Supporting Information

Efficient and Atomic-resolution Uncertainty Estimation for Neural Network Potentials Using Replica Ensemble

Wonseok Jeong,[1] Dongsun Yoo,[1] Kyuhyun Lee,[1] Jisu Jung[1] and Seungwu Han[1].*

[1]Department of Materials Science and Engineering, Seoul National University, Seoul 08826, Republic of Korea.

*hansw@snu.ac.kr

Details on DFT calculation

The DFT calculations in this work are carried by VASP employing a gradient-corrected exchange-correlation functional.[1-3] The energy cutoffs for the plane-wave basis is 350 eV and k points are sampled such that total energies and stress tensors converge to within 10 meV/atom and 10 kbar, respectively. We adopt spin-polarized calculations for structures containing Ni atoms.

Details on NNP training

Table S1. The detailed information of reference structures and RMSE for the validation set.

Structure type	Number of structures	Number of atoms	MD time (ps)	Interval (fs)	Number of training points	Temperature (K)	$RMSE_{energy}$ (meV/atom)	$RMSE_{force}$ (eV/Å)
c-Ni MD	100	32	6	60	3,200	500-1500	3.4	0.19
c-Si MD	100	64	6	60	6,400	500-1500	2.9	0.17
c-Ni static	126	4	-	-	504	0	4.7	0.00
c-Si static	126	8	-	-	1,008	0	3.6	0.11
δ-Ni_2Si MD	150	96	3	20	14,400	1000-2000	4.4	0.29
MnP-NiSi MD	150	64	3	20	9,600	1000-2000	8.1	0.45

α-$NiSi_2$ MD	150	96	3	20	14,400	1000-2000	5.4	0.42
δ-Ni_2Si static	40	96	-	-	3,840	0	1.8	0.22
MnP-NiSi static	40	64	-	-	2,560	0	3.2	0.44
α-$NiSi_2$ static	40	96	-	-	3,840	0	2.7	0.00
l-Ni_2Si MD	250	96	5	20	24,000	2000	5.1	0.45
l-NiSi MD	250	64	5	20	16,000	2000	7.0	0.41
l-$NiSi_2$ MD	250	96	5	20	24,000	2000	5.2	0.30
a-Ni_2Si MD	100	96	4	40	9,600	800	3.8	0.28
a-NiSi MD	100	64	4	40	6,400	800	7.4	0.43
a-$NiSi_2$ MD	100	96	4	40	9,600	800	8.3	0.50
l-Ni_2Si to *a*-Ni_2Si	150	96	3	20	14,400	2000-300	3.1	0.28
l-NiSi to *a*-NiSi	150	64	3	20	9,600	2000-300	6.8	0.41
l-$NiSi_2$ to *a*-$NiSi_2$	150	96	3	20	14,400	2000-300	5.5	0.49
Ni(001) surface MD	125	64	10	80	8,000	1000	2.1	0.16
Si(001) surface MD	120	128	4.8	40	15,360	1000	4.4	0.38
Interface (1) 300 K	50	74	10	200	3,700	300	2.7	0.23
Interface (1) 1000 K	375	74	6	16	27,750	1000	5.3	0.32
Interface (1) 1300 K	440	74	4.4	10	32,560	1300	5.1	0.36
Interface (1) 1500 K	600	74	6	10	44,400	1500	4.4	0.38
Interface (2) 1000 K	312	180	5	16	56,160	1000	2.5	0.30
Interface (2) 1300 K	200	180	2	10	36,000	1300	2.6	0.33
Interface (3) 1000 K	200	232	2	10	46,400	1000	3.0	0.28
Total	4,944				458,082		5.1	0.34

