Grand-canonical evolutionary algorithm for the prediction of two-dimensional materials

Benjamin C. Revard, William W. Tipton, Anna Yesypenko, and Richard G. Hennig

Phys. Rev. B **93**, 054117 — Published 26 February 2016

DOI: 10.1103/PhysRevB.93.054117

Grand Canonical Evolutionary Algorithm for the Prediction of Two-Dimensional Materials

Benjamin C. Revard,[1,2] William W. Tipton,[1] Anna Yesypenko,[1] and Richard G. Hennig[1,2]

[1] *Department of Materials Science and Engineering, Cornell University, Ithaca, New York 14853, USA*
[2] *Department of Materials Science and Engineering, University of Florida, Gainesville, FL 32611, USA*

(Dated: February 4, 2016)

Single-layer materials represent a new materials class with properties that are potentially transformative for applications in nanoelectronics and solar energy harvesting. With the goal of discovering novel two-dimensional (2D) materials with unusual compositions and structures, we have developed a grand canonical evolutionary algorithm that searches the structure and composition space while constraining the thickness of the structures. Coupling the algorithm to first principles total energy methods, we show that this approach can successfully identify known 2D materials and find novel low-energy ones. We present the details of the algorithm, including suitable objective functions, and illustrate its potential with a study of the Sn-S and C-Si binary materials systems. The algorithm identifies several new 2D structures of InP, recovers known 2D structures in the binary Sn-S and C-Si systems, and finds two new 1D Si defects in graphene with formation energies below that of isolated substitutional Si atoms.

I. INTRODUCTION

Two-dimensional (2D) materials posses properties that are distinct from those of their bulk counterparts. For example, single-layer MoS_2 exhibits a wider and direct band gap than bulk MoS_2,[1,2] and single-layer SnS_2 has interesting photocatalytic properties lacked by the bulk material.[3,4] Such unique physical properties motivate intense research interest in the field of 2D materials in the last decade. Proposed applications for 2D materials include optical sensors,[1] nanoelectronics,[5] and photocatalysts for water splitting.[3,4,6]

Several new single-layer materials have been computationally predicted in recent years.[7–10] In one common approach to 2D structure prediction, which relies on chemical intuition and substitution, the lattice sites of known 2D crystal structures are decorated with different atomic species.[9] Another method involves mining databases of bulk crystal structures to identify those with layered motifs, from which a single layer could potentially be exfoliated.[10] These approaches are quite useful, but since they are based on previously known structures, they make assumptions about the structures and compositions of 2D materials. These assumptions may unnecessarily constrain the search, leaving other potentially synthesizable single-layer materials awaiting discovery.

Recently, much practical success in the prediction of the structures of clusters and three-dimensional crystals has been achieved with global optimization techniques, such as evolutionary algorithms[11–17] and particle swarm optimization.[18] Some of these methods have been extended to search for 2D materials. Bahmann and Kortus developed an evolutionary algorithm that can search for 2D crystals,[19] and Zhou *et al.* extended an evolutionary algorithm to search for 2D structures.[20] Both search for structures with fixed stoichiometry and number of atoms per cell. Luo *et al.* extended a particle swarm optimization algorithm to search for fixed stoichiometry 2D structures that are both completely planar[21] and have finite layer thicknesses.[22] In this work, we extend our grand canonical evolutionary algorithm[13] to enable unbiased searching for the *compositions* and *structures* of novel 2D materials with low formation energies and finite layer thicknesses.

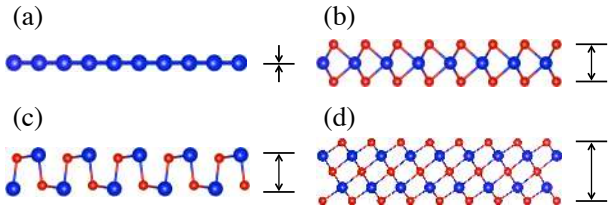

FIG. 1. Layer thicknesses of several 2D structures: (a) graphene with a thickness of zero, (b) MoS_2, (c) PbO, and (d) the slightly thicker structure of Bi_2Te_3. The structures are shown from the side, and the layer thickness is the maximum vertical distance between atoms in the cell.

The details of the grand canonical evolutionary algorithm for 2D materials are described in Sec. II. In Sec. III we demonstrate that the algorithm enables the prediction of novel 2D materials structures by applying it to a system at fixed composition (2D-InP) in Sec. III A and then to the composition space of the binary Sn-S and C-Si materials systems in Sec. III B and III C. The algorithm successfully identifies the known low-energy 2D structures in these materials systems. In addition, it discovers several new 2D structures of InP and a novel 2D structure of C_6Si with formation energy below that of previously known 2D structures in this system.

II. METHODS

A. Search Space and Objective Function

The search for novel 2D materials requires the definition of a search space and an objective function. We define a 2D material to be a crystal with a structure that is periodic in two dimensions and has a finite extent in the third dimension. The layer thickness of a 2D material is the maximum vertical distance between atoms in the structure, as illustrated in Fig. 1. Two-dimensional materials display a range of layer thicknesses, from completely planar structures like graphene and hexagonal boron nitride to thicker structures with multiple sublayers, such as metal dichalcogenides (three sublayers), group-IV monochalcogenides (four sublayers), and Bi_2Te_3 (five sublayers).[1,2,23,24] To include 2D structures with different layer thicknesses in our search, we employ a finite-thickness constraint. Hence, the search space for 2D materials consists of the configuration space of all materials that are periodic in two dimensions and exhibit a finite thickness in the third dimension.

For the objective function, we propose the formation energy of a 2D material relative to the bulk ground state phase, as defined by

$$\Delta E_{\mathrm{f}} = \frac{E_{\mathrm{2D}}}{N_{\mathrm{2D}}} - \frac{E_{\mathrm{3D}}}{N_{\mathrm{3D}}}, \qquad (1)$$

where E_{2D} and E_{3D} are the total energies of the 2D and 3D structures, respectively, and N_{2D} and N_{3D} are the numbers of atoms in the 2D and 3D unit cells.[25] Common to all synthesized free-standing 2D materials is a low formation energy with respect to the bulk ground state phase. Three well-known 2D materials illustrate this point: the formation energy of graphene is only about 56 meV/atom relative to graphite,[26] that of single-layer MoS_2 is 77 meV/atom,[27] and we calculated the formation energy of phosphorene to be 112 meV/atom relative to bulk phosphorous.

Fig. 2 summarizes the formation energies of many 2D materials, both predicted and synthesized. We observe that all 2D materials that have been synthesized as free-standing films posses formation energies less than 200 meV/atom.[6] Therefore, we take this as an empirical rule of thumb to determine whether a 2D material has a realistic chance of being experimentally synthesized without the need for stabilization by a substrate.[28,29] In Sec. III A, we discuss an alternative choice of objective function for 2D structure prediction.

