Supplementary Material: Database Construction for Two-Dimensional Material-Substrate Interfaces

Xian-Li Zhang(张现利)[1,2#], Jinbo Pan(潘金波)[1#], Xin Jin(金鑫)[1,2], Yan-Fang Zhang(张艳芳)[2,1], Jia-Tao Sun(孙家涛)[3], Yu-Yang Zhang(张余洋)[2,4] , and Shixuan Du(杜世萱)[1,2,4,5*]

[1]Beijing National Laboratory for Condensed Matter Physics, and Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China

[2]School of Physics, University of Chinese Academy of Sciences, Beijing 100049, China

[3]School of Information and Electronics, MIIT Key Laboratory for Low-Dimensional Quantum Structure and Devices, Beijing Institute of Technology, Beijing 100081, China

[4]CAS Center for Excellence in Topological Quantum Computation, University of Chinese Academy of Sciences, Beijing 100190, China

[5]Songshan Lake Materials Laboratory, Dongguan, Guangdong 523808, China

[#]Xian-Li Zhang and Jinbo Pan contributed equally to this work.

[*]Shixuan Du: sxdu@iphy.ac.cn

Matrix and index notations

In general, the primitive vectors $\boldsymbol{a}_1, \boldsymbol{a}_2$ of substrate surface or 2D material and those $\bar{\boldsymbol{a}}_1, \bar{\boldsymbol{a}}_2$ of the superlattice are related by

$$\begin{pmatrix} \bar{\boldsymbol{a}}_1 \\ \bar{\boldsymbol{a}}_2 \end{pmatrix} = \boldsymbol{M} \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{pmatrix},$$

where

$$\boldsymbol{M} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$$

is a 2×2 matrix consist with four integers. Under the prerequisite of isometric strain, the two base vectors of superlattice must have the same length (i.e., $|\bar{\boldsymbol{a}}_1| = |\bar{\boldsymbol{a}}_2|$) and keep unchanged angle with primitive cell. That is, the four matrix elements are not independent of each other. In fact, the *index notation* (*m*, *n*), which indicates $\bar{\boldsymbol{a}}_1 = m\boldsymbol{a}_1 + n\boldsymbol{a}_2$ (shown in Fig. S2a), is an alternative and equivalent representation to describe the superlattice. The relation between these two notations can be expressed as $m_{11} = m$, $m_{12} = n$, $m_{21} = -n$ and $m_{22} = \begin{cases} m & for\ square\ lattices \\ m+n & for\ hexagonal\ lattices \end{cases}$.

Calculation Methods

The DFT calculations were performed by using projector augmented wave (PAW) method[1, 2] as implemented in the Vienna ab-initio simulation package (VASP).[3] The Perdew-Burke-Ernzerhof (PBE) with generalized gradient approximation (GGA)[4] were used for the exchange-correlation functional. The vdW correction method of Grimme is considered in the calculations.[5] The energy cut-off of plane wave basis was set as the maximum energy of default cut-off in the PAW potential files of all related elements. The momentum point grid $N_k \times N_k \times 1$ based on Γ-center sampling method were dense enough (typically $N_k \times l > 50$Å) for specific interface system (Figure S4). The vacuum spacing larger than 15 Å was set to eliminate the interlayer interactions along the *z* direction due to the periodic boundary conditions. To simulate the bulk state, three atomic layers of substrate were used in the calculation and the bottom two layer were fixed in the structural relaxation. The atomic structures were relaxed until the force on each unconstrained atom is less than 0.01 eV/Å and the energy convergence value of electronic iteration is set as 10^{-5} eV.