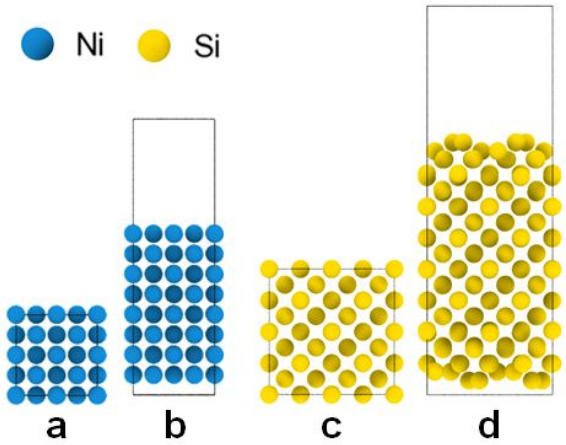

Figure S1. The training set structures. (a) Ni crystal. (b) Ni(001) surface slab. (c) Si crystal. (d) Si(001) surface slab.

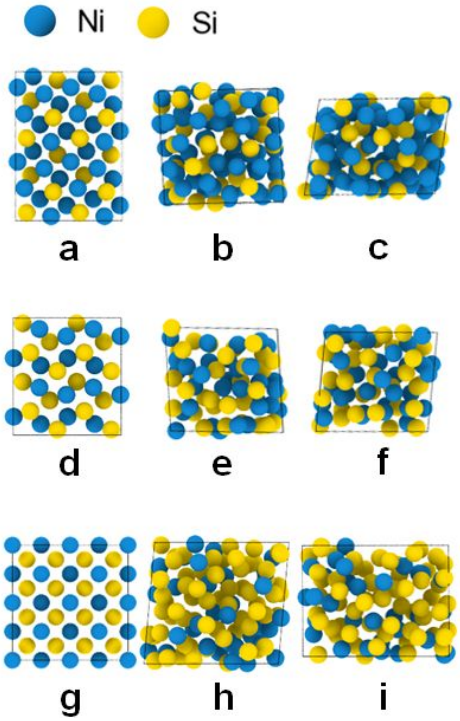

Figure S2. The training set structures. (a) δ-Ni_2Si crystal. (b) Liquid Ni_2Si. (c) Amorphous Ni_2Si. (d) NiSi crystal. (e) Liquid NiSi. (f) Amorphous NiSi. (g) α-$NiSi_2$ crystal. (h) Liquid $NiSi_2$. (i) Amorphous $NiSi_2$.

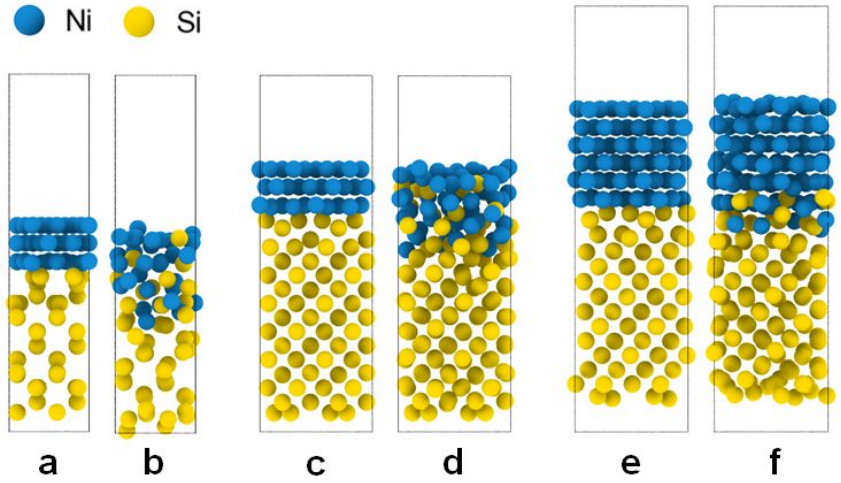

Figure S3. The training set structures. (a) Interface(1) at 0 K. (b) Interface(1) after 4.4 ps of 1300 K MD simulation. (c) Interface(2) at 0 K. (d) Interface(2) after 2 ps of 1300 K MD simulation. (e) Interface(3) at 0 K. (f) Interface(3) after 2 ps of 1000 K MD simulation.