B. 2D Evolutionary Algorithm

To search for low-energy 2D crystal structures, we modify our grand canonical genetic algorithm for structure and phase prediction (GASP)[13] code by constraining the layer thickness of the crystal structures considered by the algorithm. In the following, we provide a brief description of the general evolutionary algorithm, and then focus in detail on the changes made for the 2D search. For a complete description of GASP and some of its applications, see Refs. 13, 17, 36, and 37. The GASP code is freely available under the GPL v3 license at http://gasp.mse.ufl.edu.

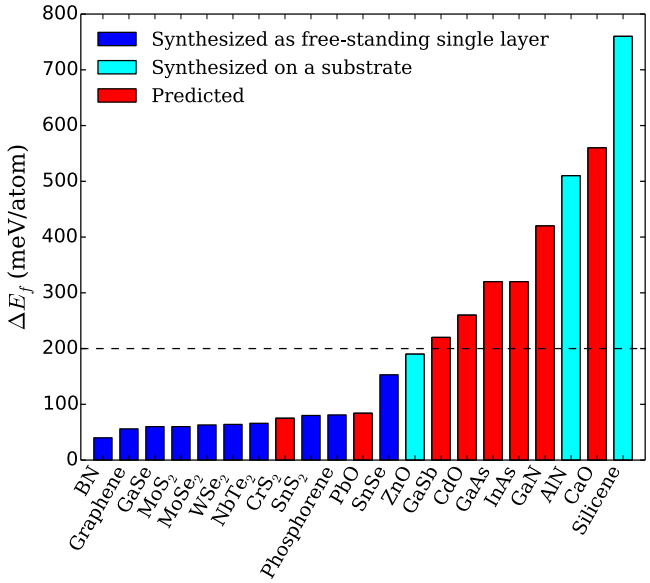

FIG. 2. Calculated formation energies relative to the bulk of several 2D materials, both predicted and synthesized. All 2D materials that have been synthesized as free-standing films have formation energies below 200 meV/atom, illustrated by the horizontal dashed line. The formation energies were calculated as follows: graphene[26] and phosphorene[30] with quantum Monte Carlo; BN,[31] MoS_2,[31] $MoSe_2$,[31] WSe_2,[31] $NbTe_2$[31] and PbO[31] with the RPA method; GaSe,[32] CrS_2,[33] SnS_2[4] and SnSe[23] with the vdW-DF-optB88 van der Waals functional; CdO,[34] CaO,[34] ZnO,[34] GaAs,[35] GaSb,[35] InAs,[35] GaN[35] and AlN[35] with the PBE functional, and silicene[25] with the LDA exchange correlation functional.

Overview. The evolutionary algorithm starts by generating a population of random structures that represent a broad sampling of the solution space. Structures are evaluated based on their relative formation energies per atom, with lower energy solutions being more favorable.[14] An offspring generation is then populated by probabilistically selecting lower energy structures to "reproduce" through biologically inspired operators such as mutation and mating. The mutation operator randomly perturbs the atomic positions and lattice vectors of the parent to create an offspring structure, and the mating operator combines spatially coherent pieces of two parent structures,[12,15,16,38–41] as illustrated in Fig. 3. When enough child structures have been created, they are evaluated for their fitness and then make an offspring generation of their own using the same evolutionary operators. This process continues until some user-defined stopping criteria are met. Since better solutions are selected more frequently to reproduce, structural traits associated with

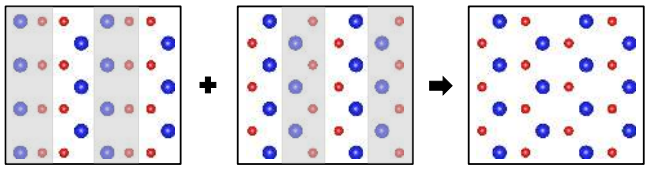

FIG. 3. Illustration of the mating operator. Sections are sliced from each parent structure, shown on the left, and combined to form an offspring structure. Supercells are shown for clarity.

low-energy crystals are propagated in subsequent generations.

Constraints. It is generally advantageous to apply some loose constraints to the parameters of structures considered by the algorithm. Constraints can guide the search for structures with particular characteristics of interest, and they can apply prior knowledge about which structures are likely to have good objective function values. For example, we know that crystal structures with atoms quite near to each other (less than about 75% of the equilibrium bond length) will have very high formation energies, and this knowledge is put to use by enforcing a constraint on the minimum interatomic distance. Constraints on the maximum and minimum lattice vector lengths can help prevent structures with unphysically short lattice vectors or unusually large or small aspect ratios from entering the population.[42] The composition, number of atoms in the unit cell, lattice vector angles and symmetries may also be constrained.

Structure representation. An evolutionary structure search will be more successful if structures are represented in the computer in a standardized fashion. This is because the mating operator is more likely to produce viable offspring if the parent structures are represented similarly.[13] In addition, it is easier to identify duplicate structures if each structure has a unique representation. The application of constraints (discussed above) to the lattice vector lengths and angles is one way to help standardize how structures are represented. Another useful method is the Niggli cell reduction algorithm,[43] which essentially transforms the representation of a structure such that its unit cell is closest to a cube.

Modifications for 2D structure search. To facilitate searching for 2D structures, we make three modifications to our evolutionary algorithm: (i) We impose a constraint on the layer thickness of structures, (ii) we modify the structure representation to be suitable for 2D materials, and (iii) we add a vacuum layer to the structures for the energy evaluation. Fig. 4 illustrates these modifications made to GASP to search for 2D structures.

First, a new constraint is imposed on the layer thickness of structures, which limits the search to the 2D regime. This constraint is an input parameter, so the user may change it to search for thinner or thicker structures as considered appropriate for the given material system. The algorithm checks all structures it generates against this new constraint, both before and after structural relaxation.

The second modification involves how 2D structures are represented. Once a new structure is created, the algorithm rotates the cell such that lattice vector $\vec{a}$ is parallel to the Cartesian x-axis and lattice vector $\vec{b}$ lies in the (x, y) plane. Since 2D materials are not periodic in the third dimension and have a finite thickness, no lattice vector along the third dimension is required. However, for compatibility with energy codes that are designed for 3D crystal structures, it is convenient to select a $\vec{c}$ lattice vector that is normal to the $(\vec{a}, \vec{b})$ plane. This is achieved by replacing the $\vec{c}$ lattice vector with its component along the Cartesian z axis. The $\vec{c}$ lattice vector is fixed during structural relaxation to reduce the computational cost of the relaxation, avoid collapse of the vacuum region, and to prevent a sheet from sliding relative to its periodic images during relaxation, which could lead to spurious minima.

Our choice of the $\vec{c}$ lattice vector also improves the success of the mating operator, which combines slices of unit cells that are taken parallel to one of the cell facets.[13,17] When the $\vec{c}$ lattice vector is normal to the 2D material structure, the slice plane is always either parallel or perpendicular to the plane of the 2D sheet in both parent structures, increasing the chances that the result will also correspond to a valid 2D structure.

When applying the Niggli cell reduction[43] to 2D structures, the reduced cell must correspond to a 2D structure. This means that the algorithm should only transform the $\vec{a}$ and $\vec{b}$ lattice vectors and, furthermore, $\vec{a}$ and $\vec{b}$ should remain in the x-y plane. To achieve this in practice, we simply increase the magnitude of the $\vec{c}$ lattice vector to an arbitrary large value before passing the cell to the Niggli reduction algorithm. After reduction, the $\vec{c}$ lattice vector is returned to its original magnitude.

