CHIN. PHYS. LETT. Vol. 38, No. 6 (2021) 066801

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 523808, China*

(Received 9 April 2021; accepted 25 April 2021; published online 25 May 2021)

Interfacial structures and interactions of two-dimensional (2D) materials on solid substrates are of fundamental importance for fabrications and applications of 2D materials. However, selection of a suitable solid substrate to grow a 2D material, determination and control of 2D material-substrate interface remain a big challenge due to the large diversity of possible configurations. Here, we propose a computational framework to select an appropriate substrate for epitaxial growth of 2D material and to predict possible 2D material-substrate interface structures and orientations using density functional theory calculations performed for all non-equivalent atomic structures satisfying the symmetry constraints. The approach is validated by the correct prediction of three experimentally reported 2D material-substrate interface systems with only the given information of two parent materials. Several possible interface configurations are also proposed based on this approach. We therefore construct a database that contains these interface systems and has been continuously expanding. This database serves as preliminary guidance for epitaxial growth and stabilization of new materials in experiments.

DOI: 10.1088/0256-307X/38/6/066801

In past decades, various high-throughput calculation frameworks have been designed to discover materials with superior and tailored properties.[1−15] They are now becoming increasingly important for discovery of new functional materials. The automated FLOW (AFLOW)[16,17] package provides a groundbreaking software framework for high-throughput calculations of structural and electronic properties of inorganic crystals. The Materials Project[18,19] is another milestone materials database containing hundreds of thousands of compounds. Due to distinct mechanical and electronic properties, 2D materials have attracted broad interest, and many high-throughput calculations have also been performed.[20−27] More than a thousand of 2D materials, which are easily or potentially exfoliatable from bulk materials, were reported by Mounet *et al.*[24] The Computational 2D materials Database (C2DB),[27] which is developed by Hasstrup, contains the stability, electronic, magnetic and optical properties of thousands of 2D materials distributed across tens of different crystal structures. All these databases provide convenient tools to search for novel functional materials in batteries, electrocatalysts, etc.

Substrates, most of the time, play non-negligible roles in fabrications and applications of 2D materials. For a 2D material epitaxially grown on a substrate, the lattice mismatch is typically unavoidable, leading to the lattice dilation and compression of the component.[28] Meanwhile, physical properties of these atomically thin materials are usually sensitive to strain.[29−32] In addition, coupling between a 2D material (adsorbate) and a substrate may affect intrinsic physical properties of the 2D material. For example, the Dirac cone of graphene (Gr) is distorted while grown on some metallic or semiconducting substrates.[33−35] However, so far, the statistical effects from the supporting substrates including atomic configurations and electronic properties have not yet been fully considered,[28,33−38] although all the effects coming from the specific configurations of

Supported by the National Key R&D program of China (Grant Nos. 2019YFA0308500, 2020YFA0308800, and 2016YFA0202300), the National Natural Science Foundation of China (Grant Nos. 51922011, 61888102, and 11974045), the Strategic Priority Research Program of Chinese Academy of Sciences (Grant Nos. XDB30000000 and XDB28000000), Beijing Institute of Technology Research Fund Program for Young Scholars, and the Fundamental Research Funds for the Central Universities.

†Xian-Li Zhang and Jinbo Pan contributed equally to this work.

*Corresponding author. Email: sxdu@iphy.ac.cn

© 2021 Chinese Physical Society and IOP Publishing Ltd

adsorbates on a substrate (e.g., twist angle, strain, buckled height) are ubiquitous, which hinders further exploration of new materials on substrates.

We note that there are several groups that reported high-throughput calculations of interfacial systems including solid/solid hetero-structures, solid/implicit solvents systems, nanoparticle/ligands systems,[39] computational framework to select optimal substrates for epitaxial growth of polymorphs,[40] and a method for predicting solid/solid interface structures.[41] However, the database of 2D materials on substrates is still in its infancy.

In this Letter, we present a high-throughput calculation workflow aiming to predict the configurations of 2D materials on substrates beyond experiments. A series of unique and irreducible interfacial structures are generated under biaxial strains and symmetry constraints. High-throughput density functional theory (DFT) calculations are performed to obtain the optimized structures and binding energies. The stable configurations that can be synthesized experimentally are predicted based on their binding energies. This method has been proven to be valid by various 2D-material-substrate interfaces, including Gr@Ru(0001),[42–44] Gr@Ir(111),[45,46] arsenene@Ag(111),[47] and extended to 2D heterostructure antimonene@$PdTe_2$,[48] which have been reported in experiments. Based on these successful examples, a database,[49] including several new materials, is created. The calculations on new configurations containing 2D material/substrate interfaces and 2D-2D interfaces are ongoing and will be included in the database in the future. We believe these results will improve our ability to explore and design 2D materials on substrates and can be extended to construct 2D heterostructures which exhibit new and unique properties.

The workflow for the interface database construction is illustrated in Fig. 1. Two input databases, which contain the potential 2D materials, and substrates are constructed from an existing materials database, i.e., C2DB,[27] Materials Project.[18,19]

At the start of the workflow, the adsorbate (a 2D material) and a substrate are chosen from the input databases. The bulk phase of a substrate is cleaved along a high Miller index surface to fit the adsorbate. In fact, a 2D material tends to form more uniform sheets on a symmetry matched surface.[50–52] The 2D material is compressed or stretched biaxially to fit the lattice of the substrate and multiple twist angles are also allowed.