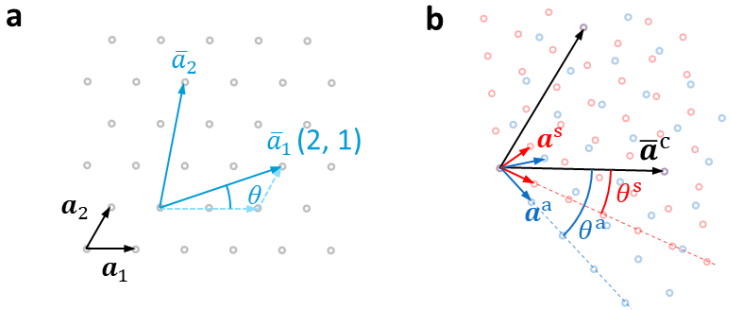

Figure S1. Definition of index notation and twist angle. **(a)** Schematic of (*m*, *n*) notation of supercell. The gray circles represent the lattice points of unitcell. **(b)** Illustration of adsorbate on substrate with a configuration of (1, 3)@(3, 2). The blue and red circles represent the unitcell lattice points of adsorbate and substrate, respectively. The $\boldsymbol{a}^a$, $\boldsymbol{a}^s$ and $\boldsymbol{a}^c$ are basis vectors of adsorbate, substrate and their composite structure, respectively. The θ^a and θ^s are angles between $\boldsymbol{a}^a$, $\boldsymbol{a}^s$ and $\bar{\boldsymbol{a}}^c$, respectively. The twist angle θ is defined as $\theta = \theta^a - \theta^s$.

Table S1. The effective range of twist angle under different symmetries of adsorbate and substrate. Here S_{R} and S_{M} are rotation and mirror symmetries, respectively. The superscript of a and s indicate the adsorbate and substrate, respectively. $LCM\{S_{\mathrm{R}}^{\mathrm{a}}, S_{\mathrm{R}}^{s}\}$ are the lowest common multiple (LCM) of $S_{\mathrm{R}}^{\mathrm{a}}$ and $S_{\mathrm{R}}^{\mathrm{s}}$.

$LCM\{S_{\mathrm{R}}^{\mathrm{a}}, S_{\mathrm{R}}^{\mathrm{s}}\}$	$(S_{\mathrm{R}}^{\mathrm{a}}, S_{\mathrm{R}}^{\mathrm{s}}) or (S_{\mathrm{R}}^{\mathrm{s}}, S_{\mathrm{R}}^{\mathrm{a}})$	$S_{\mathrm{M}}^{\mathrm{a}} \wedge S_{\mathrm{M}}^{\mathrm{s}}$	Range of twist angle
1	(1, 1)	N	$[0°, 360°)$
		Y	$[0°, 180°]$
2	(1, 2) (2, 2)	N	$[0°, 180°)$
		Y	$[0°, 90°]$
3	(1, 3) (3, 3)	N	$[0°, 120°)$
		Y	$[0°, 60°]$
4	(1, 4) (2, 4) (4, 4)	N	$[0°, 90°)$
		Y	$[0°, 45°]$
6	(1, 6) (2, 3) (2, 6) (3, 6) (6, 6)	N	$[0°, 60°)$
		Y	$[0°, 30°]$

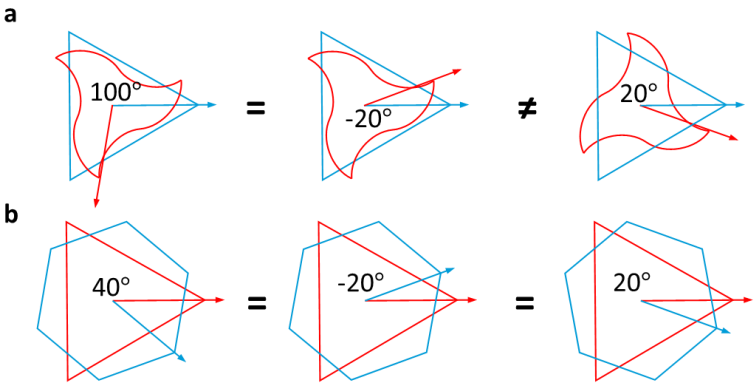

Figure S2. Schematics for the reduction of the twist angle out of range. **(a)** The effective twist angle range is [0 °, 120 °) due to the C_3 and C_{3v} symmetries of two materials, respectively. **(b)** The two materials possess C_{3v} and C_{6v} symmetries, respectively. Therefore, the effective twist angle is in the range of [0 °, 30 °]. The blue and red indicate the symmetries of two materials.

