Cluster expansion method for adsorption: Application to hydrogen chemisorption on graphene

Marcel H. F. Sluiter*

Laboratory for Advanced Materials, Institute for Materials Research, Tohoku University, Sendai, 980-8577 Japan

Yoshiyuki Kawazoe[†]

Institute for Materials Research, Tohoku University, Sendai, 980-8577 Japan (Received 20 April 2003; published 21 August 2003)

A systematic method is presented for determining the ground state of absorbed species on a substrate based on a cluster expansion of the configurational energy. It is shown that the method can determine the ground state of a strongly relaxing system using a few first-principles total energy calculations of small cells only. The method is applied to a particularly challenging case, the two-sided hydrogen chemisorption on a free standing graphene sheet, where, as a function of hydrogen coverage, the carbon hybridization goes from sp^2 to sp^3 . The method should require still fewer calculations and yield still more accurate results in the case of physisorption where longer-ranged strain effects are less important.

DOI: 10.1103/PhysRevB.68.085410

PACS number(s): 68.35.Md, 05.50.+q, 05.70.Np, 68.43.-h

I. INTRODUCTION

Chemisorption and physisorption are important areas of research in view of the many practical applications such as in catalysis and separation. When an isolated atom or molecule is absorbed on to a surface, there are usually only a few high symmetry configurations (structures) that are likely to be stable, see, e.g., Ref. 1. While relaxations and reconstructions make simulating the absorption of a single species onto a surface nontrivial, the number of distinct topologies remains rather limited. However, when many species absorb, this is no longer the case. The interactions between absorbed species in addition to the surface-species interactions make that very large periodic cells can become ground states even when the interactions themselves are rather short ranged. This is a well-known phenomenon in the study of the ground states of bulk substitutional alloys.² Fortunately, for bulk allovs there is an algorithm that has been highly successful for finding the most likely ground states.³ The algorithm is based on a cluster expansion of the energy of an alloy,⁴ and it is described in detail below. Studies on bulk alloys have shown that the cluster expansion converges most rapidly when strain effects are small.⁵ The question in applying the cluster expansion algorithm for the determination of ground-state configurations in absorption problems then is whether the large strains can be handled. As a test a rather severe example was selected: the case of two-sided hydrogen chemisorption onto a single graphene sheet.

In the graphene sheet the carbon valence orbitals are in the sp^2 hybridized state associated with in-plane bonds, whereas after chemisorption sp^3 hybridization occurs which demands tetrahedral coordination. Clearly, chemisorption must cause very large strains in the initially flat graphene sheet. Moreover, assuming a horizontally positioned graphene sheet, there are three possibilities associated with every carbon atom—no hydrogen, hydrogen above, and hydrogen below, making for a very large number of configurations even when a modest number of carbon sites is only considered. Chemisorption of hydrogen is also of interest in its own right and several studies have been devoted to it.¹ It should be remarked, however, that the large hydrogen absorption capacity reported for carbon nanotube materials^{6,7} is based on physisorption of H₂ molecules, rather than on chemisorption of single H atoms. However, as physisorption of H₂ on to graphene¹ and single-wall carbon nanotubes^{8–10} involves only minor strains, such a calculation is not nearly as stringent a test as chemisorption of H. It should be noted that in order to include finite-temperature effects, not only configurational aspects must be considered. Because adatoms have relatively much freedom of movement, it is to be expected that vibrational aspects play a significant role also.¹¹

II. CLUSTER EXPANSION ALGORITHM

The cluster expansion assumes that the energy can be written as a rapidly converging sum over cluster contributions where contributions from larger clusters become negligible. A practical example of this idea is common in organic chemistry where the formation enthalpy of a molecule can be expressed to a good approximation as a sum of contributions from nearest neighbor pairs (read bonds) only. In the case of ethane, one would count one C—C pair and six C—H pairs and estimate that the formation energy with respect to the isolated atoms is given as the sum of the single C—C bond energy and the six C—H bond energies. Contributions from larger clusters such as triplets usually provide only minor corrections.

