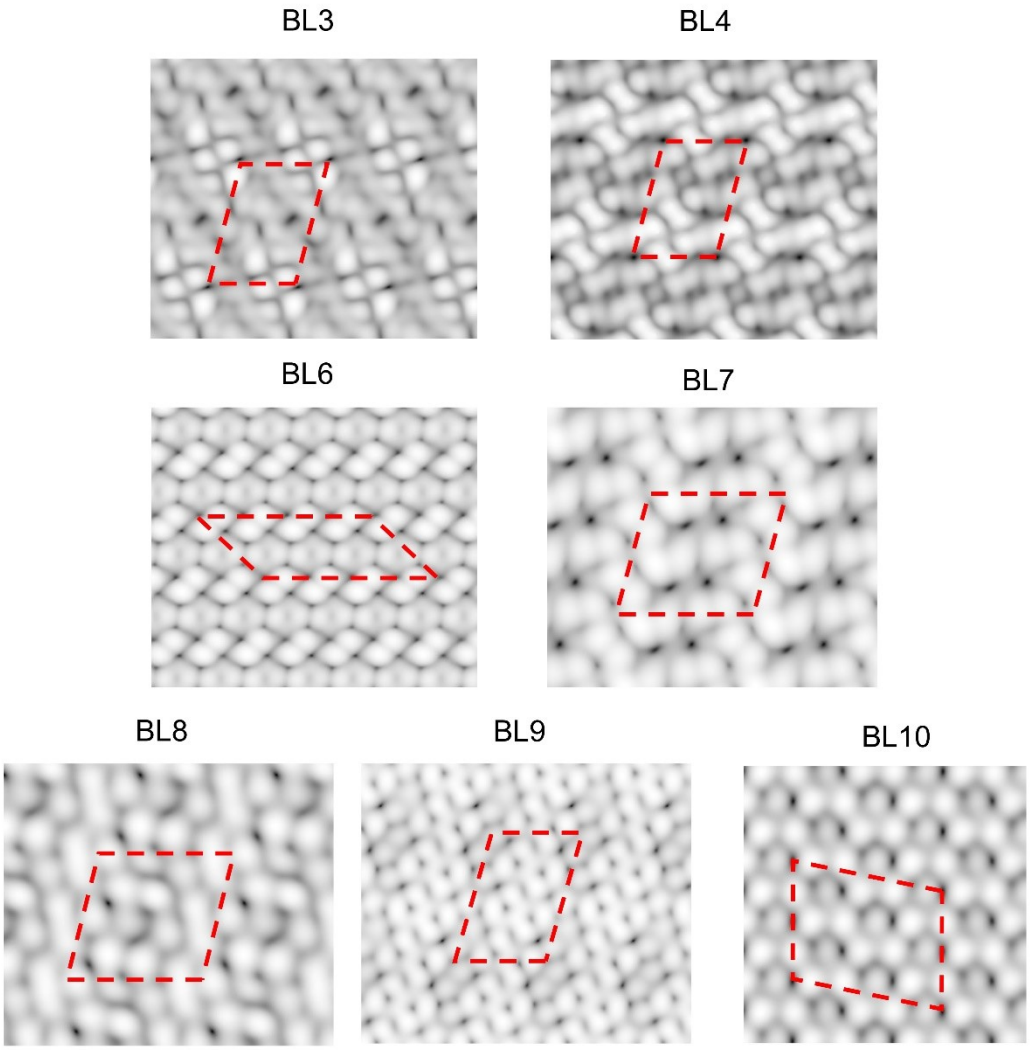

Figure S12. Simulated STM images of BL3~BL9 on Ag(100), which was performed under the constant-current mode. The red dashed lines denote the primitive cell.