The final modification is related to the evaluation of the energy of the 2D structures using codes that are designed for 3D bulk structures and employ periodic boundary conditions. Before passing a newly created 2D structure to an external code for relaxation and energy calculation, the algorithm sets the $\vec{c}$ lattice parameter of the unit cell such that the spacing between periodic images of the 2D material are sufficiently separated by vacuum to prevents spurious interactions between the periodic images.

C. Density Functional Calculations

To accurately relax the 2D candidate structures and determine their energy, we perform density functional theory (DFT) calculations with the Vienna *ab initio* simulation package (VASP).[44,45] The interactions between valence electrons and ionic cores are described by the projector augmented-wave (PAW) method.[46,47] The core electron states described by the PAW potentials are $1s^2$ for C, $1s^2 2s^2 2p^6$ for Si, P, and S, [Kr]$4d^{10}$ for Sn, and

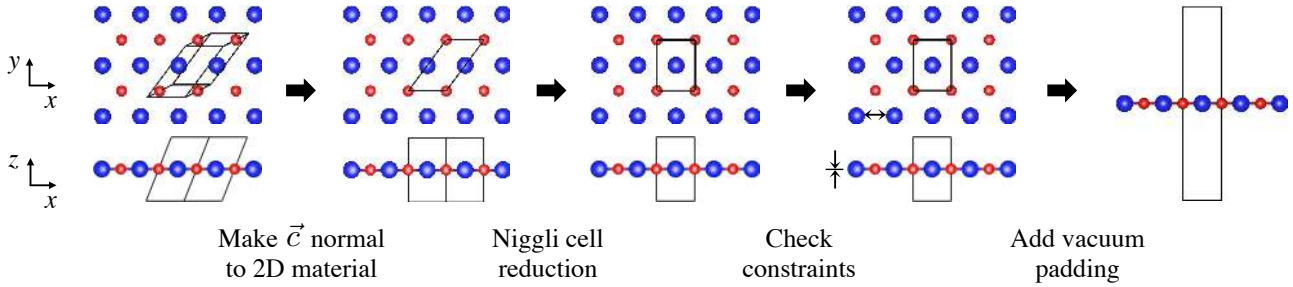

FIG. 4. Modifications made to GASP to prepare a candidate 2D structure for an energy calculation. First the structure is rotated to lie within the x-y plane and a $\vec{c}$ lattice vector normal to the plane is chosen. The algorithm then applies Niggli cell reduction to obtain the most cubic representation of the structure. The reduced structure is checked against the constraints, including the layer thickness constraint. If it passes, vacuum padding is added to the cell before it is relaxed and its energy evaluated by an external code. Afterwards, the vacuum padding is removed from the relaxed cell, Niggli cell reduction is applied, and the constraints are checked again before it is added to the offspring generation.

TABLE I. DFT parameters to converge the energy to within 1 meV/atom for the three materials systems under consideration.

System	Cutoff energy (eV)	k-points density (Å)	Vacuum padding (Å)
InP	300	35	16
Sn-S	550	55	12
C-Si	500	50	12

[Kr] for In. Tab. I summarizes the values of the DFT parameters used for the three considered materials systems. For each system, the energy was converged to within 1 meV/atom for each parameter relative to the energy obtained with a maximum reference value of the parameter. The maximum reference values for the cutoff energy, plane-wave basis set and vertical vacuum padding were 700 eV, a k-point density of 70 per Å^{-1}, and 30 Å, respectively.

All structural relaxations during the evolutionary algorithm search used the Perdew-Burke-Ernzerhof (PBE) approximation for the exchange-correlation functional.[48] Following the structure search, we perform structural relaxations and calculate the energies of all 2D structures found by the algorithm, as well as of the known ground-state bulk structures, using the computationally more demanding non-local vdW-DF-optB88 exchange-correlation functional.[49–51] This accurately accounts for the dispersion interactions that are important for the layered bulk structures of some of the materials. The $\vec{c}$ lattice vector is kept fixed during all structural relaxations.

For all structure searches, the number of atoms in the simulation cell was allowed to vary. Equilibrium bond lengths for the minimum interatomic distance constraints, as discussed in Sec. II B, are obtained from the relaxed ground state bulk structures in each system.

For the InP 2D structure searches, we fixed the stoichiometry of all structures encountered in the search to In:P = 1:1 and employed an upper bound of 12 atoms in the cell to limit the size of the search space and the computational cost of the DFT calculations. To explore the energy landscape as a function of layer thickness, we carried out a total of five evolutionary algorithm structure searches for InP, each with a successively larger layer thickness constraint, ranging from 2 to 6 Å. In each search, the algorithm was stopped after 500 successful structural relaxations and subsequent energy calculations.

For the Sn-S and C-Si structure searches, we allowed the stoichiometry to vary between the pure elements, and 1000 relaxations and energy calculations were performed in the structure search for each of these materials systems. For the Sn-S search, we chose a layer thickness constraint of 4 Å, which is slightly larger than the layer thicknesses of the known 2D structures in this system. Structures were permitted with up to 15 atoms in the cell.

For the C-Si search, we used a fairly conservative layer thickness constraint of 2 Å because the previously predicted 2D structures in this system have all been nearly or completely planar. The known elemental 2D structures of planar hexagonal graphene and buckled hexagonal silicene, with layer thicknesses of 0.00 Å and 0.45 Å, respectively, were provided to the algorithm in the initial generation. Up to 12 atoms per cell were allowed in the C-Si search.

III. RESULTS AND DISCUSSION

A. Indium Phosphide

Bulk indium phosphide is a direct gap semiconductor and occurs in the zincblende crystal structure.[53] A novel tetragonal structure of 2D InP was recently proposed by Zhuang *et al.*, and its formation energy was calculated to be comparable to that of the 2D buckled hexagonal structure previously proposed by Şahin *et al.*[35,54] Tong *et al.*[55] predicted a new 2D structure of InP to be lower in energy than both the buckled hexagonal and tetrag-