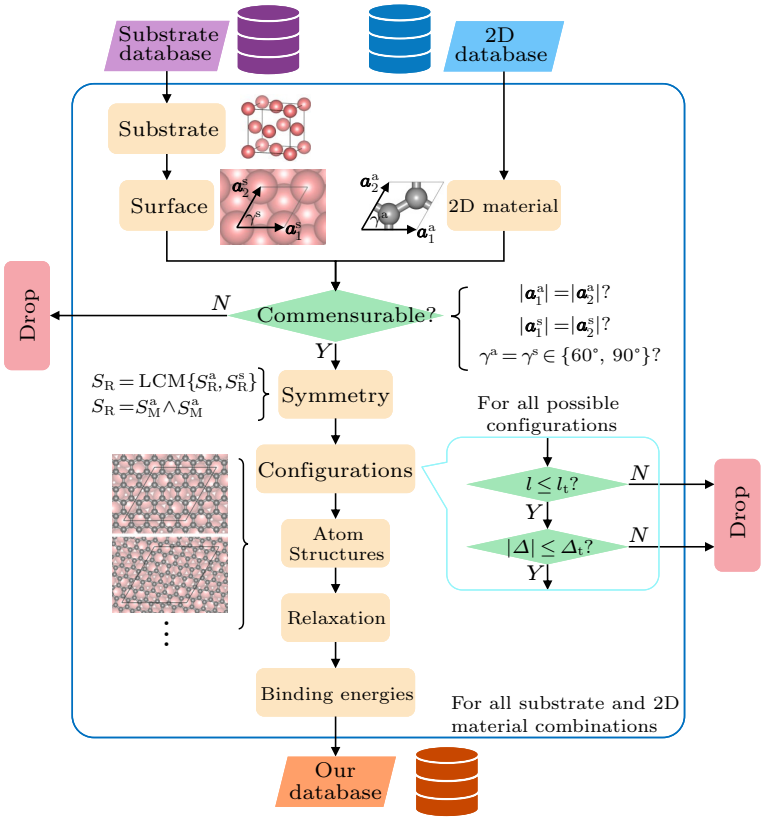

Fig. 1. Workflow for 2D-surface interface structures prediction. The interfaces of the substrate and the adsorbate, selected from the two input databases, were filtered by the commensurability and symmetry requirements. The DFT calculations were performed to obtain the stable structures and binding energies.

Under these prerequisites, the lattice of the cleaved substrate and adsorbate should meet (shown in Fig. 1)

$$|\boldsymbol{a}_1^{\rm a}| = |\boldsymbol{a}_2^{\rm a}|, \quad |\boldsymbol{a}_1^{\rm s}| = |\boldsymbol{a}_2^{\rm s}|, \quad \gamma_{\rm a} = \gamma_{\rm s} \in \{60^\circ, 90^\circ\},$$

where $|\boldsymbol{a}_1^{\rm a}|$ and $|\boldsymbol{a}_2^{\rm a}|$ represent the lengths of vectors of the adsorbate, $|\boldsymbol{a}_1^{\rm s}|$ and $|\boldsymbol{a}_2^{\rm s}|$ are the lengths of vectors of the substrate, $\gamma_{\rm a}$ and $\gamma_{\rm s}$ are the angles between two adsorbate vectors and between two substrate vectors, respectively. That is, the substrate and the adsorbate should be the same 2D Bravais lattices, hexagonal or square. In the following, we use the hexagonal lattices as examples.

We then generated the superlattices for the adsorbate and substrate, respectively. Due to the constraints of the commensurate supercell, the index notation $\boldsymbol{M} = (m, n)$ is used to describe the superlattice, where m and n are two integers. The details about the index notation are discussed in the Supplementary Material. The relative length L can be expressed as $L = |\bar{\boldsymbol{a}}_1|/|\boldsymbol{a}_1|$, where $|\bar{\boldsymbol{a}}_1|$ is the vector length of supercell. The large value of irrational L is rarely reported in literature. Thus, we only keep those configurations with integer L when L is greater than $\sqrt{50}$ in the following.

We then consider the constraint conditions of the symmetries of the adsorbate and the substrate. Different from a 3D lattice, only the rotation axes perpendicular to the surface and the mirror plane normal to the surface are allowed in the 2D system. If the mirror symmetry does not exist in any of the two materials, the effective range of twist angle can be described as $[0^\circ, \frac{360^\circ}{\mathrm{LCM}(S_{\rm R}^{\rm a}, S_{\rm R}^{\rm s})})$, where $\mathrm{LCM}(S_{\rm R}^{\rm a}, S_{\rm R}^{\rm s})$ is the lowest common multiple (LCM) of rotation symmetries of the adsorbate ($S_{\rm R}^{\rm a}$) and the substrate ($S_{\rm R}^{\rm s}$) (shown in Table S1 and Fig. S2). In contrast, if the mirror symmetry exists in both the substrate and the adsorbate, the effective range of the twist angle could further reduce to $[0^\circ, \frac{180^\circ}{\mathrm{LCM}(S_{\rm R}^{\rm a}, S_{\rm R}^{\rm s})}]$. It should be noted that the range of the twist angle is a right-open interval in the former case and a close interval in the latter.