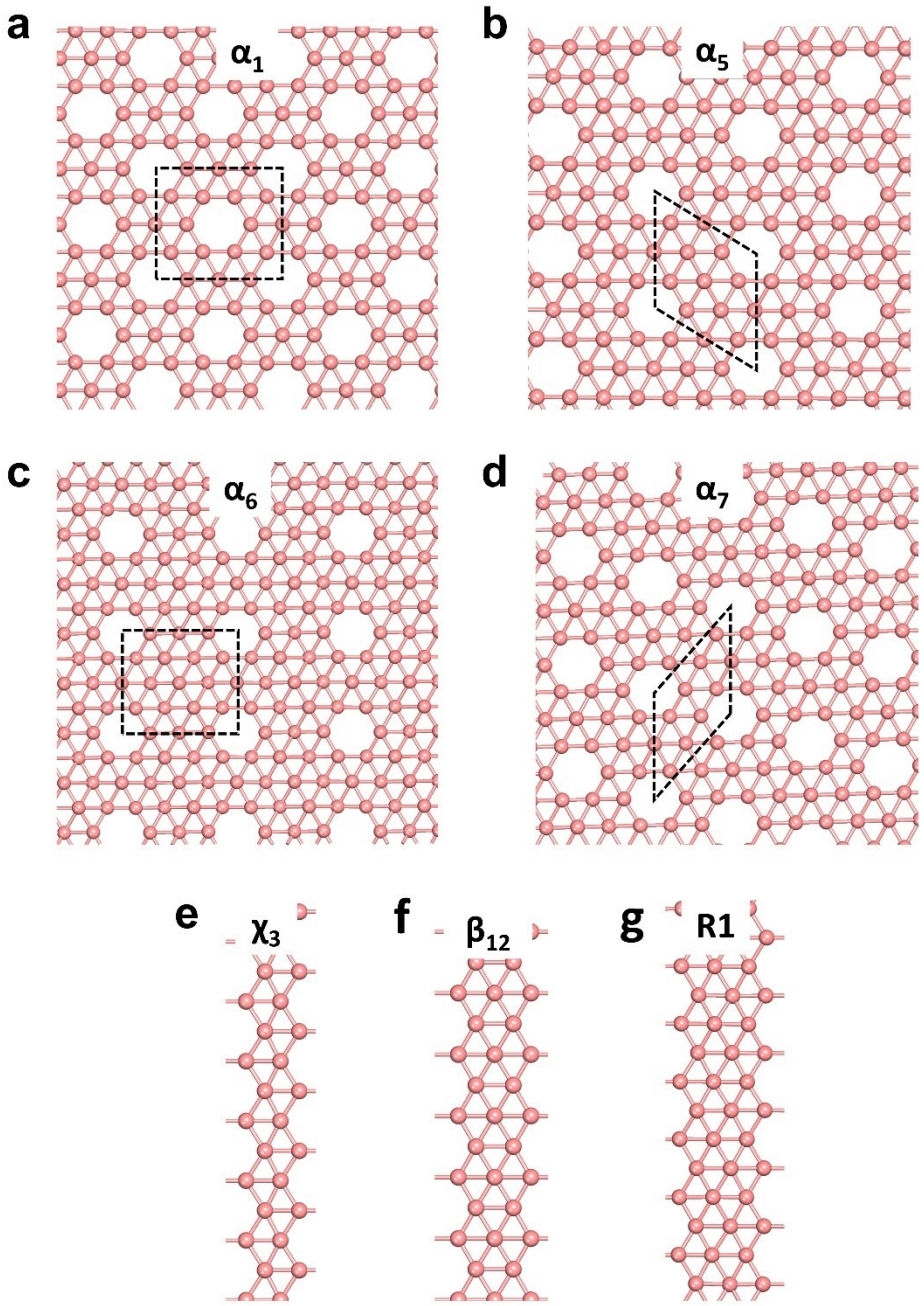

Figure S13. The atomic patterns of MLs: (a) α_1, (b) α_5, (c) α_6, and (d) α_7; and three ribbons: (e) χ_3, (f) β_{12}, and (g) R1, composed of 2, 2.5, and 3 boron chains, respectively. The pink colors denote boron atoms. The black dashed lines denote the unit cell.