In the case of bulk alloys, an expansion in terms of nearest-neighbor pairs only can explain some of the most commonly occurring ground states, but generally more distant pairs and clusters involving more than just two atoms have to be considered. For a bulk alloy with a fixed underlying crystal structure, each structure can be specified purely in terms of the occupancy of each site. In actuality one might represent the occupancy of a site i as a vector of dimension m, where m is the number of components, e.g., in case there are atom types A, B, and C, the occupancy is

$$\sigma_i = \begin{cases} (100), & A \text{ atom at site } i \\ (010), & B \text{ atom at site } i \\ (001), & C \text{ atom at site } i. \end{cases}$$
(1)

The concentrations c of the atomic species are then given simply the average of σ over all sites i, the expectation value of σ ,

$$c^{P} = \frac{1}{N} \sum_{i} \sigma_{i}^{P} = \langle \sigma^{P} \rangle, \qquad (2)$$

where the superscript *P* refers to the atomic species (*A*, *B*, or *C*) and the column index in the vector σ_i , and *N* represents the number of sites. The cluster correlations ξ_{α} are expressed as products of occupancies^{4,12} of the sites that make up the cluster α [see also Eq. (10) in Ref. 3], e.g., a nearest-neighbor pair (*p*1) correlation

$$\xi^{PQ}_{\alpha=p1} = \frac{1}{2zN} \sum_{i,j} \ ' \sigma^{P}_{i} \sigma^{Q}_{j} = \langle \sigma^{P}_{i} \sigma^{Q}_{j} \rangle_{p1}, \qquad (3)$$

where P and Q indicate the atomic species, z is the number of nearest-neighbor pairs per site (coordination number), and the prime in the sum refers to the restriction that sites *i* and *j* are nearest neighbors. The factor 2 is a combinatorial factor that takes care of the double counting that results from the reversal of indices *i* and *j*. In the thermodynamic limit, a description in terms of σ_i would be of little use because the specification of any arbitrary structure requires an infinite number of occupancies. Fortunately, we are interested in ground states, which tend to have periodic structures with unit cells with a limited number of sites only, so that such a structure can be represented in terms of just three translation vectors and only a limited number of occupancies. In such a unit cell it is feasible also to actually count the number of pairs and other clusters of a given type and determine exactly the correlations. The cluster expansion of the energy⁴ is given simply by

$$E^{s} = \sum_{\alpha} \xi^{s}_{\alpha} J_{\alpha}, \qquad (4)$$

where E^s is the energy per site for structure *s*, the sum runs over clusters α , and J_{α} is the effective cluster interaction (ECI) associated with cluster α . It should be noted that while the ECIs are generally unknowns, the energy of structures can be computed with first-principles methods such as density functional electronic structure methods, see e.g., Ref. 13. The correlations for a structure can be found simply by inspection. For the special case of a completely random structure, such as in a solid solution at high temperature, a particularly simple expression can be obtained

$$\xi_{\alpha}^{R} = \langle \sigma_{i} \rangle \langle \sigma_{j} \rangle \dots \langle \sigma_{m} \rangle = c^{N_{\alpha}}, \qquad (5)$$

where *R* indicates the random state, *i*, *j*, ... *m* are the N_{α} number of sites that form cluster α .

Having energies for a set of structures, the ECIs can be obtained by inverting the cluster expansion, provided that the summation over clusters α can be truncated. Moreover, to be of use, the cluster expansion must converge with cluster size, so that retaining in the sum terms corresponding to compact clusters with a few sites only gives a sufficiently accurate representation of the energy. Intuition suggests this to be the case as was illustrated by the chemical example for computing the formation enthalpy, which used pairwise terms only. However, in bulk alloys it is known that proper accounting of relaxation requires longer-ranged interactions and for superlattice structures in the long wave limit a reciprocal space treatment might be needed.⁵ Fortunately, we are interested in chemisorbed structures with relatively small periodicities so that a set of ECIs pertaining to compact clusters can be expected to suffice. The issue then is to select a small but adequately accurate set of clusters to represent the energies of structures. It is not really necessary to represent accurately all structures, but only those that are stable, or are almost stable.^{3,14,15} We can allow large errors for very unstable structures because in reality such structures will not occur.