From the total set in Table 1, 70% is randomly selected for training and the remaining 30% is used in model validation. The architecture of the reference NNP is 70-70-70-1 for both Ni and Si, with 20,022 parameters in total. We train the model by fitting to total energies and atomic forces from the DFT calculations. We also employ a weighting scheme aiming at uniform training of reference atomic structures.[4] Since the symmetry functions are highly correlated with each other, decorrelating them with principal component analysis (PCA) accelerates convergence in training. Thus, we transform symmetry function vectors by PCA without truncating dimensions. After the transformation, variances of all components are normalized (whitened). We also apply a L2 regularization to avoid overfitting. The atomic configurations are visualized with the OVITO package.[5]

Table S2. NNP training hyperparameters

Hyperparameter	Applied
NN structure	70-70-70-1
Optimizer	Adam
Learning rate	0.0001
Batch size	20
Force coefficient	0.1
L2 regulation coefficient	0.00000001
PCA whitening level	0.0001
GDF σ	0.05
Weight modifier	$x/(1+\exp(-bx+c))$
Weight modifier parameters for Si	$b = 1$, $c = 5$
Weight modifier parameters for Ni	$b = 1$, $c = 7.5$

NNP validation with materials properties

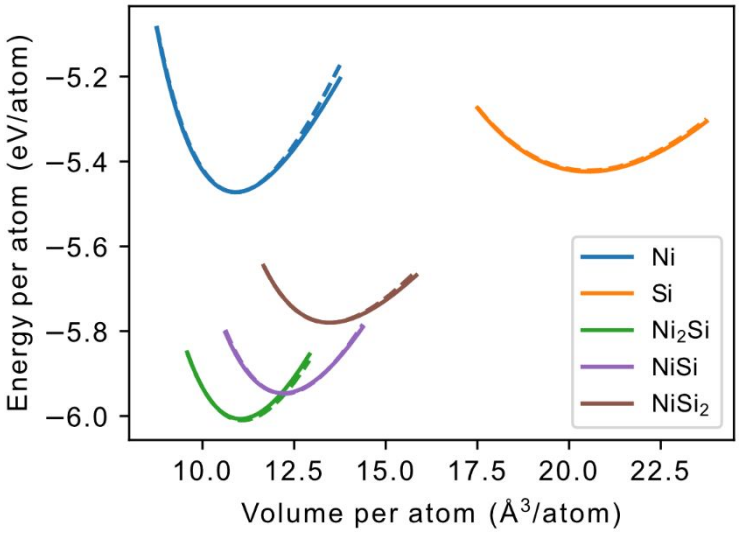

Figure S4. The equation of state calculated by DFT (solid lines) and NNP (dashed lines).

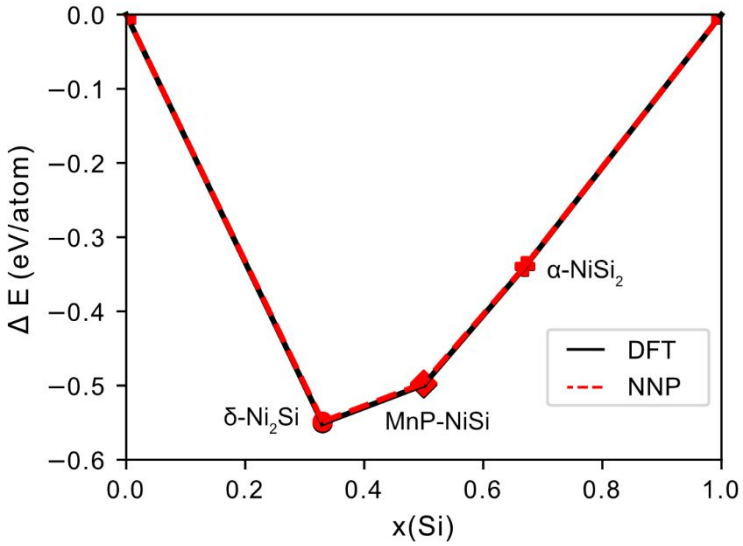

Figure S5. Convex hull of Ni_2Si, NiSi, $NiSi_2$ relative to Ni and Si calculated by DFT and NNP.