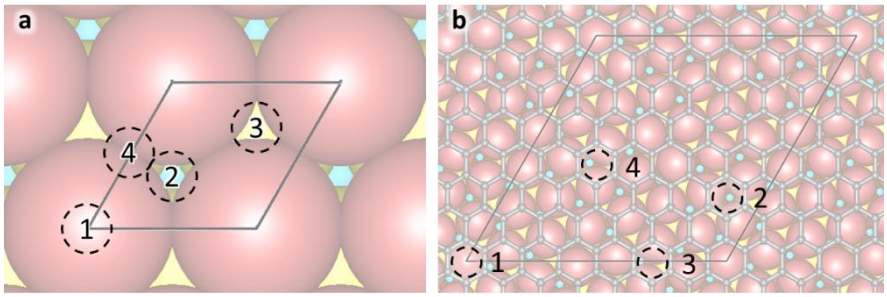

Figure S3. Various adsorption sites are considered in the workflow. **(a)** Four adsorption sites on the hexagonal substrate. **(b)** The adsorption sites mentioned in (a) are implicitly included in the large supercell (typically ≥15 Å).

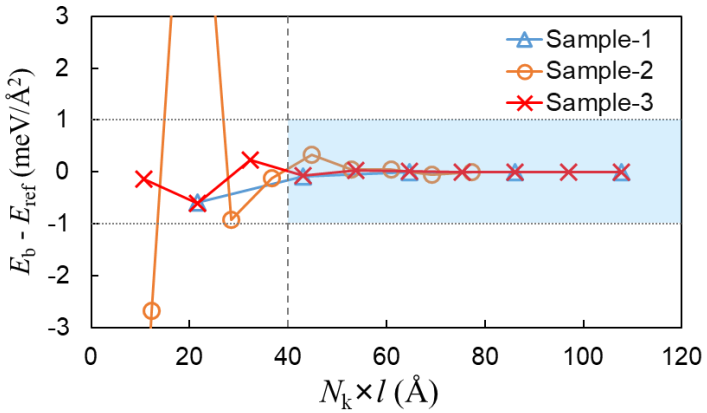

Figure S4. Typical K-sample test results of Sb@$PdTe_2$. The K points are taken dense enough (gray area) to ensure that binding energies of various structures converge to 1 meV/Å^2.

Table S2. Several experimental configurations and their calculated binding energy of graphene on Ir(111). The ranking *R* of specific structure sorted by binding energy is nearly independent of the functional.

M^{Gr}	$M^{Ir(111)}$	Gr@Ir(111)	θ (°)	PBE D3		LDA		opt-b88	
				E_b (meV/Å^2)	$R_{PBE\ D3}$	E_b (meV/Å^2)	R_{LDA}	E_b (meV/Å^2)	$R_{opt-b88}$
(4, 0)	(1, 3)	4@√13	13.90	38	1	21	2	31	3
(10, 0)	(9, 0)	10@9	0.00	37	2	21	1	31	2
(3, 4)	(5, 1)	√37@√31	25.77	30	12	15	12	22	13
(4, 3)	(5, 1)	√37@√31	16.34	30	13	14	13	22	14
(3, 2)	(4, 0)	√19@4	23.41	29	15	13	16	24	11
(3, 0)	(1, 2)	3@√7	19.11	25	16	8	19	19	17
(2, 0)	(1, 1)	2@√3	30.00	-9	36	-28	38	-4	33

Table S3. Top 20 configurations of buckled arsenene on Ag(111). The experimental formations were labeled in orange background. The binding energies of buckled arsenene (b-As) on Ag(111) are significantly higher than those of flat arsenene (f-As) (shown in Table S1), which indicates that arsenene on Ag(111) surface is more energy favorable to form a buckled structure.