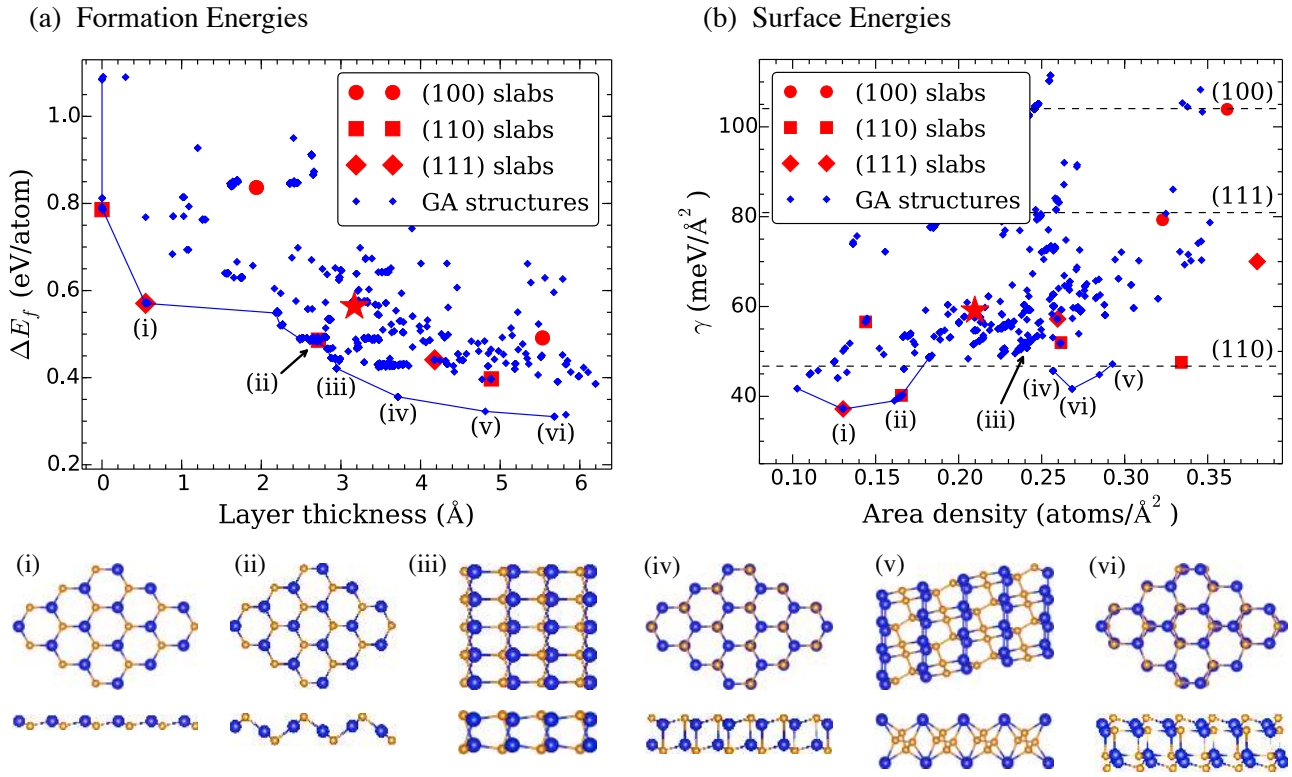

FIG. 5. Structure search for 2D InP. In both plots, the small blue diamonds indicate 2D structures found by GASP, and the larger red symbols correspond to 2D structures that were obtained by taking slabs from high symmetry planes of the bulk InP structure. The large red star near the center of each plot represents the previously proposed tetragonal structure.[35] In (a), the formation energies of the 2D structures with respect to the bulk phase is plotted versus their layer thicknesses. The blue line connects the lowest energy structures found by the algorithm at various thicknesses. Several structures that lie on this line are labeled (i) through (v) in both plots. In (b), the surface energies of the 2D structures are plotted versus their area densities. The three horizontal dashed lines designate the calculated surface energies of three facets of the bulk material. Top and side views of the low-energy 2D structures (i) to (vi) found by the algorithm are shown below the plots.

onal structure, and Susuki[56] recently reported a novel 2D structure of InP with the same layer thickness as the structure reported by Tong *et al.* and nearly degenerate formation energy. To determine if other low-energy 2D InP structures exist, we employ the evolutionary algorithm to do several 2D structure searches with varying layer thickness constraints. Fig. 5 displays the results of these 2D structure searches.

Since we are searching for 2D materials, we can treat the layer thickness as a second objective function that we seek to minimize (that is, in addition to the formation energy). The set of solutions that optimize the trade-off between layer thickness and formation energy form the Pareto front for this system. These structures are connected by a blue line in Fig. 5(a), which shows the formation energy versus layer thickness.

Several low formation energy structures that lie on the Pareto front are shown in Fig. 5 below the plots, labeled (i) through (vi). The buckled hexagonal structure shown in (i) was previously proposed by Şahin *et al.*[54] and was also recovered by relaxing a monolayer from the (111) plane of the bulk crystal. The orthorhombic structure shown in (ii) has been reported by Tong *et al.*[55] and is also obtained by relaxing a bilayer from the (110) plane of the bulk crystal. The structure labeled (iii) has orthorhombic symmetry and 8 atoms in the cell, and the hexagonal bilayer structure in (iv) is equivalent to two stacked buckled hexagonal structures, with one inverted. The triclinic structure shown in (v) has 8 atoms per cell, and (vi) is a more complex bilayer structure with monoclinic symmetry and 12 atoms in the cell. Table II contains the parameters for these six structures. The previously proposed tetragonal structure[35] is also shown, and it is located above the Pareto front.

We compare the results of the structure searches with an alternative approach to generating candidate 2D materials, in which thin slabs from the bulk structure are relaxed. Slabs of the bulk zincblende structure from the (100), (110) and (111) planes were padded with vacuum and relaxed, and the resulting 2D structures are shown as red symbols in Fig. 5. The low energy of several of the relaxed slabs of the bulk material shows that this simple approach can efficiently provide useful candidates for the prediction of novel 2D structures – it recovered two low formation energy structures lying on the Pareto front of this system. However, this technique misses the lowest

TABLE II. Structural parameters and formation energies of six low-energy 2D InP materials and the lowest energy C-Si material found by the evolutionary algorithm. We have used 3D space groups to describe these finite-thickness 2D structures that lack periodicity in the direction normal to the 2D sheet. In the representations given here, the $\vec{c}$ lattice vector is normal to the plane of the 2D sheet. Symmetry information was obtained with the FINDSYM software package.[52]

	Space group	a (Å)	b (Å)	c (Å)	Atomic positions		ΔE_f (meV/atom)
InP (i)	$P3m1$ (156)	4.209	4.209	3.146	In 1(b)	$z = -0.413$	571
					P 1(c)	$z = 0.413$	
InP (ii)	$Pmn2_1$ (31)	4.095	6.057	5.210	In 2(a)	$y = 0.290, z = -0.447$	484
					P 2(a)	$y = -0.388, z = 0.250$	
InP (iii)	$Abm2$ (39)	4.902	6.770	5.547	In 4(c)	$y = 0.214, z = 0.299$	421
					P 4(c)	$y = 0.340, z = -0.234$	
InP (iv)	$P\bar{3}m1$ (164)	4.240	4.240	6.317	In 2(d)	$z = 0.322$	356
					P 2(d)	$z = -0.206$	
InP (v)	P-1 (2)	5.039	5.979	7.414	In 2(i)	$x = 0.459, y = -0.316, z = -0.175$	322
					In 2(i)	$x = 0.461, y = 0.179, z = -0.186$	
					P 2(i)	$x = -0.047, y = 0.424, z = -0.374$	
					P 2(i)	$x = -0.209, y = 0.165, z = -0.488$	
InP (vi)	$C2/m$ (12)	7.115	12.572	8.278	In 8(j)	$x = 0.322, y = 0.351, z = -0.343$	310
					In 4(i)	$x = 0.375, z = -0.244$	
					P 8(j)	$x = 0.311, y = 0.341, z = 0.335$	
					P 4(i)	$x = 0.289, z = 0.157$	
C_6Si (ii)	$Cmm2$ (35)	16.600	2.456	2.842	C 4(d)	$x = 0.400, z = 0.404$	414
					C 4(d)	$x = -0.229, z = 0.282$	
					C 4(d)	$x = -0.142, z = 0.346$	
					Si 2(b)	$z = -0.282$	

energy structures available at higher thicknesses, labeled (iii) - (vi) in Fig. 5(a).