Restricting ourselves to accurately representing the most stable structures only allows us to use a cluster expansion with yet a smaller set of clusters. This is evident in the chemical model for the formation energy of molecules also. The chemical model does not describe nearly as well atomic assemblies in which the atomic valences are not respected as it describes stable molecules. Once a choice is made of which clusters to use in the cluster expansion and additionally a set of structures is selected and the corresponding energies have been computed, the ECIs are obtained by inversion. Straightforward matrix inversion can be applied when the matrix ξ is square,¹⁶ that is, when the number *n* of unknown ECIs to be solved is equal to the number of structures n^{s} for which energies have been computed. This method is not very good because the cluster expansion is truncated and the neglected ECIs "pollute" the ECIs retained in the cluster expansion. Therefore, there should be more structures than ECIs, so that the inversion process can average out over the neglected ECIs. Hence, the ECI J_{α} for a cluster α are calculated with

$$\sum_{s} w^{s} \left[\Delta E^{s}(V) - \sum_{\alpha=1}^{n} J_{\alpha}(V) \xi_{\alpha}^{s} \right]^{2} = \text{minimal}, \qquad (6)$$

where *w* represents the weights assigned to each structure and *V* indicates volume or, more generally, strain. Naturally, ECIs can also be defined with reference to other thermodynamic potentials and the cluster expansion itself is applicable to a wide range of physical properties such as bulk and other elastic moduli, lattice parameters, and other properties that are sensitive to local atomic order. It has even been applied for finding structures with large band gaps.¹⁷ The method of assigning weights is more or less arbitrary as long as groundstate structures get a much larger weight than unstable structures. The method used here assigns weights according to

$$w^{s} = \frac{1}{1 + \omega \left(\frac{d^{s}}{\langle d \rangle}\right)^{2}},\tag{7}$$

where d^s represents the energy difference of a structure *s* to the convex hull formed by the ground-state ordered structures, and $\langle d \rangle$ is that energy difference averaged for all structures. Of course, d^s takes the value 0 if *s* is a ground state. ω is a factor, which is assigned the smallest positive value that insures that the total energies of all structures are in the correct order, as in the spirit of Ref. 14. In the actual calculations ω was given the value of 10. A slightly less arbitrary weighting method, which uses the idea of importance sampling, would give the structures a weight according to the likelihood of occurrence,

$$w^{s}(T) = \frac{e^{-\beta\Delta E^{s}} \text{ for } m/N^{s}_{sym}}{\sum_{s'} e^{-\beta\Delta E^{s'}} \text{ for } m/N^{s'}_{sym}},$$
(8)

where $\beta = (k_B T)^{-1}$, *T* is temperature and N_{sym}^s is the number of symmetry elements per site for structure *s*. In this scheme, the weights, and therefore the ECIs, are temperature dependent.

As was already mentioned, ECIs associated with clusters consisting of many sites should be much smaller than those associated with effective pair interactions. This physical intuition can be imposed on the ECIs in an elegant manner by adding an extra set of equations to those of Eqs. (6). These equations have the appearance of a cost or penalty function, with the ECIs corresponding to larger clusters having a larger penalty,

$$\sum_{\alpha=1}^{n} [W_{\alpha}J_{\alpha}]^2 = \text{minimal}, \qquad (9)$$

where W_{α} is the penalty for correlation α . A possible way to assign this penalty is with

$$W_{\alpha} = K z^{N_{\alpha}}, \tag{10}$$

where z is some number greater than unity, which forces the larger clusters to have smaller ECIs, and K is a scale factor that should be chosen such that the ECIs decay with cluster size without sacrificing too much in terms of quality of fit according to Eq. (6). As the contributions from larger clusters converged reasonably well, Eq. (9) was not actually used in this study.