$\boldsymbol{M}^{\mathrm{b-As}}$	$\boldsymbol{M}^{\mathrm{Ag(111)}}$	b-As@Ag(111)	l (Å)	Δ	θ (°)	E_{b} (meV/Å^2)
(4, 0)	(5, 0)	4@5	14.45	0.06%	0.00	59
(0, 4)	(5, 0)	4@5	14.45	0.06%	60.00	59
(-4, 6)	(1, 6)	√28@√43	18.94	-0.80%	48.48	59
(4, 2)	(1, 1)	√28@√43	18.94	-0.80%	11.52	59
(-2, 6)	(1, 6)	√28@√43	18.94	-0.80%	26.70	59
(2, 4)	(6, 1)	√28@√43	18.94	-0.80%	33.30	59
(-1, 6)	(4, 4)	√31@√48	20.02	-0.39%	38.95	59
(0, 5)	(5, 2)	5@√39	18.04	-0.02%	43.90	59
(0, 5)	(2, 5)	5@√39	18.04	-0.02%	16.10	59
(1, 5)	(4, 4)	√31@√48	20.02	-0.39%	21.05	59
(-2, 4)	(2, 3)	√12@√19	12.59	0.72%	53.41	59
(2, 2)	(3, 2)	√12@√19	12.59	0.72%	6.59	59
(-3, 6)	(1, 6)	√27@√43	18.94	1.02%	37.59	59
(3, 3)	(6, 1)	√27@√43	18.94	1.02%	22.41	59
(1, 5)	(7, 0)	√31@7	20.22	0.64%	51.05	58
(5, 1)	(7, 0)	√31@7	20.22	0.64%	8.95	58
(1, 5)	(5, 3)	√31@7	20.22	0.64%	29.26	58
(-1, 6)	(3, 5)	√31@7	20.22	0.64%	30.74	58
(-1, 6)	(5, 3)	√31@7	20.22	0.64%	47.16	58
(1, 5)	(3, 5)	√31@7	20.22	0.64%	12.84	58

Table S4. Top 10 configurations of flat arsenene on Ag(111).

$\boldsymbol{M}^{\mathrm{f-As}}$	$\boldsymbol{M}^{\mathrm{Ag(111)}}$	f-As@Ag(111)	l (Å)	Δ	θ (°)	E_{b} (meV/Å^2)
(1, 4)	(5, 3)	√21@7	20.22	0.98%	27.32	39
(2, 2)	(5, 0)	√12@5	14.45	-4.59%	30.00	39
(4, 0)	(3, 4)	4@√37	17.57	0.52%	25.28	38
(2, 0)	(3, 0)	2@3	8.67	-0.84%	0.00	37
(1, 3)	(4, 2)	√13@√28	15.29	-2.98%	27.00	37
(1, 3)	(2, 4)	√13@√28	15.29	-2.98%	5.21	36
(2, 1)	(4, 0)	√7@4	11.56	-0.06%	19.11	35
(2, 3)	(6, 1)	√19@√43	18.94	-0.55%	29.00	35
(3, 2)	(6, 1)	√19@√43	18.94	-0.55%	15.82	34
(3, 0)	(2, 3)	3@√19	12.59	-3.95%	23.41	34

[1] Blöchl P E 1994 *Phys. Rev. B* 50 17953

[2] Kresse G and Joubert D 1999 *Phys. Rev. B* 59 1758

[3] Kresse G and Furthmüller J 1996 *Phys. Rev. B* 54 11169

[4] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* 77 3865

[5] Grimme S, Antony J, Ehrlich S and Krieg H 2010 *The Journal of Chemical Physics* 132