As can be seen in Fig. 5(a), the lowest formation energy found by the algorithm continuously decreases as the layer thickness increases. In other words, thicker slabs tend to be thermodynamically favored, and the bulk structure is the thermodynamic ground state. We expect most materials systems to follow this general pattern. For this reason, formation energy relative to the bulk is not an ideal objective function when searching for physically-realizable 2D materials, and we must enforce a layer thickness constraint (see Sec. II B) to prevent the algorithm from finding ever-thicker structures.

A possible solution to this issue is to use the surface energy of 2D structures as the objective function. We define the surface energy of a 2D material relative to the energy of the lowest-energy bulk material with the same composition,

$$\gamma = \frac{N_{2D}}{2A}\Delta E_{\mathrm{f}}, \tag{2}$$

where N_{2D} is the number of atoms in the 2D cell, $A = |\vec{a} \times \vec{b}|$ is area of the $(\vec{a}, \vec{b})$ facet of the 2D cell and ΔE_f is the formation energy relative to the bulk of the 2D material. Our definition of surface energy is motivated by considering slabs cut from a bulk crystal. For thick slabs, γ corresponds to the surface energy of the bulk material. For thinner slabs (with layer thickness less than about 10 Å), surface reconstructions are no longer constrained to match the underlying bulk structure, and additional reconstructions can occur through the interactions between the top and bottom surfaces. The 2D structure with the most optimal objective function value, as defined in Eq. 2, corresponds to the lowest energy reconstruction.

Fig. 5(b) shows the surface energies of all 2D InP structures found by the algorithm versus their area densities. The choice of area density is motivated by experimental considerations as it is an experimentally accessible growth parameter, in contrast to the layer thickness. For this reason, and to provide an additional perspective of the 2D InP structures, we plot the surface energy as a function of area density rather than layer thickness. First of all, the six low-energy structures, (i) - (vi), also display low surface energies. Second, as can be seen in the figure, there are two minimums in surface energy with respect to area density, one around 0.14 atoms/Å^2 and a second one at about 0.27 atoms/Å^2. The structures near the low-density minimum correspond to single-layer structures, while the higher-density minimum is composed of structures that are better described as bilayer structures.

We observe that the lowest surface energies in Fig. 5(b) are about 40 meV/Å^2, which is roughly four times the values calculated by Björkman *et al.* for many transition metal dichalcogenides, whose bulk structures comprise weakly bonded layers.[31] The structure of bulk InP forms a 3D bonded network and lacks layered motifs, leading to the high surface energies of the 2D structures. The high formation and surface energies of the 2D InP structures considered here indicate that it is unlikely that these 2D materials could be synthesized as free-standing films,[6] although stabilization on substrates may be feasible.[28]

Overall, the surface energy provides an alternative objective function that does not require imposing a thickness constraint during the evolutionary algorithm search

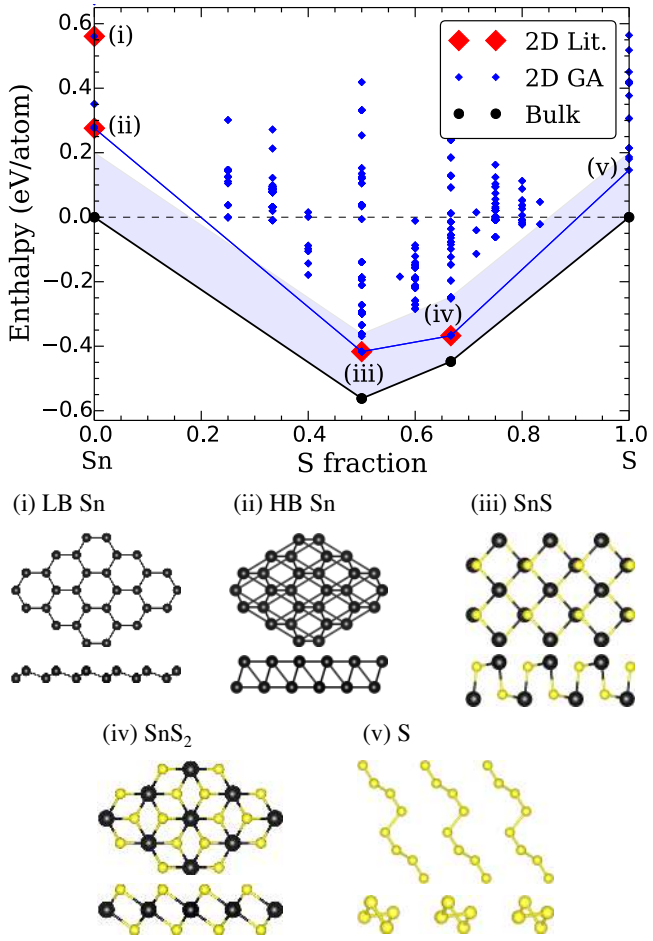

FIG. 6. Results of the search for 2D structures in the Sn-S system. The black circles denote the ground state bulk structures, and the lines connecting them form the convex hull for the bulk system. The light blue shading indicates the region less than 200 meV/atom above the bulk convex hull. The blue diamonds denote 2D structures found by the evolutionary algorithm, and the blue line segments form the convex hull for the 2D structures. The red diamonds, labeled (i) - (iv), denote 2D structures that have previously been reported in the literature. Top and side views of the structures labeled in the plot are shown below.

for 2D materials. This new objective function could be extended to variable composition 2D structure searches as well.

B. Tin-Sulfur 2D Phase Diagram

The grand canonical evolutionary algorithm enables the search for both composition and structure of unknown materials. To test this capability, we apply the algorithm to the binary Sn-S system, which exhibits two previously studied 2D materials, SnS and SnS_2, at intermediate composition. A two-dimensional structure of SnS_2 with potential application as a photocatalyst for water splitting has recently been synthesized by Sun *et al.*[3] and computationally characterized by Zhuang *et al.*[4] Tritsaris *et al.* and Singh *et al.* predicted the electronic properties of single-layer SnS, and calculated the binding energy between adjacent layers to be only 56 meV per unit cell[57] and the formation energy to be 144 meV/atom.[23]

To find the lowest energy structures across the composition range, we use the phase diagram searching mode of the evolutionary algorithm, as described by Tipton *et al.*[13] Two main modifications to the algorithm are required to search for low-energy structures and compositions across a phase diagram. First we permit structures with arbitrary stoichiometries in the search. Second, the objective function is replaced; instead of using the energy per atom, the objective function is now defined as a structure's vertical distance from the lowest convex hull known to the algorithm.