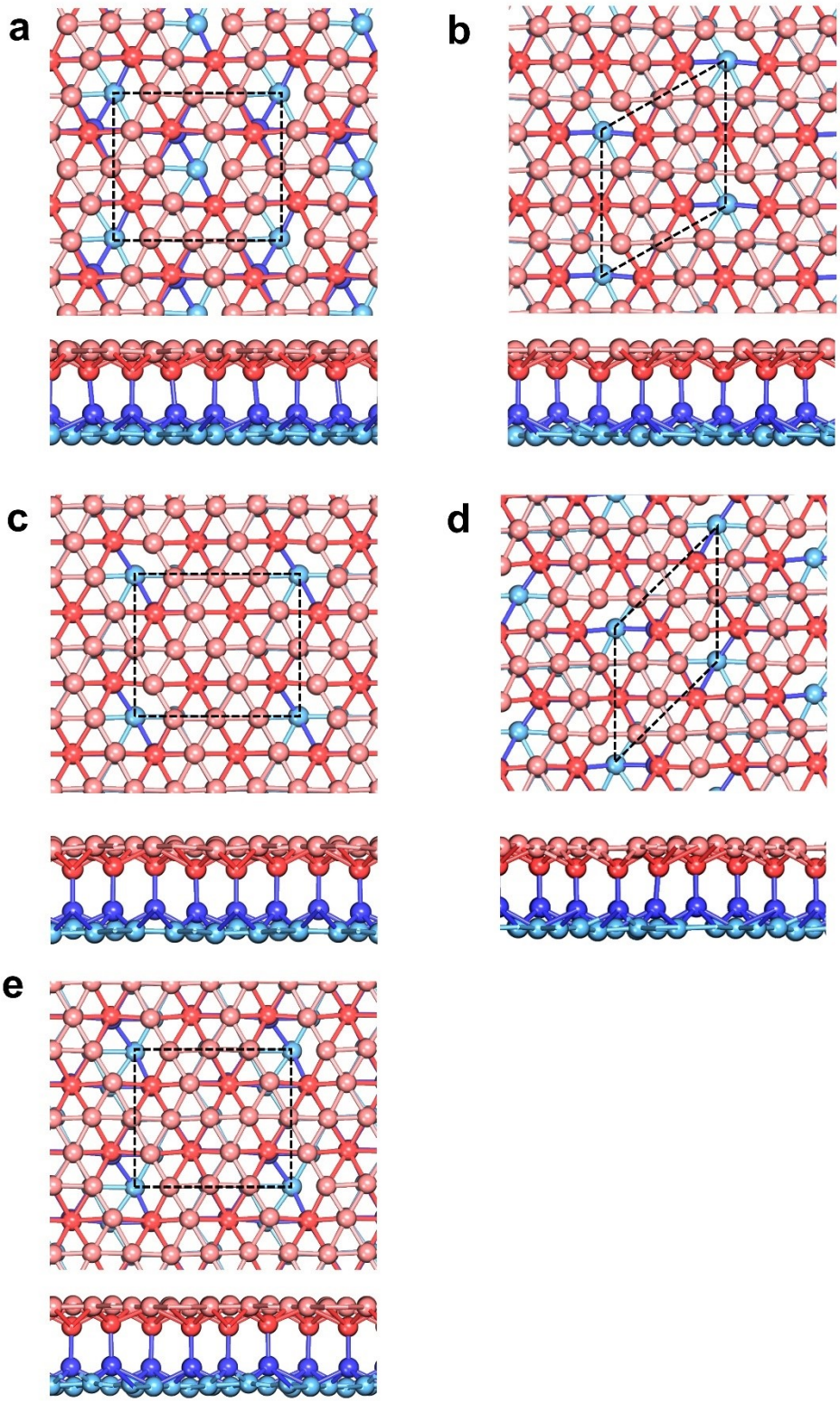

Figure S14. The top and side views of atomic patterns of (a) BL-α_1, (b) BL-α_5, (c) BL-α_6, (d) BL-α_7 and (e) BL-$\alpha_1\alpha$. by AB stacking. The pink and light cyan colors denote topmost and bottommost boron atoms, respectively. The upper and lower boron atoms for the formation of interlayer bonds are colored by red and dark cyan, respectively. The black dashed lines denote the unit cell.