Three different Ni silicides amorphous structures are generated with DFT and NNP respectively such that each of them has an atomic number ratio corresponding to Ni_2Si, NiSi, and $NiSi_2$. The initial positions of the Ni and Si atoms are randomly spread at the cell, then 1.6 ps premelting process was run under DFT 5000 K MD with NPT ensemble. Uncorrelated structures from premelting procedures are chosen and the 2000 K NPT melting procedure was carried out for 10 ps with DFT and NNP separately. The amorphous structures are then obtained by separate DFT and NNP quenching procedure of 2000 K to 300 K simulations with the quenching rate of 425K/ps. Finally, the structures were relaxed until the residual force is less than 0.02 eV/Å. The total radial distribution function (RDF) of liquid structures are time-averaged for the snapshots from the melting procedure. To obtain the RDF of

amorphous structures, we carried out the 800 K DFT NPT MD simulations with melt and quenched structure for 4 ps and the total RDF was time-averaged.

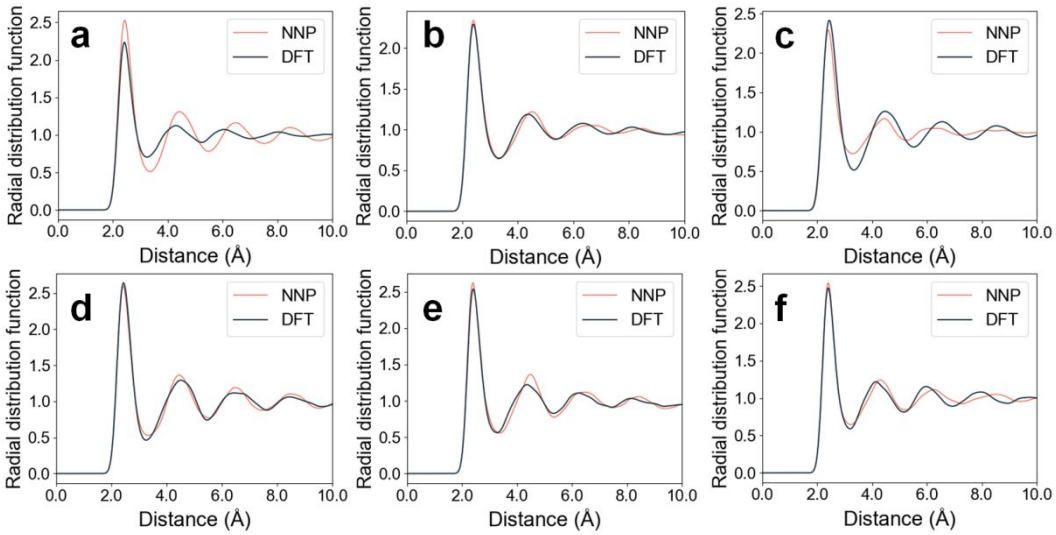

Figure S6. The time-averaged total radial distribution function (RDF) of liquid Ni silicides which have stoichiometry corresponding (a) Ni_2Si. (b) NiSi. (c) $NiSi_2$. The time-averaged total RDF of and amorphous Ni silicides which have stoichiometry corresponding (d) Ni_2Si. (e) NiSi. (f) $NiSi_2$.

Table S3. Total energy per atom and volume per formula unit of Ni silicides obtained by separate melt and quench procedure. The values inside the parentheses are the percent error of NNP energy and cell volume with respect to reference DFT value.