Subsequently, a series of configurations are generated by combining all possible superlattices of the adsorbate and the substrate. However, the simple combination operation takes a large number of configurations. Some of them are unreasonable or computationally expensive. To balance the computational accuracy and cost, the oversized system is excluded when the lattice constant l is larger than the limitation $l_{\rm t}$ (typical value is 30 Å). Noteworthily, we assume that the rigid substrate is not strained, i.e., the lengths of vectors of the commensurate structure are equal to those of the superlattice of the substrate ($l \equiv |\bar{a}_1^{\rm s}|$).

We also screen the configurations with a large lattice mismatch. In the definition of mismatch $\Delta = (l - |\bar{a}_1^{\rm a}|)/|\bar{a}_1^{\rm a}| = (|\bar{a}_1^{\rm s}| - |\bar{a}_1^{\rm a}|)/|\bar{a}_1^{\rm a}|$, Δ is equal to the strain of the adsorbate [Fig. 2(a)]. That is, a negative value of Δ implies that the adsorbate is compressed, and vice versa. In general, materials will be unstable under large strains. We plot the $E_{\rm b}$ vs Δ of several configurations in Figs. 2(b)–2(c). The normalized binding energy is defined as $E_{\rm b} = (E^{\rm a} + E^{\rm s} - E^{\rm c})/S$, where S is the area of the superlattice; $E^{\rm a}$, $E^{\rm s}$ and $E^{\rm c}$ are energies of the adsorbate, the substrate and their composite structure, respectively. The configurations having larger binding energies are more stable according to the definition. There is an approximate negative correlation between $E_{\rm b}$ and the absolute value of mismatch $|\Delta|$. The $|\Delta|$ of the most stable 10 configurations (red points) are always less than 3% (gray area region). Therefore, we drop the likely unstable configurations if $|\Delta| < \Delta_{\rm t}$, where the $\Delta_{\rm t}$ is the limitation of mismatch. Interestingly, the most stable configuration could deviate from the smallest mismatch [shown in Figs. 2(b)–2(c)]. Thus, the smallest Δ is insufficient to guarantee the most stable configuration.

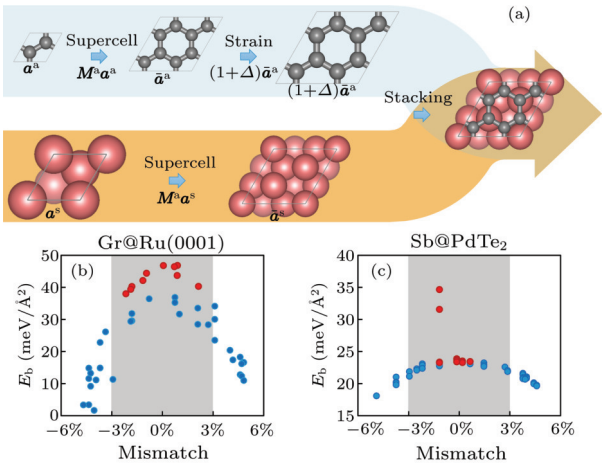

Fig. 2. The construction of 2D material and substrate interface. (a) An illustration of building supercells under Δ strains and stacking processes. The binding energy $E_{\rm b}$ vs lattice mismatch for the configurations of (b) Gr@Ru(0001) and (c) Sb@$PdTe_2$. The red spots represent the ten most stable configurations.

The stacking process is illustrated in Fig. 2(a). The superlattices of an adsorbate or a substrate are built by multiplying their primitive cell with a superlattice matrix. Furthermore, the superlattice of the adsorbate is multiplied by $(1 + \Delta)$ to match the magnitude of the substrate before stacking the superlattices of the adsorbate and the substrate to obtain a commensurate structure. Moreover, different adsorption sites are considered on the substrate [shown in Fig. S3(a)]. In fact, several adsorption sites are implicitly included in the large supercell [shown in Fig. S3(b)]. Thus, we only construct different structures to include all the possible adsorption sites for the small superlattices (typically for $l < 15$ Å).

Finally, we perform structural optimizations and total energy calculations for each atomic structure based on DFT calculations. The optimized structure and binding energy of configurations are added to the database. The configurations are sorted by the binding energy normalized to the area. For the interface systems that have several configurations with significantly larger binding energies than others, we consider that they are the stable configurations. Otherwise, the configurations are deemed to be stable if the normalized binding energy is larger than 80% of the most stable structure.

Table 1. Top 10 configurations with large binding energies (E_b) of Gr@Ru(0001). $\boldsymbol{M}^{Gr}$ and $\boldsymbol{M}^{Ru(0001)}$ are the superlattice matrices of the adsorbate and substrate, respectively. The experimentally reported configurations are labeled in the rows 1–3 in italic type. Here l is the lattice constant of supercell, Δ and θ are the mismatch and the twist angle, respectively, between the adsorbate and substrate.