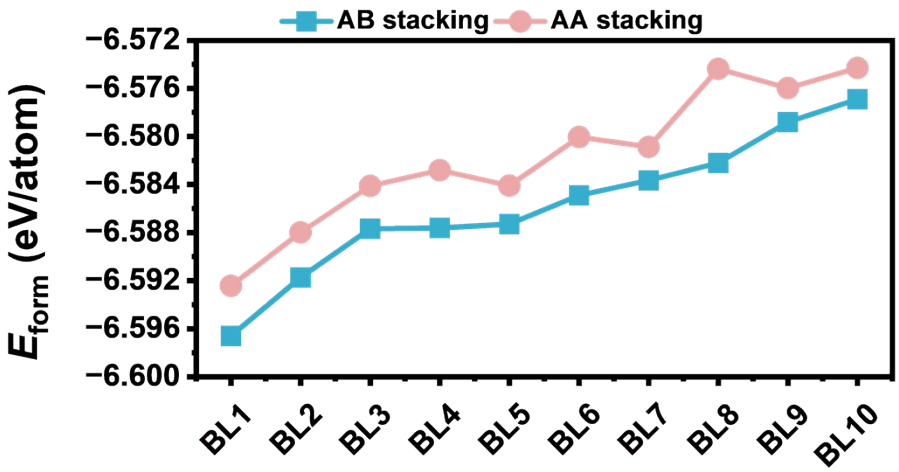

Figure S15. The comparison of E_{form} by AA and AB stacking styles for BL1~BL10

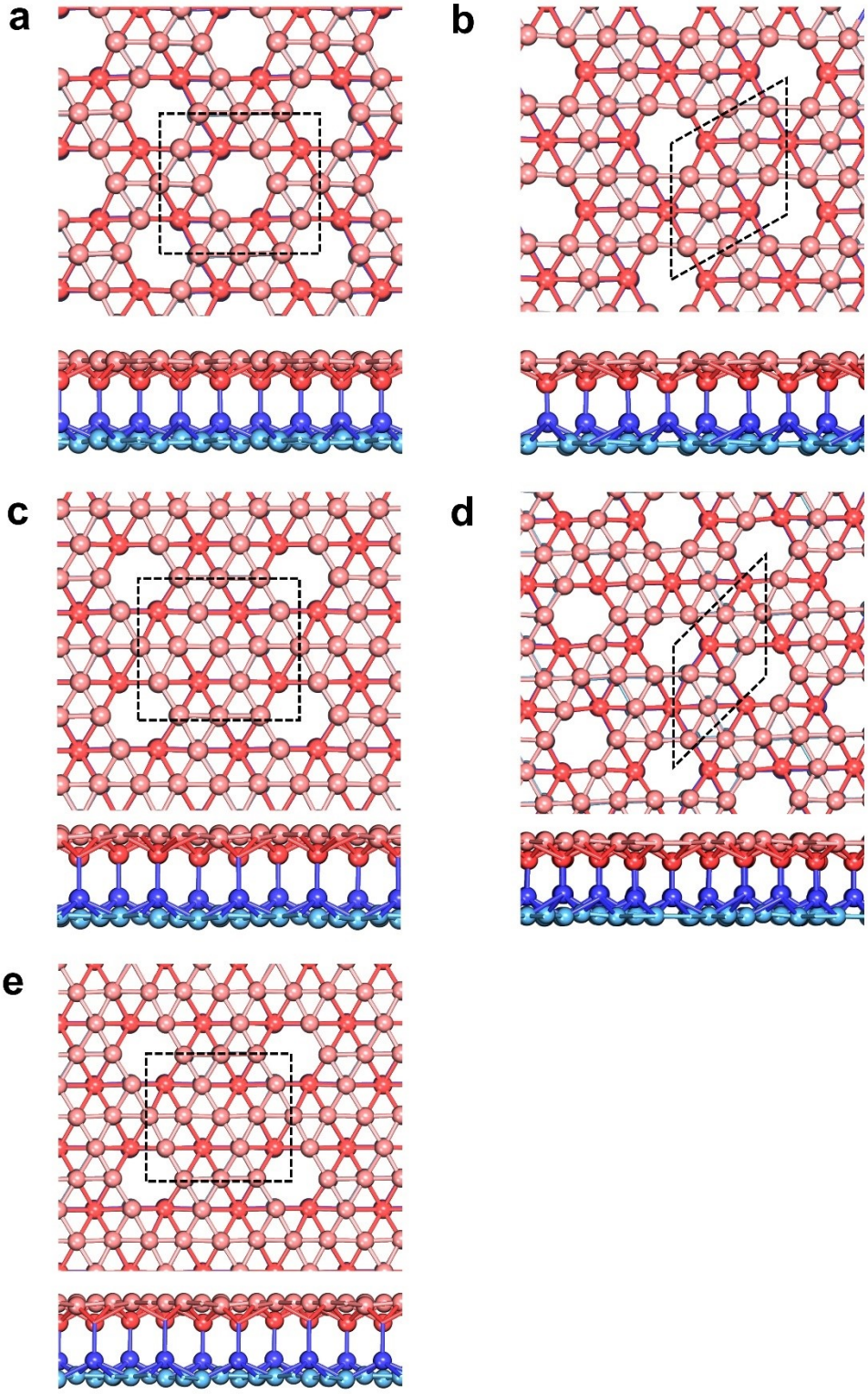

Figure S16. The top and side views of atomic patterns of (a) BL-α_1, (b) BL-α_5, (c) BL-α_6, (d) BL-α_7 and (e) BL-$\alpha_1\alpha$. by AA stacking. The pink and light cyan colors denote topmost and bottommost boron atoms, respectively. The upper and lower boron atoms for the formation of interlayer bonds are colored by red and dark cyan, respectively. The black dashed lines denote the unit cell.