System	Method	Energy (eV/atom)	Volume ($Å^3$ /formula unit)
a-Ni_2Si	DFT	-5.88	34.1
	NNP	-5.86 (-0.4 %)	34.6 (1.5 %)
a-NiSi	DFT	-5.75	24.7
	NNP	-5.84 (1.6 %)	23.8 (-3.7 %)
a-$NiSi_2$	DFT	-5.57	42.0
	NNP	-5.63 (1.1 %)	40.4 (-3.9 %)

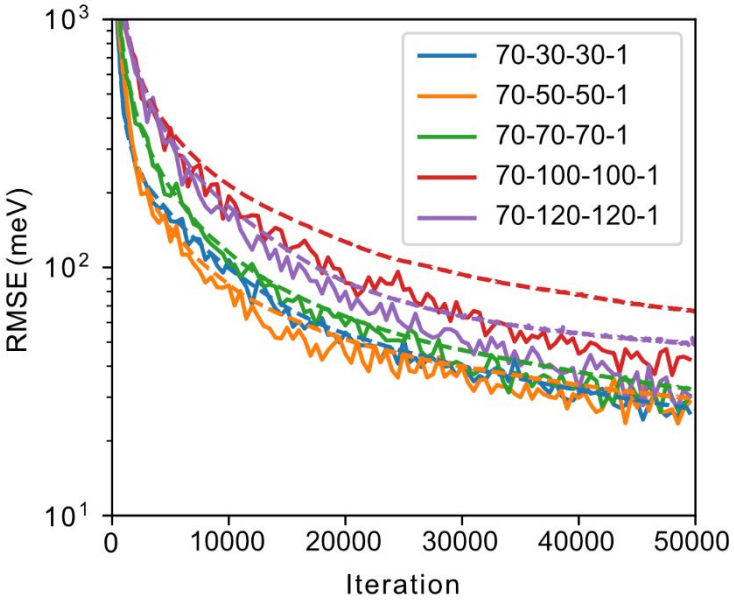

Figure S7. The training set (solid lines) and validation set (dashed lines) root mean square error (RMSE) of atomic energy during the replica NNP training iterations.

r-NNP validation with materials properties

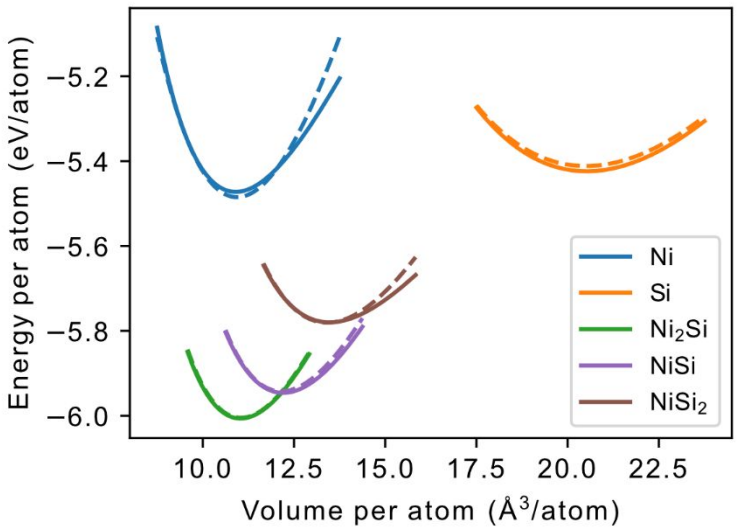

Figure S8. The equation of state calculated by DFT (solid lines) and r-NNP (dashed lines).