$\boldsymbol{M}^{Gr}$	$\boldsymbol{M}^{Ru(0001)}$	Gr@Ru(0001)	l (Å)	Δ	θ (deg)	E_b (meV/Å^2)
(12,0)	*(11,0)*	*12@11*	*29.87*	*0.89%*	*0.00*	*47*
(11,0)	*(10,0)*	*11@10*	*27.15*	*0.06%*	*0.00*	*47*
(13,0)	*(12,0)*	*13@12*	*32.58*	*1.60%*	*0.00*	*46*
(6,1)	(6,0)	$\sqrt{43}$@6	16.29	0.71%	7.59	46
(10,0)	(9,0)	10@9	24.44	−0.94%	0.00	44
(5,0)	(1,4)	5@$\sqrt{21}$	12.44	0.88%	10.89	44
(5,1)	(5,0)	$\sqrt{31}$@5	13.58	−1.16%	8.95	42
(6,0)	(1,5)	6@$\sqrt{31}$	15.12	2.14%	8.95	40
(5,3)	(5,2)	7@$\sqrt{39}$	16.96	−1.81%	5.68	40
(5,2)	(5,1)	$\sqrt{39}$@$\sqrt{31}$	15.12	−1.87%	7.15	39
(9,0)	(8,0)	9@8	21.72	−2.16%	0.00	38

Table 2. Top 20 configurations with large binding energies of Gr@Ir(111). The experimentally reported configurations are labeled in the rows 1, 3, 12, 13, 15 and 16 in italic type.

$\boldsymbol{M}^{Gr}$	$\boldsymbol{M}^{Ir(111)}$	Gr@Ir(111)	l (Å)	Δ	θ (deg)	E_b (meV/Å^2)
(4,0)	*(1,3)*	*4@*$\sqrt{13}$	*9.87*	*0.05%*	*13.90*	*38*
(5,1)	(5,0)	$\sqrt{31}$@5	13.69	−0.33%	8.95	37
(10,0)	*(9,0)*	*10@9*	*24.64*	*−0.11%*	*0.00*	*37*
(5,3)	(5,2)	7@$\sqrt{39}$	17.10	−0.98%	5.68	36
(7,0)	(2,5)	7@$\sqrt{39}$	17.10	−0.98%	16.10	35
(5,2)	(5,1)	$\sqrt{39}$@$\sqrt{31}$	15.25	−1.05%	7.15	35
(3,5)	(5,2)	7@$\sqrt{39}$	17.10	−0.98%	22.11	35
(5,2)	(1,5)	$\sqrt{39}$@$\sqrt{31}$	15.25	−1.05%	25.05	35
(11,0)	(10,0)	11@10	27.38	0.90%	0.00	34
(9,0)	(8,0)	9@8	21.91	−1.34%	0.00	34
(6,1)	(6,0)	$\sqrt{43}$@6	16.43	1.56%	7.59	31
(3,4)	*(5,1)*	$\sqrt{37}$@$\sqrt{31}$	*15.25*	*1.59%*	*25.77*	*30*
(4,3)	*(5,1)*	$\sqrt{37}$@$\sqrt{31}$	*15.25*	*1.59%*	*16.34*	*30*
(5,0)	(1,4)	5@$\sqrt{21}$	12.55	1.72%	10.89	29
(3,2)	*(4,0)*	$\sqrt{19}$*@4*	*10.95*	*1.85%*	*23.41*	*29*
(3,0)	*(1,2)*	*3@*$\sqrt{7}$	*7.24*	*−2.12%*	*19.11*	*25*
(4,4)	(4,3)	$\sqrt{48}$@$\sqrt{37}$	16.66	−2.55%	4.72	23
(8,0)	(7,0)	8@7	19.17	−2.88%	0.00	19
(1,6)	(4,3)	$\sqrt{43}$@$\sqrt{37}$	16.66	2.96%	27.13	16
(1,6)	(3,4)	$\sqrt{43}$@$\sqrt{37}$	16.66	2.96%	17.70	16

We demonstrate our workflow in graphene on Ru(0001), which is an extensively investigated system.[42,44,53–57] The first ten configurations with largest binding energies are presented in Table 1. The first three of them, i.e., (12,0)@(11,0), (11,0)@(10,0) and (13,0)@(12,0), have been fabricated experimentally in the previous work.[44,53,54] Here, the indexes before and after @ correspond to the superlattice index of adsorbate and substrate, respectively. The configuration of a 25 × 25 graphene honeycomb sitting on a 23 × 23 ruthenium far exceeds the threshold of lattice constant l_t, thus has not been considered in our calculations. However, the configuration (25,0)@(23,0) could be regarded as the combination of (13,0)@(12,0) and (12,0)@(11,0). Therefore, the configuration of (13,0)@(12,0) with the lattice constant (32.6 Å) is slightly larger than l_t, and it is calculated as an exception. These agreements make it reasonable to use the binding energy as the criterion for predicting the experimentally possible configurations of 2D materials on substrates. Interestingly, we find some configurations [e.g., (6,1)@(6,0), (5,0)@(1,4)] with large

binding energies and non-zero twist angles that have not been reported so far. We believe that some of these configurations may have great potential to be observed if more experiments have been carried out. For some other systems, nucleation centers have preferred orientations, which may also affect the final configurations. Nevertheless, the predicted configurations with relatively strong binding energy may be obtained by a specific method, e.g., high-temperature annealing.

We also demonstrate the workflow in graphene on an Ir(111) surface, which is a medium interacting system. Due to the relative weak van der Waals (vdW) bonding, graphene shows multi-oriented superstructures on the Ir(111) surface.[45,46] The two most stable configurations predicted by high-throughput calculations in this work have been observed experimentally (listed in Table 2). The energy ranking R of the rest experimental results are in the range from 11 to 20 (listed in Table S2). It is worth noting that the (2,0)@(1,1) configuration possesses a negative binding energy, implying the energetically unfavorable, which is in contrast to the experimental results.[45,46] Moreover, there are also a series of large binding energy configurations that have the potential to be fabricated in the future.