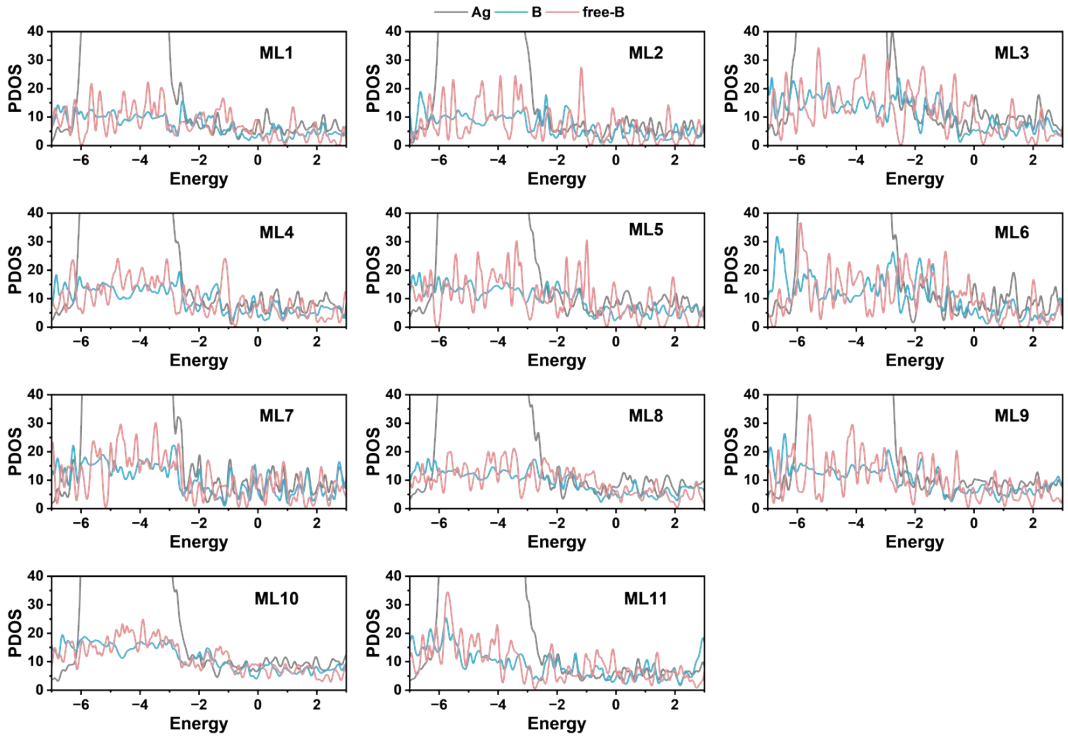

Figure S17. Projected density of states of borophene on Ag(100) and free-standing borophene for ML1 to ML11.

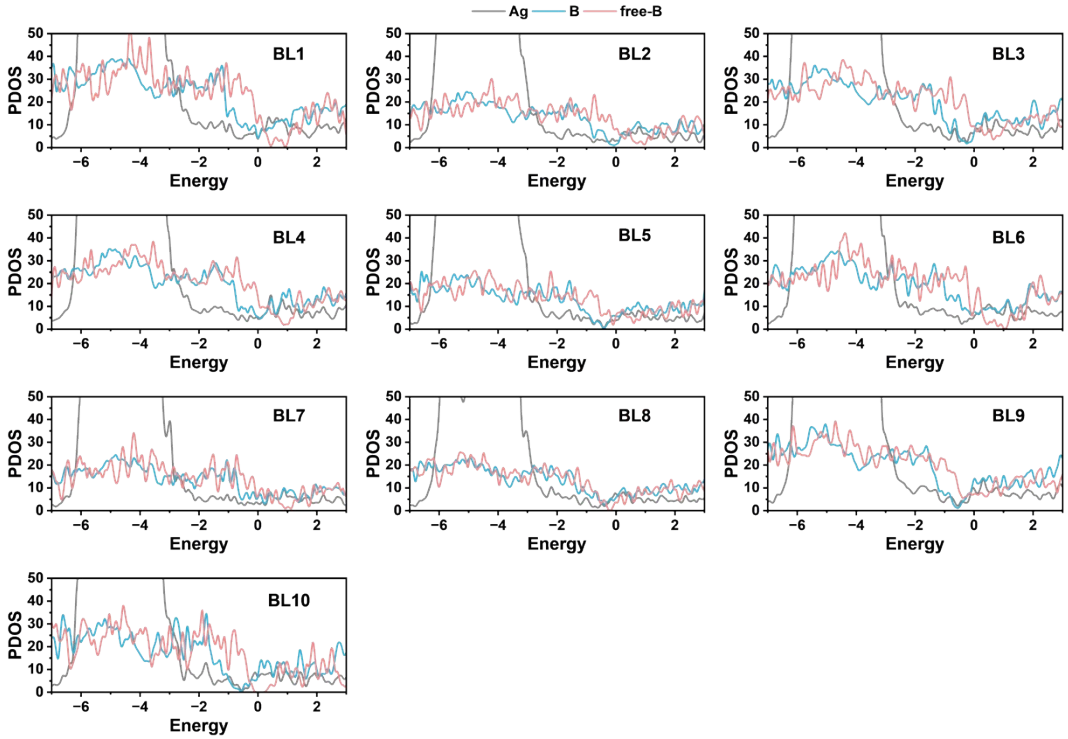

Figure S18. Projected density of states of borophene on Ag(100) and free-standing borophene for BL1 to BL10.