An overdetermined system, such as Eq. (6) or Eqs. (6) and (9), is most reliably inverted with singular value decomposition (SVD). Aside from inverting an overdetermined system,^{18,19} SVD also allows us to simultaneously examine how significant the contribution is from a particular ECI.³ This means that we could start out with trying to obtain more ECIs than there are structures. Naturally, SVD can determine out of a set of n^s structures at utmost n^s ECIs, but generally not all of these ECIs will much contribute to an accurate

representation of the energy. Usually a nearest neighbor pair ECI contributes much, but e.g., a cluster with more than four sites could be expected to contribute very little. Some of the n^s ECIs, which do not much improve the fit, can then simply be omitted and a new SVD inversion is done using the most important ECIs only. Again there may be ECIs that contribute little to improve the fit and one can select an even smaller set of ECIs. Thus, SVD allows one to find the most significant ECIs from a set that is larger than the set of structures and in this way it helps selection of ECIs. Now we have approximate values for the ECIs; they are not exact because of the limited set of energies and the limited set of clusters employed.

It is important to note here that now that approximate ECIs are known, the energy of any structure can be predicted with the cluster expansion requiring as only extra information a knowledge of the correlations of this structure. As the latter is purely geometrical information, it can be tabulated easily for a very large set of structures including all those possible up to some given unit cell size. It is this feature, of easily computing energies from ECIs and ECIs from energies that is at the heart of the cluster expansion algorithm for ground-state searching.

In the algorithm one starts by computing from *ab initio* the energies of some structures, usually structures with small unit cells so that the *ab initio* calculations can be done speedily. One then extracts ECIs and uses these ECIs to compute the energies of all the tabulated structures. Usually, one then discovers that the cluster expansion predicts that there are structures with lower energies than those included in the prior *ab initio* calculation. The newly predicted stable structures are now computed by the *ab initio* method. Next again a cluster expansion is performed and the process repeats itself.

Every time this cycle of extracting ECIs and finding new low-energy structures is repeated, more ab initio energies become available. As a consequence, the cluster expansion gradually becomes more accurate and it becomes less likely that new ground-states can be found within the table of structures. When any new stable structures are no longer found by the cluster expansion the latter might be converged. This can be verified by repeating the above algorithm using a different starting set of structures, or by checking how well the cluster expansion is able to predict the energy of a structure that was not used for the extraction of the ECIs. The latter, the predictive error test,³ can be carried out conveniently as follows. Now that one has a set of n^s structures for which the energy was computed ab initio, one removes one structure from this set and extracts the ECIs from the set of $n^{s}-1$ structures. The ECIs are then used to predict the energy of the removed structure and this value is compared with the known ab initio value. Naturally, the two results should be close; if they are not; the cluster expansion is not likely to be converged. Then, one must consider using a larger set of clusters in the cluster expansion and as this larger set of clusters has more degrees of freedom, usually more ab initio calculations are going to be needed as well. When one has computed the predictive error by removing each structure in turn, and one is satisfied with the predictive error for every structure, and moreover one finds the same ground-states even when one starts with a completely different set of structures, then the search for the ground-state structures is assumed to be converged. Experience shows that convergence can be reached when the number of actually ab initio calculated structures is orders of magnitude smaller than the number of tabulated structures considered in the cluster expansion. As the ab initio calculations tend to be computationally demanding and as this algorithm is rather easily implemented in a computer program, this method of searching for the ground-state can be orders of magnitude faster than a direct search. Conversely, the current method can consider many more structures than a straightforward trial and error search by *ab initio* calculation. Actual calculations for bulk binary alloys show that a few tens of structures^{3,20} suffice for the highly sensitive test of a composition-temperature phase diagram.

III. COMPUTATIONAL DETAILS AND METHOD

Ab initio calculations of the hydrogen chemisorbed graphene structures were performed within the local density approximation using the Vienna ab initio simulation program (VASP).^{13,21} An orthorhombic unit cell was used in which the volume was held constant at 203 Å³. This resulted in a separation distance between the graphene layers of about 10 Å, which was verified to be sufficient not to pollute the energies with spurious intergraphene layer contributions. The exchange-correlation functional with generalized-gradient corrections²² was used. The calculations were performed using fully nonlocal optimized ultrasoft pseudopotentials²³ with energy cutoffs of 287 (200) eV and augmentation charge cutoff energies of 550 (400) eV for C (H). The cutoff energy for the wave functions has been set at 360 eV. Integrations in reciprocal space used а $6 \times 6 \times 3$ Monkhorst-Pack²⁴ grid in the first Brillouin zone, giving rise to 27 k points in the irreducible section. The Hermite-Gauss smearing method of Methfessel and Paxton²⁵ of order 1 has been used to accelerate k point convergence. The a and blattice parameters of the orthorhombic cell and the internal coordinates were optimized for all structures using the conjugate gradient method. Structural optimizations were considered converged when the greatest magnitude of the force on any atom was less than 1 meV/Å. The orthorhombic cell with eight carbon atoms, see Fig. 1, allows for $3^8 = 6561$ structures. The structures were given an identification number Q computed with