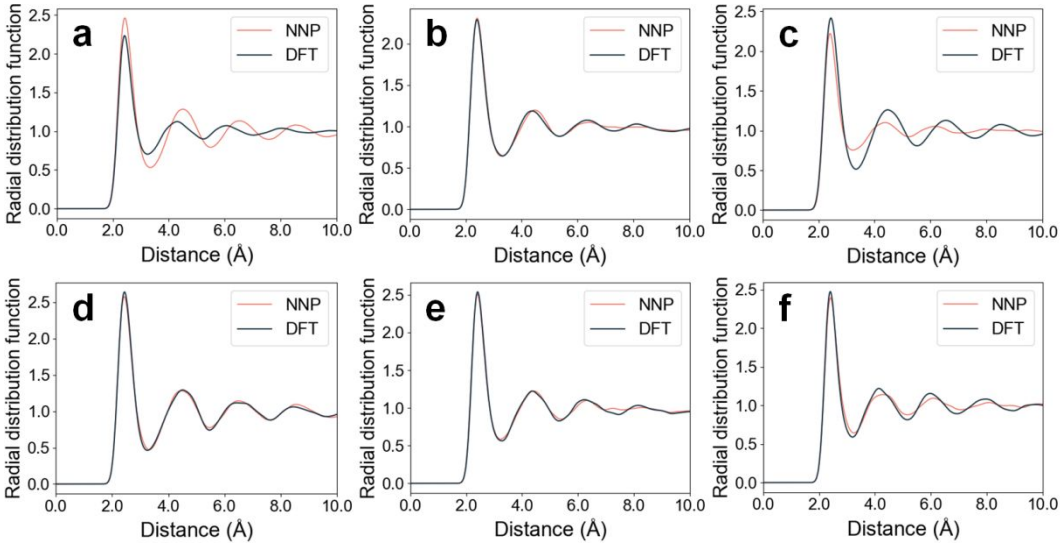

Figure S9. The time-averaged total radial distribution function (RDF) of liquid Ni silicides which have stoichiometry corresponding (a) Ni_2Si. (b) NiSi. (c) $NiSi_2$. The time-averaged total RDF of and amorphous Ni silicides which have stoichiometry corresponding (d) Ni_2Si. (e) NiSi. (f) $NiSi_2$.

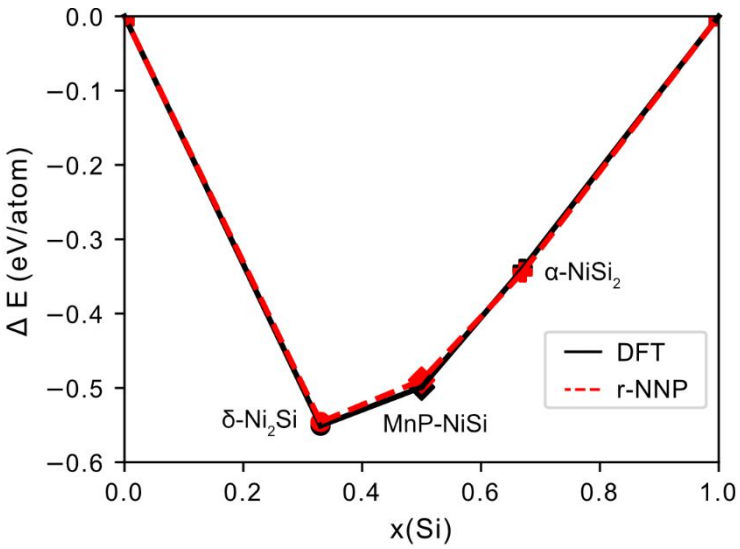

Figure S10. Convex hull of Ni_2Si, NiSi, $NiSi_2$ relative to Ni and Si calculated by DFT and r-NNP.

Table S4. Total energy per atom and volume per formula unit of Ni silicides obtained by separate melt and quench procedure. The values inside the parentheses are the percent error of r-NNP energy and cell volume with respect to reference DFT value.