It is well known that the choice of exchange-correlation functional is important for systems with medium and weak interaction, i.e., adsorption height and binding energy will be significantly affected. Nevertheless, stable configurations predicted by the high-throughput calculation method are independent of the choice of exchange-correlation function. The robustness of our method is checked by testing different exchange-correlation functionals. The binding energies and their ranking R of three different functionals (i.e., DFT-D3, LDA and optB88-vdW) have been listed in Table S2. The binding energies of the specific configurations obtained using different functionals are significantly different, but their ranking shows negligible dependence on the functionals used.

Table 3. Top 10 configurations with large binding energies of Sb@$PdTe_2$. The experimentally reported configurations are listed in rows of 1 and 2 in italic type and they have significantly higher binding energies than other configurations.

$\boldsymbol{M}^{\mathrm{Sb}}$	$\boldsymbol{M}^{\mathrm{PdTe_2}}$	Sb@$PdTe_2$	l (Å)	Δ	θ (deg)	E_{b} (meV/Å^2)
(1,0)	*(1,0)*	*1@1*	*4.07*	*−1.17%*	*0.00*	*35*
(0,1)	*(1,0)*	*1@1*	*4.07*	*−1.17%*	*60.00*	*32*
(−4,8)	(3,5)	$\sqrt{48}$@7	28.49	−0.15%	51.79	24
(4,4)	(7,0)	$\sqrt{48}$@7	28.49	−0.15%	30.00	24
(0,6)	(4,3)	6@$\sqrt{37}$	24.75	0.19%	34.72	24
(−3,6)	(2,4)	$\sqrt{27}$@$\sqrt{28}$	21.53	0.64%	49.11	23
(−4,4)	(5,3)	$\sqrt{48}$@7	28.49	−0.15%	8.21	23
(−3,7)	(3,4)	$\sqrt{37}$@$\sqrt{37}$	24.75	−1.17%	50.57	23
(0,6)	(3,4)	6@$\sqrt{37}$	24.75	0.19%	25.28	23
(−3,8)	(3,5)	7@7	28.49	−1.17%	43.57	23

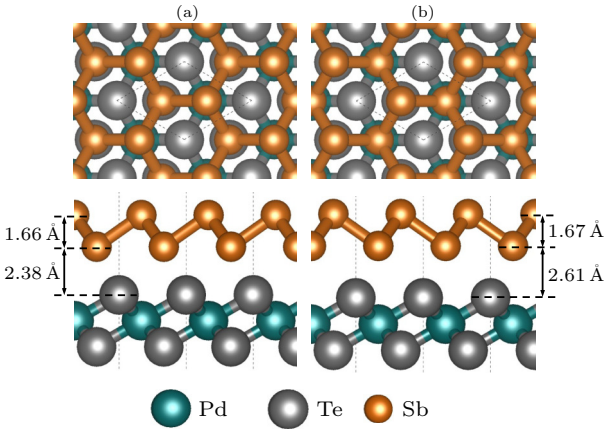

Fig. 3. The two most stable structures of Sb@$PdTe_2$ predicted by high-throughput calculations. The configuration of (a) $(0,1)$@$(1,0)$–R60° was reported in the previous work[48] and the configuration of (b) $(1,0)$@$(1,0)$–R0° is another possible configuration predicted by our high-throughput calculations.

Our workflow is also applicable to lower-symmetry systems. Unlike graphene, freestanding group-VA monolayers have a buckled honeycomb structure, which holds the three-fold rotational symmetry. The available twist angle θ is in the range of $[0°, 60°]$ (listed in Table S1). We calculate the possible configurations of buckled antimonene (b-Sb) on a $PdTe_2$ substrate and present the results in Table 3. The $PdTe_2$ is one of the typical layered materials. Therefore, we consider the (001) surface as the substrate, which is not specially labeled in the following. The two superstructures with the largest binding energy, which are (1,0)@(1,0) and (0,1)@(1,0) (shown in Fig. 3), have significant larger binding energies (35 and 32 meV/Å^2) than other configurations ($\leq$24 meV/Å^2). Therefore, both structures are grouped into the most stable configurations. It is noted that these two composite structures cannot be transformed into each other through symmetry operation or changing the adsorption site. That is, two different orientations of antimonene with the same period could form on the $PdTe_2$ surface.

However, only (0,1)@(1,0) was reported in the previous work.[48] As (1,0)@(1,0) and (0,1)@(1,0) yield the same scanning tunneling microscope (STM) images and low energy electron diffraction (LEED) pattern, we believe that the (1,0)@(1,0) may also exist but may not have been identified correctly in Ref. [44]. It is worthy to note that the 2D materials heterostructures could be predicted by considering 2D phase $PdTe_2$. In this case, all atoms of 2D $PdTe_2$ (contrastively, only one atom layer in bulk $PdTe_2$) are relaxed.

The same situation exists in the arsenene on the Ag(111) system. The 4×4 buckled arsenene on 5×5 Ag(111) was fabricated experimentally.[47] A series of linear boundaries are interpreted as the small shift of the domains. However, considering the three-fold rotational symmetry in both the buckled arsenene (b-As) and the Ag(111) surface, the two nonequivalent configurations of (4,0)@(5,0) and (0,4)@(5,0) (as shown in Table S3), with similar binding energies, can be expected to coexist. The experimentally observed linear bright features are more likely to be the grain boundary of these two configurations.