Table S3. The Name, boron coverage (*BC*), formation energy (E_{form}), lattice parameters (a, b, and γ), supercell types (referred as $r_{11}r_{21}r_{21}r_{22}$ in which r_{11}, r_{21}, r_{21} and r_{22} are elements of redefine matrix), multiple (*M*) values, structural components of unit cell, order parameters (OP_2, OP_4, OP_6), the averaged coordination number (ave_cn), averaged boron-boron bond length (ave_bond), energy (Δ*E*, meV/atom) and force (Δ*F*, eV/Å) error between the results from NN prediction and DFT calculations, averaged number of electrons obtained from Ag (C, e/atom), lattice mismatch (LM) and the energies of free-standing borophene (E_B, eV/atom) for ML1~ML10 in the first energy valley. The Structures denote the type and corresponding number of different borophene phases included in the unit cell of mixed-phase borophenes. The larger lattice constant is defined as a, the other is b.

Name	**ML1**	**ML2**	**ML3**	**ML4**	**ML5**	**ML6**	**ML7**	**ML8**	**ML9**	**ML10**
BC	2.7	2.737	2.8	2.72	2.769	2.857	2.667	2.731	2.778	2.793
E_{form}	-6.5371	-6.5325	-6.5320	-6.5316	-6.5299	-6.5256	-6.5220	-6.5200	-6.518	-6.5177
a	17.573	14.445	36.997	20.428	29.462	20.428	29.744	14.731	19.380	21.032
b	10.416	11.912	8.667	11.912	10.416	11.556	8.667	14.731	15.558	14.445
γ	65.8	112.8	51.3	59.0	135.0	81.9	61.0	90.0	131.6	52.8
Supercell	6-123	4-314	10803	55-14	10-2-23	7104	9503	51-15	63-52	7-243
M	20	19	30	25	26	28	27	26	27	29
Structures	**1**χ_3**1**β_{12}	**2**χ_3**1**β_{12}	**1**χ_3**2**β_{12}	**3**χ_3**1**β_{12}	**1**χ_3**1**β_{12}	β_{12}	χ_3	**1**χ_3**2**β_{12}-dislocation	**1**χ_3**1**β_{12} **1**R1	Hybridized -**2**χ_3**1**β_{12}
OP_2	1.9374	1.8473	1.8586	1.8376	1.8452	1.8749	1.8933	1.8544	1.8076	1.8805
OP_4	1.3583	1.3824	1.3802	1.3691	1.3643	1.3918	1.4022	1.3824	1.3374	1.3839
OP_6	1.0842	1.1836	1.1732	1.1652	1.1513	1.1856	1.1793	1.1835	1.1312	1.1572
ave_cn	4.67	4.62	4.71	4.59	4.67	4.8	4.5	4.73	4.8	4.62
ave_bond	1.716	1.696	1.694	1.712	1.724	1.69	1.709	1.716	1.714	1.685
Δ*E*	0.952	1.041	1.229	3.067	4.631	0.695	2.061	3.416	4.467	3.457
Δ*F*	0.176	0.289	0.299	0.246	0.279	0.202	0.242	0.275	0.21	0.269
C (e/atom)	0.0265	0.0298	0.0274	0.0274	0.0290	0.0262	0.0272	0.0268	0.0258	0.0270
LM	1.47%	1.61%	0.96%	2.19%	1.55%	1.34%	3.56%	2.34%	1.46%	1.74%
E_B	-6.2998	-6.2897	-6.2942	-6.2967	-6.2987	-6.2832	-6.2918	-6.2924	-6.2931	-6.2876

Table S4. The Name, boron coverage (*BC*), formation energy (E_{form}), supercell types (referred as $r_{11}r_{21}r_{21}r_{22}$ in which r_{11}, r_{21}, r_{21} and r_{22} are elements of redefine matrix), lattice parameters (a, b, and γ), multiple (*M*) values, structural components of unit cell, order parameters (OP_2, OP_4, OP_6), the averaged coordination number (ave_cn), averaged boron-boron bond length (ave_bond), energy (Δ*E*, meV/atom), force (Δ*F*, eV/Å) error between the results from NN prediction and DFT calculations, averaged number of electrons obtained from Ag (C, e/atom), lattice mismatch (LM) and the energies of free-standing borophene (E_B, eV/atom) for ML11~ML20 in the second energy valley. The Structures denote the type and corresponding number of different borophene phases included in the unit cell of mixed-phase borophenes. The larger lattice constant is defined as a, the other is b.