Fig. 6 shows the energies of the 2D structures in the Sn-S system found by the evolutionary algorithm relative to the energies of the ground state structures of bulk Sn and S. To visualize the formation energy of the 2D structures with respect to the bulk Sn-S phases at all compositions, the convex hull of the bulk structures is shown as well. The evolutionary algorithm recovers both the high[58] and low[59,60] buckled hexagonal structures of stanene, labeled (i) and (ii), respectively. The known 2D structures of SnS and SnS_2 were also recovered by the algorithm. 2D SnS occurs in the distorted rocksalt structure[23,57] and 2D SnS_2 displays the $1T$ structure common to many transition-metal dichalcogenides.[4] These two structures have the lowest formation energy relative to the bulk phase of any of the 2D structures found by the algorithm in this system. The lowest energy structure of pure S found by the algorithm is not a 2D structure at all, but rather consists of rows of 1D polymeric chains lying in a plane. This structure is labeled (v) in Fig. 6.

C. Carbon-Silicon 2D Phase Diagram

Graphene has been the subject of intense research since it was first successfully synthesized by Novoselov *et al.*[61] Two-dimensional silicon, or silicene, has also garnered attention more recently.[62–65] In addition, nanosheets of SiC have been experimentally reported,[66] and two different structures of buckled SiC monolayers have been predicted by Menon *et al.*[67] Li *et al.*[68] and Zhou *et al.*[69] have predicted completely planar structures of single-layer SiC_2. Because of the potential existence of several 2D structures in the C-Si system, we studied it using a 2D phase diagram search. We seeded the algorithm only with the known structures of planar hexagonal graphene and buckled hexagonal silicene in the initial generation. No other known 2D structures were provided.

Fig. 7 shows the energies of the 2D structures in the C-Si system found by the evolutionary algorithm relative to the energies of the ground state structures of bulk C and Si. Similar to Fig. 6 for the Sn-S system, we show the

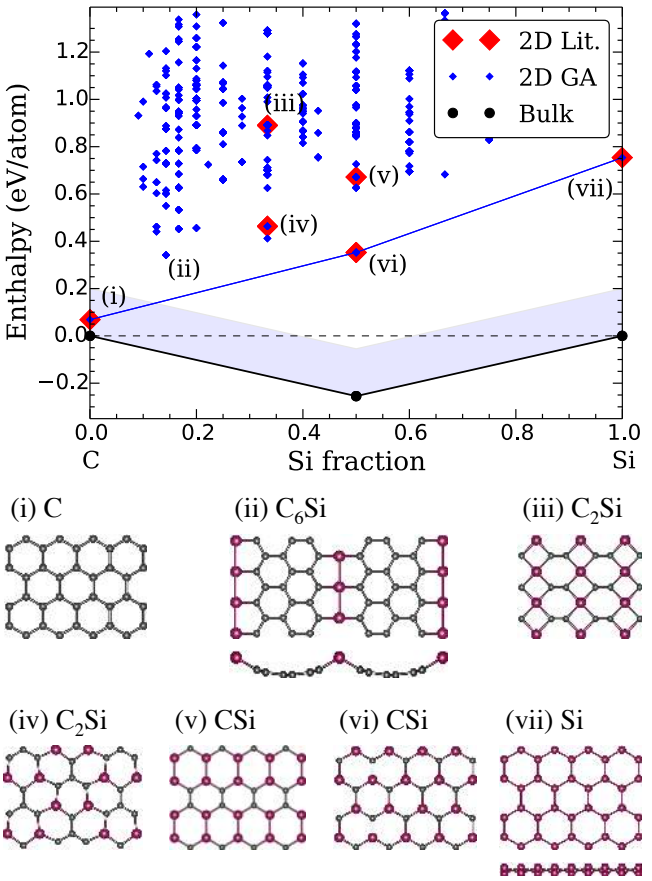

FIG. 7. Results of the search for 2D structures in the C-Si system. The symbols have the same meanings as in Fig. 6. The structures labeled (i) - (vii) in the plot are illustrated below. For the completely planar structures (i) and (iii)-(vi), side views are omitted.

convex hull of the bulk and 2D C-Si structures. Planar hexagonal graphene is labeled (i). The lowest energy C-Si structure found by the algorithm, labeled (ii), consists of nanoribbons of graphene joined by rows of Si atoms. The parameters for this structure are given in Table II.

The structures labeled (iii) - (vi) have all been previously reported in the literature and were all recovered by our evolutionary algorithm. Labeled (iii) is a structure of C_2Si, which consists of Si atoms bonded to four C atoms in a plane; this structure was predicted by Li *et al.*[68] A lower energy planar hexagonal structure of C_2Si, labeled (iv), was predicted by Zhou *et al.*[69] A planar hexagonal structure of CSi is labeled (v); Menon *et al.*[67] predicted a slightly buckled version of this structure. We found that the buckled version does not correspond to a local minimum in the energy landscape defined by the Hamiltonian (VASP-PBE) used in our study. Another lower energy planar hexagonal structure of CSi, consisting of alternating C and Si atoms, is labeled (vi); Lin[66] reported the synthesis of nanoflakes consisting of a few layers of this structure. The structure of buckled hexagonal silicene is labeled (vii).

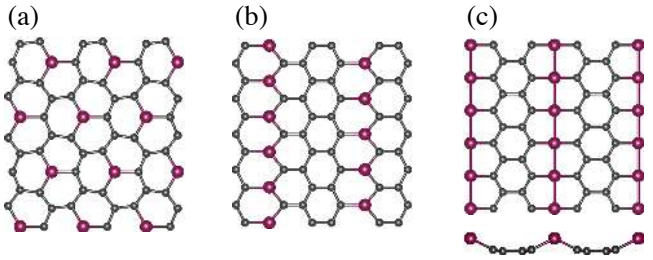

FIG. 8. Examples of three types of Si defects in graphene found by the evolutionary algorithm. The structures in (a) and (b) are completely planar and can be described as graphene with substitutional Si atoms. In (a), the substitutional Si atoms are located as far from each other as possible, while in (b) they are arranged in rows. In (c), the defect consists of 1D chains of four-fold coordinated Si atoms that are bonded to each other and two C atoms, with a slight distortion of the planarity of the graphene sheet.

With the exception of graphene, all of the 2D structures shown in Fig. 7 lie at least 400 meV/atom above the convex hull of the bulk system. We also note that the convex hull for 2D structures in this system is significantly skewed with respect to the bulk convex hull. Comparing the calculated formation energies of the structures at the endpoints of the 2D convex hull illustrates this point: graphene has a formation energy of only 69 meV/atom, while silicene has a formation energy of 754 meV/atom. Due to this gradient in thermodynamic instability across the composition range, we predict that carbon-rich 2D structures in this system hold the most promise for experimental synthesis.