We also evaluate the possibility of flattened arsenene epitaxially grown on the substrate, which is an analogue to planar graphene. We compare the buckled arsenene (b-As) and the flat arsenene (f-As) on an Ag(111) substrate (shown in Tables S3 and S4). It is found that the binding energies of the buckled arsenene are much larger than those of the flat arsenene indicating that the b-As is energetically favorable to be synthesized on the Ag(111) substrate.

It should be noted that tens of configurations of b-As@Ag(111) show similar binding energies (shown in Table S3). Among them, only the (4,0)@(5,0) and (0,4)@(5,0) of b-As@Ag(111) were synthesized under specific experimental conditions. Other possible configurations (shown in Table S3) also have the potential to be fabricated on the Ag(111) substrate. On the other hand, however, several different configurations which are energetically favorable, signifying multiple phases of arsenene, could also be formed on the Ag(111) surface, and more precise experimental conditions are required to fabricate the specific large-scale monocrystalline structure.

In summary, we have built a computational database to predict the atomic structures of the 2D material-substrate interface. High-throughput DFT calculations are performed to obtain the energetically favorable configurations. The validity of this method is verified by several reported systems, e.g., Gr@Ru(0001), Gr@Ir(111), Sb@$PdTe_2$. Based on this method, we have predicted several possible configurations that could be potentially synthesized in experiments. The database provides a convenient and powerful tool for discovery of suitable substrates to grow 2D materials and novel 2D heterostructures.

References

[1] Sarmadian N, Saniz R, Partoens B, Lamoen D, Volety K, Huyberechts G and Paul J 2014 *Phys. Chem. Chem. Phys.* **16** 17724
[2] Zhang J, Zeng Q, Oganov A R, Dong D and Liu Y 2014 *Phys. Lett. A* **378** 3549
[3] Zhang S, Xie M, Li F, Yan Z, Li Y, Kan E, Liu W, Chen Z and Zeng H 2016 *Angew. Chem. Int. Ed.* **55** 1666
[4] Emery A A, Saal J E, Kirklin S, Hegde V I and Wolverton C 2016 *Chem. Mater.* **28** 5621
[5] Kweun J M, Li C, Zheng Y, Cho M, Kim Y Y and Cho K 2016 *Appl. Surf. Sci.* **370** 279
[6] Wang J, Yang X, Wang G, Ren J, Wang Z, Zhao X and Pan Y 2017 *Comput. Mater. Sci.* **134** 190
[7] Miyata M, Ozaki T, Takeuchi T, Nishino S, Inukai M and Koyano M 2018 *J. Electron. Mater.* **47** 3254
[8] Restuccia P, Levita G, Wolloch M, Losi G, Fatti G, Ferrario M and Righi M C 2018 *Comput. Mater. Sci.* **154** 517
[9] Barreteau C, Crivello J C, Joubert J M and Alleno E 2019 *Comput. Mater. Sci.* **156** 96
[10] Saal J E, Kirklin S, Aykol M, Meredig B and Wolverton C 2013 *JOM* **65** 1501
[11] Emery A A and Wolverton C 2017 *Sci. Data* **4** 170153
[12] Kirklin S, Saal J E, Hegde V I and Wolverton C 2016 *Acta Mater.* **102** 125
[13] Schatschneider B, Monaco S, Liang J J and Tkatchenko A 2014 *J. Phys. Chem. C* **118** 19964
[14] Bradlyn B, Elcoro L, Cano J, Vergniory M G, Wang Z J, Felser C, Aroyo M I and Bernevig B A 2017 *Nature* **547** 298
[15] Vergniory M G, Elcoro L, Felser C, Regnault N, Bernevig B A and Wang Z J 2019 *Nature* **566** 480
[16] Curtarolo S, Setyawan W, Hart G L W, Jahnatek M, Chepulskii R V, Taylor R H, Wanga S D, Xue J K, Yang K S, Levy O, Mehl M J, Stokes H T, Demchenko D O and Morgan D 2012 *Comput. Mater. Sci.* **58** 218
[17] Calderon C E, Plata J J, Toher C, Oses C, Levy O, Fornari M, Natan A, Mehl M J, Hart G, Nardelli M B and Curtarolo S 2015 *Comput. Mater. Sci.* **108** 233
[18] Jain A, Hautier G, Moore C J, Ong S P, Fischer C C, Mueller T, Persson K A and Ceder G 2011 *Comput. Mater. Sci.* **50** 2295
[19] Jain A, Ong S P, Hautier G, Chen W, Richards W D, Dacek S, Cholia S, Gunter D, Skinner D, Ceder G and Persson K A 2013 *APL Mater.* **1** 011002
[20] Ashton M, Paul J, Sinnott S B and Hennig R G 2017 *Phys. Rev. Lett.* **118** 106101
[21] Lebegue S, Bjorkman T, Klintenberg M, Nieminen R M and Eriksson O 2013 *Phys. Rev. X* **3** 031002
[22] Miro P, Audiffred M and Heine T 2014 *Chem. Soc. Rev.* **43** 6537
[23] Tada T, Takemoto S, Matsuishi S and Hosono H 2014 *Inorg. Chem.* **53** 10347
[24] Mounet N, Gibertini M, Schwaller P, Campi D, Merkys A, Marrazzo A, Sohier T, Castelli I E, Cepellotti A, Pizzi G and Marzari N 2018 *Nat. Nanotechnol.* **13** 246
[25] Olsen T, Okugawa T, Torelli D, Deilmann T and Thygesen K S 2019 *Phys. Rev. Mater.* **3** 024005
[26] Choudhary K, Kalish I, Beams R and Tavazza F 2017 *Sci. Rep.* **7** 5179
[27] Haastrup S, Strange M, Pandey M, Deilmann T, Schmidt P S, Hinsche N F, Gjerding M N, Torelli D, Larsen P M, Riis-Jensen A C, Gath J, Jacobsen K W, Mortensen J J, Olsen T and Thygesen K S 2018 *2D Mater.* **5** 042002
[28] Bae S H, Kum H, Kong W, Kim Y, Choi C, Lee B, Lin P,