Name	**ML11**	**ML12**	**ML13**	**ML14**	**ML15**	**ML16**	**ML17**	**ML18**	**ML19**	**ML20**
BC	3.6	3.5	3.529	3.538	3.524	3.52	3.478	3.448	3.455	3.481
E_{form}	-6.5285	-6.528	-6.5258	-6.5247	-6.5207	-6.5203	-6.5179	-6.5174	-6.5173	-6.5172
a	16.846	16.343	16.846	18.499	19.380	20.428	23.292	23.292	14.445	23.292
b	14.731	11.556	11.912	12.920	11.912	11.912	11.912	14.445	12.920	12.257
γ	137.7	135.0	135.0	114.8	49.4	59.0	136.2	46.0	79.7	52.1
Supercell	53-51	4-404	5-3-14	54-42	6314	55-14	74-41	8134	43-24	8-133
M	20	16	17	26	21	25	23	29	22	27
Structures	δ_6	δ_6	δ_6	δ_6	δ_6	δ_6	δ_6	δ_6	δ_6	δ_6
OP_2	1.7079	1.7861	1.7993	1.7211	1.8124	1.7632	1.8972	1.9019	1.9094	1.9177
OP_4	1.0112	0.9787	1.0677	1.0327	1.0678	1.0472	1.1215	1.1441	1.1403	1.1414
OP_6	0.7283	0.7022	0.7707	0.7485	0.7685	0.7636	0.8076	0.8327	0.8332	0.8322
ave_cn	6	6	6	6.01	6.07	6.1	6	6	6.01	6.02
ave_bond	1.797	1.817	1.771	1.797	1.816	1.811	1.822	1.815	1.819	1.819
Δ*E*	1.602	1.552	4.473	0.552	1.752	1.781	1.758	2.802	1.785	1.444
Δ*F*	0.276	0.316	0.29	0.263	0.273	0.252	0.273	0.255	0.292	0.279
C (e/atom)	0.0108	0.0106	0.0090	0.0095	0.0077	0.0083	0.0075	0.0067	0.0085	0.0079
LM	0.64%	1.98%	0.16%	1.26%	2.07%	2.13%	3.30%	1.74%	3.01%	3.88%
E_B	-6.2439	-6.2439	-6.2450	-6.2439	-6.2437	-6.2423	-6.2445	-6.2440	-6.2429	-6.2445

Table S5. The Name, boron coverage (*BC*), formation energy (E_{form}), supercell types (referred as $r_{11}r_{21}r_{21}r_{22}$ in which r_{11}, r_{21}, r_{21} and r_{22} are elements of redefine matrix), lattice parameters (a, b, and γ), multiple (*M*) values, structural components of unit cell, order parameters (OP_2, OP_4, OP_6), the averaged coordination number (ave_cn), averaged boron-boron bond length (ave_bond), energy (ΔE, meV/atom) and force (ΔF, eV/Å) error between the results from NN prediction and DFT calculations, averaged number of electrons obtained from Ag (C, e/atom), lattice mismatch (LM) and the energies of free-standing borophene (E_B, eV/atom) for BL1~BL10 in the third energy valley. The Structures denote the type and corresponding number of different borophene phases included in the unit cell of mixed-phase borophenes. The larger lattice constant is defined as a, the other is b.