System	Method	Energy (eV/atom)	Volume (Å^3 /formula unit)
a-Ni_2Si	DFT	-5.88	34.1
	NNP	-5.86 (-0.3 %)	34.5 (1.1 %)
a-NiSi	DFT	-5.75	24.7
	NNP	-5.76 (0.2 %)	24.6 (-0.2 %)
a-$NiSi_2$	DFT	-5.57	42.0
	NNP	-5.54 (-0.5 %)	41.1 (-2.2 %)

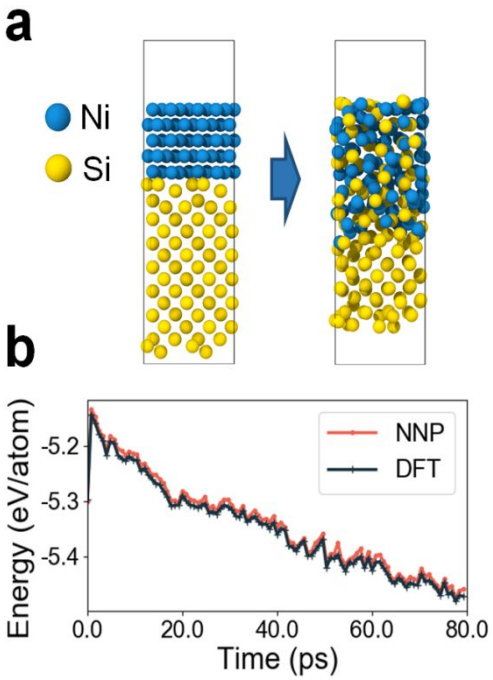

Figure S11. (a) Relatively small Ni-Si interface structure with 100 Ni atoms and 120 Si atoms. (b) Comparison of the DFT and r-NNP total energies per atom along MD trajectories. The red line represents the energy from 1000 K NNP MD. The black line represents the DFT energy of r-NNP MD. The trajectories are sampled by an 800 fs interval.

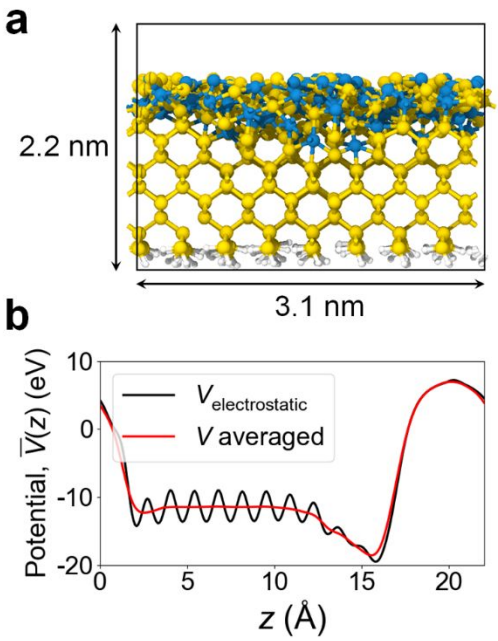

Figure S12. (a) Model structure of the Ni_xSi_y/Si interface slab for the Schottky barrier height estimation. The dangling bonds of bulk-terminated Si surface are passivated with H atoms. (b) The black line represents the electrostatic potential of the Ni_xSi_y/Si interface slab averaged to the z-axis. The red line represents the macroscopic average of the potential with a sampling window set to the layer spacing of the Si(001) slab.

REFERENCES

(1) Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Liquid Metals. *Phys. Rev. B.* **1993**, *47*, 558–561.

(2) Blöchl, P. E. Projector Augmented-wave Method. *Phys. Rev. B.* **1994**, *50*, 17953–17979.

(3) Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(4) Jeong, W.; Lee, K.; Yoo, D.; Lee, D.; Han, S. Toward Reliable and Transferable Machine Learning Potentials: Uniform Training by Overcoming Sampling Bias. *J. Phys. Chem. C* **2018**, *122*, 22790−22795.

(5) Stukowski, A. Visualization and Analysis of Atomistic Simulation Data with OVITO–the Open Visualization Tool. *Model. Simul. Mater. Sci. Eng.* **2010**, *18*, 015012.