Park Y and Kim J 2019 *Nat. Mater.* **18** 550
[29] Zhang S, Yan Z, Li Y, Chen Z and Zeng H 2015 *Angew. Chem. Int. Ed.* **54** 3112
[30] Scalise E, Houssa M, Pourtois G, Afanas'ev V V and Stesmans A 2012 *Nano Res.* **5** 43
[31] Peng X, Wei Q and Copple A 2014 *Phys. Rev. B* **90** 085402
[32] Amorim B, Cortijo A, de Juan F, Grushin A G, Guinea F, Gutiérrez-Rubio A, Ochoa H, Parente V, Roldán R, San-Jose P, Schiefele J, Sturla M and Vozmediano M A H 2016 *Phys. Rep.* **617** 1
[33] Gao Y, Zhang Y Y and Du S 2019 *J. Phys.: Condens. Matter* **31** 194001
[34] Giovannetti G, Khomyakov P A, Brocks G, Karpan V M, van den Brink J and Kelly P J 2008 *Phys. Rev. Lett.* **101** 026803
[35] Yan J, Wu L, Ma R S, Zhu S, Bian C, Ma J, Huan Q, Bao L, Mao J, Du S and Gao H J 2019 *2D Mater.* **6** 045050
[36] Dedkov Y S, Fonin M, Rudiger U and Laubschat C 2008 *Phys. Rev. Lett.* **100** 107602
[37] Robinson B J, Giusca C E, Gonzalez Y T, Kay N D, Kazakova O and Kolosov O V 2015 *2D Mater.* **2** 015005
[38] Su L Q, Yu Y F, Cao L Y and Zhang Y 2015 *Nano Res.* **8** 2686
[39] Mathew K, Singh A K, Gabriel J J, Choudhary K, Sinnott S B, Davydov A V, Tavazza F and Hennig R G 2016 *Comput. Mater. Sci.* **122** 183
[40] Ding H, Dwaraknath S S, Garten L, Ndione P, Ginley D and Persson K A 2016 *ACS Appl. Mater. Interfaces* **8** 13086
[41] Gao B, Gao P Y, Lu S H, Lv J, Wang Y C and Ma Y M 2019 *Sci. Bull.* **64** 301
[42] Pan Y, Shi D X and Gao H J 2007 *Chin. Phys.* **16** 3151
[43] Martoccia D, Willmott P R, Brugger T, Björck M, Günther S, Schlepütz C M, Cervellino A, Pauli S A, Patterson B D, Marchini S, Wintterlin J, Moritz W and Greber T 2008 *Phys. Rev. Lett.* **101** 126102
[44] Iannuzzi M, Kalichava I, Ma H, Leake S J, Zhou H, Li G, Zhang Y, Bunk O, Gao H, Hutter J, Willmott P R and Greber T 2013 *Phys. Rev. B* **88** 125433
[45] Meng L, Wu R T, Zhang L Z, Li L F, Du S X, Wang Y L and Gao H J 2012 *J. Phys.: Condens. Matter* **24** 314214
[46] Loginova E, Nie S, Thürmer K, Bartelt N C and McCarty K F 2009 *Phys. Rev. B* **80** 085430
[47] Shah J, Wang W, Sohail H M and Uhrberg R I G 2020 *2D Mater.* **7** 025013
[48] Wu X, Shao Y, Liu H, Feng Z, Wang Y L, Sun J T, Liu C, Wang J O, Liu Z L, Zhu S Y, Wang Y Q, Du S X, Shi Y G, Ibrahim K and Gao H J 2017 *Adv. Mater.* **29** 1605407
[49] 2D material and substrate interfaces database [Online] (accessed 2, 2021). http://n11.iphy.ac.cn/2dinterface.html
[50] Zhao L, Rim K T, Zhou H, He R, Heinz T F, Pinczuk A, Flynn G W and Pasupathy A N 2011 *Solid State Commun.* **151** 509
[51] He R, Zhao L, Petrone N, Kim K S, Roth M, Hone J, Kim P, Pasupathy A and Pinczuk A 2012 *Nano Lett.* **12** 2408
[52] Xu X, Zhang Z, Dong J, Yi D, Niu J, Wu M, Lin L, Yin R, Li M, Zhou J, Wang S, Sun J, Duan X, Gao P, Jiang Y, Wu X, Peng H, Ruoff R S, Liu Z, Yu D, Wang E, Ding F and Liu K 2017 *Sci. Bull.* **62** 1074
[53] Marchini S, Günther S and Wintterlin J 2007 *Phys. Rev. B* **76** 075429
[54] de Parga A L V, Calleja F, Borca B, Passeggi M C G, Hinarejos J J, Guinea F and Miranda R 2008 *Phys. Rev. Lett.* **100** 056807
[55] Wang B, Günther S, Wintterlin J and Bocquet M L 2010 *New J. Phys.* **12** 043041
[56] Que Y D, Xiao W D, Fei X M, Chen H, Huang L, Du S X and Gao H J 2014 *Appl. Phys. Lett.* **104** 093110
[57] Silva C C, Iannuzzi M, Duncan D A, Ryan P T P, Clarke K T, Kuchle J T, Cai J Q, Jolie W, Schlueter C, Lee T L and Busse C 2018 *J. Phys. Chem. C* **122** 18554

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.