Examining more closely the carbon-rich 2D structures found by the evolutionary algorithm, we identify three families of low-energy structures in this region of the phase diagram; representative members are illustrated in Fig. 8. Each of the three families corresponds to a different Si defect in graphene, with the members of a family having different defect densities. Fig. 8(a) and (b) shows that the first two defects are substitutional Si atoms, either isolated or arranged in rows. Fig. 8(c) shows that the third defect consists of 1D chains of four-fold coordinated Si atoms; each Si atoms is bonded to two other Si atoms along the straight chain and two C atoms. This defect interrupts the hexagonal graphene lattice and leads to out-of-plane distortions.

To analyze the formation of these defects, we consider the defect formation reaction

$$\mathrm{C}_n + \mathrm{Si}_m \rightleftharpoons \mathrm{C}_n\mathrm{Si}_m, \tag{3}$$

where C_n represents n carbon atoms forming graphene, Si_m represents m silicon atoms forming bulk diamond cubic silicon, and $\mathrm{C}_n\mathrm{Si}_m$ stands for the defective graphene structure, containing n carbon atoms and m silicon atoms. Given the reaction above, we determine the defect formation energy per Si atom as

$$\Delta E^f_{\mathrm{def}} = \frac{E(\mathrm{C}_n\mathrm{Si}_m) - mE(\mathrm{Si}) - nE(\mathrm{C})}{m}, \tag{4}$$

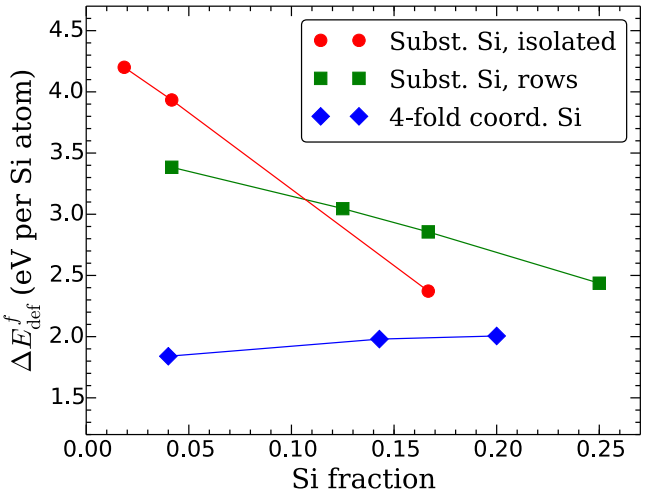

FIG. 9. Defect formation energies per Si atom for three types of defects found by the evolutionary algorithm in graphene, as a function of Si concentration.

where $E(C_nSi_m)$ is the total energy of the defective graphene structure, E(Si) is the energy per atom of bulk silicon, and E(C) is the energy per atom of graphene.

Fig. 9 shows the formation energies per Si atom as a function of Si fraction for the three types of defects in graphene identified with the evolutionary algorithm. Using the leftmost two data points in each curve, we estimate the formation energies of the isolated defects by linearly extrapolating to zero Si fraction. The resulting estimates of the defect formation energies are 4.4 eV/Si for an isolated substitutional Si atom, 3.6 eV/Si for a row of substitutional Si atoms, and 1.8 eV/Si for a chain of four-fold coordinated Si atoms. Interestingly, the formation energies of both 1D Si defects are lower than that of isolated Si substitutions, confirming the previously observed tendency of Si defect atoms to cluster in graphene.[70] In addition, the 1D defect formed by a chain of four-fold coordinated Si atoms exhibits significantly lower formation energy than the substitutional Si defects. These results demonstrate that the evolutionary algorithm can be employed to discover defect structures in 2D materials.

IV. CONCLUSION

We developed a grand canonical evolutionary algorithm for discovering low-energy structures in the emergent class of 2D materials. The algorithm enables both fixed and variable composition structure searches for 2D materials with finite thickness. The constraint on the layer thickness of the 2D structures in the search is tunable by the user. We applied the algorithm to search for 2D structures of InP, and it recovered the known buckled hexagonal structure, as well as several novel bilayer structures with lower formation energies. We further carried out variable composition searches on the Sn-S and C-Si binary systems, and the algorithm recovered the previously reported 2D structures in both of these systems. For the C-Si system, the algorithm also finds several structures corresponding to Si defects in graphene, including two new 1D defects with formation energies below that of a substitutional Si atom in graphene. Based on these successes, we believe the evolutionary algorithm for structure prediction is a useful tool to take the first step toward the computational discovery and design of novel 2D materials.

V. ACKNOWLEDGMENTS

We thank A. Singh for helpful discussions. This work was supported by the NSF under the CAREER award No. DMR-1056587 and under award No. ACI-1440547. This research used computational resources provided by the University of Florida Research Computing (http://researchcomputing.ufl.edu), the Texas Advanced Computing Center under Contract No. TG-DMR050028N and the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575. B.C.R. was supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1144153.

[1] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, Physical Review Letters **105**, 136805 (2010).
[2] A. Splendiani, L. Sun, Y. Zhang, T. Li, J. Kim, C. Y. Chim, G. Galli, and F. Wang, Nano Letters **10**, 1271 (2010).
[3] Y. Sun, H. Cheng, S. Gao, Z. Sun, Q. Liu, Q. Liu, F. Lei, T. Yao, J. He, S. Wei, and X. Y., Angewandte Chemie International Edition **51**, 8727 (2012).
[4] H. L. Zhuang and R. G. Hennig, Physical Review B **88**, 115314 (2013).
[5] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, Nature Nanotechnology **6**, 147 (2011).
[6] A. K. Singh, K. Mathew, H. L. Zhuang, and R. G. Hennig, The Journal of Physical Chemistry Letters (2015).
[7] H. Şahin, S. Cahangirov, M. Topsakal, E. Bekaroglu, E. Akturk, R. T. Senger, and S. Ciraci, Phys. Rev. B **80**, 155453 (2009).
[8] C. Ataca, H. ahin, and S. Ciraci, The Journal of Physical Chemistry C **116**, 8983 (2012).
[9] H. L. Zhuang and R. G. Hennig, JOM **66**, 366 (2014).
[10] S. Lebègue, T. Björkman, M. Klintenberg, R. M. Nieminen, and O. Eriksson, Physical Review X **3**, 031002 (2013).
[11] C. J. Pickard and R. J. Needs, Journal of Physics: Condensed Matter **23**, 053201 (2011).