Name	**BL1**	**BL2**	**BL3**	**BL4**	**BL5**	**BL6**	**BL7**	**BL8**	**BL9**	**BL10**
BC	6.312	6.444	6.333	6.292	6.375	6.261	6.250	6.438	6.417	6.600
E_{form}	-6.597	-6.592	-6.588	-6.588	-6.587	-6.585	-6.584	-6.582	-6.579	-6.577
a	11.912	17.573	17.334	17.334	11.912	23.292	11.912	11.912	17.334	14.731
b	11.556	16.846	11.912	11.912	11.556	11.912	11.556	11.556	11.912	11.556
γ	76.0	49.6	76.0	76.0	76.0	438.	76.0	76.0	76.0	101.3
Supercell	4104	6135	6014	6014	4104	7-441	4104	4104	6014	5-104
M	16	27	24	24	16	23	16	16	24	20
Structures	**2**BL-α_5 **1**BL-α_1 **1**BL-$\alpha_1\alpha_6$	**3**BL-α_5 **3**BL-α_7	**3**BL-α_5 **1**BL-α_1 **2**BL-$\alpha_1\alpha_6$	**3**BL-α_5 **2**BL-α_1 **1**BL-$\alpha_1\alpha_6$	**2**BL-α_5 **2**BL-$\alpha_1\alpha_6$	BL-α_7	**2**BL-α_5 **2**BL-α_1	**2**BL-α_5 **1**BL-$\alpha_1\alpha_6$ **1**BL-α_6	**3**BL-α_5 **2**BL-$\alpha_1\alpha_6$ **1**BL-α_6	BL-α_5
OP_2	1.666	1.663	1.671	1.661	1.67	1.656	1.644	1.711	1.676	1.624
OP_4	0.793	0.791	0.795	0.792	0.794	0.787	0.784	0.803	0.795	0.771
OP_6	0.599	0.594	0.6	0.596	0.599	0.583	0.589	0.595	0.6	0.579
ave_cn	5.62	5.66	5.64	5.6	5.69	5.63	5.56	5.75	5.73	5.73
ave_bond	1.769	1.758	1.769	1.766	1.768	1.767	1.766	1.772	1.77	1.753
Δ*E*	1.894	2.222	1.909	2.39	2.142	2.418	2.303	1.212	1.77	1.774
Δ*F*	0.318	0.293	0.299	0.331	0.314	0.301	0.304	0.331	0.31	0.187
C (e/atom)	0.0103	0.0095	0.0099	0.0105	0.0110	0.0064	0.0106	0.0096	0.0096	0.0087
LM	0.95%	1.32%	0.84%	1.01%	1.11%	2.80%	1.12%	1.15%	0.82%	2.12%
E_B	-6.4654	-6.4807	-6.4684	-6.4616	-6.4701	-6.4726	-6.4534	-6.4823	-6.4780	-6.4955

References

1. G. Kresse and J. Hafner, Ab-Initio Molecular-Dynamics Simulation of the Liquid-Metal Amorphous-Semiconductor Transition in Germanium, *Phys. Rev. B*, 1994, **49**, 14251-14269.
2. G. Kresse and J. Furthmuller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, *Comp. Mater. Sci.*, 1996, **6**, 15-50.
3. P. E. Blochl, Projector Augmented-Wave Method, *Phys. Rev. B*, 1994, **50**, 17953-17979.
4. P. E. Blochl, P. Margl and K. Schwarz, Ab initio molecular dynamics with the projector augmented wave method, *Acs Sym. Ser.*, 1996, **629**, 54-69.
5. J. P. Perdew, K. Burke and M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.*, 1996, **77**, 3865-3868.
6. F. Li, P. Jin, D.-e. Jiang, L. Wang, S. B. Zhang, J. Zhao and Z. Chen, B80 and B101–103 clusters: Remarkable stability of the core-shell structures established by validated density functionalsa), *J. Chem. Phys.*, 2012, **136**.
7. S. Grimme, J. Antony, S. Ehrlich and H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *J. Chem. Phys.*, 2010, **132**.
8. D. Zhang, P. Yi, X. Lai, L. Peng and H. Li, Active machine learning model for the dynamic simulation and growth mechanisms of carbon on metal surface, *Nat. Commun.*, 2024, **15**, 344.
9. C. Shang and Z. P. Liu, Stochastic Surface Walking Method for Structure Prediction and Pathway Searching, *J. Chem. Theory. Comput.*, 2013, **9**, 1838-1845.
10. C. Shang, X. J. Zhang and Z. P. Liu, Stochastic surface walking method for crystal structure and phase transition pathway prediction, *Phys. Chem. Chem. Phys.*, 2014, **16**, 17845-17856.
11. S.-D. Huang, C. Shang, P.-L. Kang, X.-J. Zhang and Z.-P. Liu, LASP: Fast global potential energy surface exploration, *WIREs Comput. Mol. Sci.*, 2019, **9**, e1415.
12. P. J. Steinhardt, D. R. Nelson and M. Ronchetti, Bond-orientational order in liquids and glasses, *Phys. Rev. B*, 1983, **28**, 784-805.
13. X. J. Zhang, C. Shang and Z. P. Liu, Pressure-induced silica quartz amorphization studied by iterative stochastic surface walking reaction sampling, *Phys. Chem. Chem. Phys.*, 2017, **19**, 4725-4733.