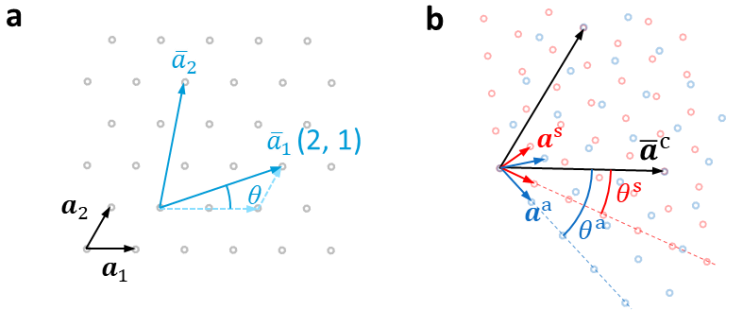

Figure S1. Definition of index notation and twist angle. **(a)** Schematic of (*m*, *n*) notation of supercell. The gray circles represent the lattice points of unitcell. **(b)** Illustration of adsorbate on substrate with a configuration of (1, 3)@(3, 2). The blue and red circles represent the unitcell 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	$[0°, 30°]$

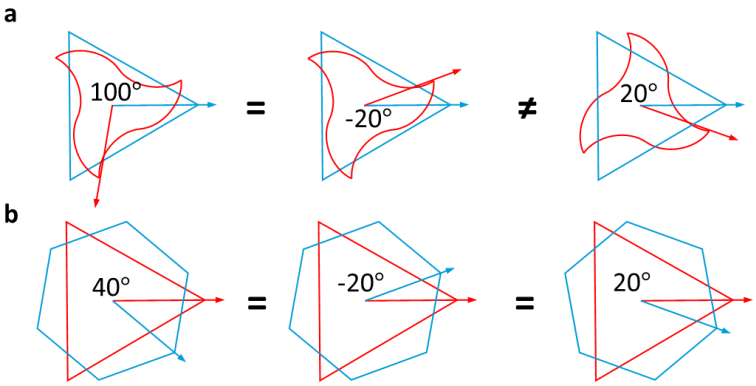

Figure S2. Schematics for the reduction of the twist angle out of range. **(a)** The effective twist angle range is [0 °, 120 °) due to the C_3 and C_{3v} symmetries of two materials, respectively. **(b)** The two materials possess C_{3v} and C_{6v} symmetries, respectively. Therefore, the effective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

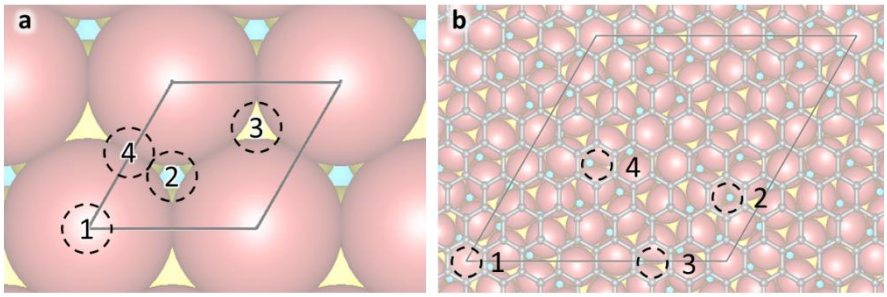

Figure S3. Various adsorption sites are considered in the workflow. **(a)** Four adsorption sites on the hexagonal substrate. **(b)** The adsorption sites mentioned in (a) are implicitly included in the large supercell (typically ≥15 Å).

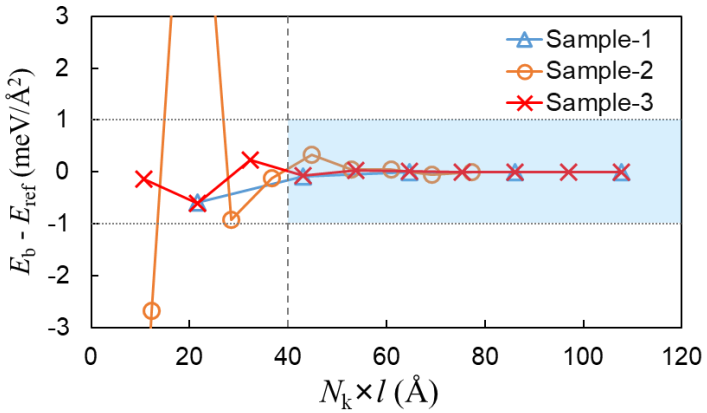

Figure S4. Typical K-sample test results of Sb@$PdTe_2$. The K points are taken dense enough (gray area) to ensure that binding energies of various structures converge to 1 meV/Å^2.

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.

$\boldsymbol{M}^{Gr}$	$\boldsymbol{M}^{Ir(111)}$	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