[12] C. W. Glass, A. R. Oganov, and N. Hansen, Computer Physics Communications **175**, 713 (2006).
[13] W. W. Tipton and R. G. Hennig, Journal of Physics: Condensed Matter **25**, 495401 (2013).
[14] G. Trimarchi, A. J. Freeman, and A. Zunger, Physical Review B **80**, 092101 (2009).
[15] D. M. Deaven and K. M. Ho, Physical Review Letters **75**, 288 (1995).
[16] R. L. Johnston, Dalton Transactions , 4193 (2003).
[17] B. C. Revard, W. W. Tipton, and R. G. Hennig, in *Prediction and Calculation of Crystal Structures* (Springer, 2014) pp. 181–222.
[18] Y. Wang, J. Lv, L. Zhu, and Y. Ma, Physical Review B **82**, 094116 (2010).
[19] S. Bahmann and J. Kortus, Computer Physics Communications **184**, 1618 (2013).
[20] X.-F. Zhou, X. Dong, A. R. Oganov, Q. Zhu, Y. Tian, and H.-T. Wang, Physical Review Letters **112**, 085502 (2014).
[21] X. Luo, J. Yang, H. Liu, X. Wu, Y. Wang, Y. Ma, S.-H. Wei, X. Gong, and H. Xiang, Journal of the American Chemical Society **133**, 16285 (2011).
[22] W. Luo, Y. Ma, X. Gong, and H. Xiang, Journal of the American Chemical Society **136**, 15992 (2014).
[23] A. Singh and R. Hennig, Applied Physics Letters **105**, 042103 (2014).
[24] D. Teweldebrhan, V. Goyal, and A. Balandin, Nano Letters **10**, 1209 (2010).
[25] H. L. Zhuang and R. G. Hennig, Applied Physics Letters **101**, 153109 (2012).
[26] L. Spanu, S. Sorella, and G. Galli, Physical Review Letters **103**, 196401 (2009).
[27] H. Zhuang and R. Hennig, The Journal of Physical Chemistry C **117**, 20440 (2013).
[28] A. K. Singh, H. L. Zhuang, and R. G. Hennig, Physical Review B **89**, 245431 (2014).
[29] A. K. Singh and R. G. Hennig, Applied Physics Letters **105**, 051604 (2014).
[30] L. Shulenburger, A. D. Baczewski, Z. Zhu, J. Guan, and D. Tomanek, Nano Letters **15**, 8170 (2015).
[31] T. Björkman, A. Gulans, A. V. Krasheninnikov, and R. M. Nieminen, Phys. Rev. Lett. **108**, 235502 (2012).
[32] H. L. Zhuang and R. G. Hennig, Chemistry of Materials **25**, 3232 (2013).
[33] H. L. Zhuang, M. D. Johannes, M. N. Blonsky, and R. G. Hennig, Applied Physics Letters **104**, 022116 (2014).
[34] H. L. Zhuang and R. G. Hennig, Applied Physics Letters **103**, 212102 (2013).
[35] H. L. Zhuang, A. K. Singh, and R. G. Hennig, Physical Review B **87**, 165415 (2013).
[36] W. W. Tipton, C. R. Bealing, K. Mathew, and R. G. Hennig, Physical Review B **87**, 184114 (2013).
[37] W. W. Tipton, C. A. Matulis, and R. G. Hennig, Computational Materials Science **93**, 133 (2014).
[38] T. S. Bush, C. R. A. Catlow, and P. D. Battle, Journal of Materials Chemistry **5**, 1269 (1995).
[39] S. M. Woodley, P. D. Battle, J. D. Gale, and C. R. A. Catlow, Physical Chemistry Chemical Physics **1**, 2535 (1999).
[40] N. L. Abraham and M. I. J. Probert, Physical Review B **73**, 224104 (2006).
[41] A. R. Oganov and C. W. Glass, The Journal of Chemical Physics **124**, 244704 (2006).
[42] C. J. Pickard and R. J. Needs, Physical Review Letters **97**, 045504 (2006).
[43] I. Křivý and B. Gruber, Acta Crystallographica Section A **32**, 297 (1976).
[44] G. Kresse and J. Hafner, Physical Review B **47**, 558 (1993).
[45] G. Kresse and J. Furthmüller, Physical Review B **54**, 11169 (1996).
[46] P. E. Blöchl, Physical Review B **50**, 17953 (1994).
[47] G. Kresse and D. Joubert, Physical Review B **59**, 1758 (1999).
[48] J. P. Perdew, K. Burke, and M. Ernzerhof, Physical Review Letters **77**, 3865 (1996).
[49] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, Physical Review Letters **92**, 246401 (2004).
[50] G. Román-Pérez and J. M. Soler, Physical Review Letters **103**, 096102 (2009).
[51] J. Klimeš, D. R. Bowler, and A. Michaelides, Phys. Rev. B **83**, 195131 (2011).
[52] H. T. Stokes and D. M. Hatch, Journal of Applied Crystallography **38**, 237 (2005).
[53] A. Jain, S. Ong, G. Hautier, W. Chen, W. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. Persson, APL Materials **1**, 011002 (2013).
[54] H. Şahin, S. Cahangirov, M. Topsakal, E. Bekaroglu, E. Akturk, R. T. Senger, and S. Ciraci, Physical Review B **80**, 155453 (2009).
[55] C.-J. Tong, H. Zhang, Y.-N. Zhang, H. Liu, and L.-M. Liu, Journal of Materials Chemistry A **2**, 17971 (2014).
[56] T. Suzuki, Applied Physics Letters **107**, 213105 (2015).
[57] G. A. Tritsaris, B. D. Malone, and E. Kaxiras, Journal of Applied Physics **113**, 233507 (2013).
[58] P. Rivero, J.-A. Yan, V. M. García-Suárez, J. Ferrer, and S. Barraza-Lopez, Physical Review B **90**, 241408 (2014).
[59] Y. Xu, B. Yan, H.-J. Zhang, J. Wang, G. Xu, P. Tang, W. Duan, and S.-C. Zhang, Physical Review Letters **111**, 136804 (2013).
[60] F.-f. Zhu, W.-j. Chen, Y. Xu, C.-l. Gao, D.-d. Guan, C.-h. Liu, D. Qian, S.-C. Zhang, and J.-f. Jia, Nature Materials (2015).
[61] K. Novoselov, A. Geim, S. Morozov, D. Jiang, Y. Zhang, S. Dubonos, I. Grigorieva, and A. Firsov, Science **306**, 666 (2004).
[62] P. Vogt, P. De Padova, C. Quaresima, J. Avila, E. Frantzeskakis, M. C. Asensio, A. Resta, B. Ealet, and G. Le Lay, Physical Review Letters **108**, 155501 (2012).
[63] B. Aufray, A. Kara, S. Vizzini, H. Oughaddou, C. Leandri, B. Ealet, and G. Le Lay, Applied Physics Letters **96**, 183102 (2010).
[64] G. G. Guzmán-Verri and L. C. Lew Yan Voon, Physical Review B **76**, 075131 (2007).
[65] C. C. Liu, W. Feng, and Y. Yao, Physical Review Letters **107**, 076802 (2011).
[66] S. Lin, The Journal of Physical Chemistry C **116**, 3951 (2012).
[67] M. Menon, E. Richter, A. Mavrandonakis, G. Froudakis, and A. N. Andriotis, Physical Review B **69**, 115322 (2004).
[68] Y. Li, F. Li, Z. Zhou, and Z. Chen, Journal of the American Chemical Society **133**, 900 (2010).
[69] L. Zhou, Y. Zhang, and L. Wu, Nano Letters **13**, 5431 (2013).
[70] T. Susi, J. Kotakoski, D. Kepaptsoglou, C. Mangler, T. C. Lovejoy, O. L. Krivanek, R. Zan, U. Bangert, P. Ayala, J. C. Meyer, *et al.*, Physical Review Letters **113**, 115501 (2014).