$$Q = \sum_{i=0}^{l} 3^{i} \sigma_{i}, \qquad (11)$$

where the position of the sites is as in Fig. 1 and the occupancy σ takes values 0, 1, 2 when H is above, H is below, or when there is no H associated with the site *i*. For example, structure 0000 corresponds to the case where there is a hydrogen atom above each carbon atom, while in structure 3280 there is a hydrogen atom below each carbon, and structure 6560 indicates a graphene sheet without hydrogen. Application of symmetry reduces the number of structures to just 528. However, even 528 structures are a daunting num-

ber for direct ab initio calculations, especially because each structure needs to be carefully structurally optimized. It is important to note that although the system consists of both hydrogen and carbon atoms, each structure can be represented by the occupancies σ_i of the carbon atoms only since all hydrogen atoms form only a single bond with one particular carbon atom. The correlations for the 528 structures could be easily tabulated using a computer program. The maximal clusters considered for the cluster expansion are the hexagon and the centered triangle. These maximal clusters were generated on the principle of the complete set of most compact clusters that are found within a sphere of a given radius. Vul and de Fontaine²⁶ gave a theoretical justification for this principle. We used a radius of 2.5 times the nearestneighbor carbon-carbon distance (which is about 1.4 Å). In Fig. 1 the hexagon can be recognized in the sites labeled 2 through 7 and the sites 1, 2, 3, and 5 represent the centered triangle. Considering all possible subclusters, also including the so-called empty cluster, gives a total of 15 clusters. As there are three species the number of ECIs is 104.

IV. RESULTS AND DISCUSSION

SVD was used to determine which of the 104 ECIs are most significant for representing the energy. After the cluster expansion had converged, meaning no new ground-states could be found and the predictive error reached a value of 12% of the range of formation enthalpies, about 46 ECIs were found to contribute significantly. At that point total energies for about 56 structures were used, see Table I. Only 31 ab initio calculations were required for these 56 structures because of symmetry between hydrogen atoms above and below carbon. This is more than one order of magnitude less than a full calculation of all the possible ground-states with the periodicity considered here would have required. The large strains are clearly reflected in the large number of ECIs that are required to describe the energetics, the relatively large predictive error and also in the rather large values of ECIs associated with multisite clusters.

The most stable structures are shown in Fig. 2. It is important to note that all these structures have 1:1 hydrogen carbon ratios. The most stable structure 0820 is completely in accordance with intuition because each carbon atom can adopt the ideal sp^3 bonding angles without any strain. The next most stable structures are not so easily guessed. Structures 0984 and 0328 are 56 and 103 meV/carbon atom above the 0820 structure. Both 0984 and 0328 are buckled. In the case of the former, the buckling puts one half of a carbon hexagon above the original graphene sheet and the other half below it, giving rise to a corrugated pattern where the grooves accommodate some of the strain involved in the sp^2 to sp^3 conversion. Structure 0328 also relieves some of the strain, but in this case by dimerization along one direction.

The most interesting result is that all structures have positive excess energies relative to the fully covered 0820 and the completely uncovered 6560 (pure graphene) structures, see Fig. 3. This means that none of the structures with partial hydrogen coverage can be stable with respect to pure graphene structure 0820, and hence that partial coverage is



FIG. 1. (Color) Top view of the orthorhombic cell with eight carbon atoms. For pure graphene the a, b, and c lattice parameters are 4.26, 4.92, and 9.68 Å, respectively.

not favored. The random state, which the cluster expansion indicates to be most stable when equal amounts of hydrogen are above and below the graphene sheet, is not stable with respect to structures 6560 and 0820 as is illustrated by the solid line in Fig. 3. It should be noted that at complete coverage the random state is higher in energy than structure 0820, in spite of both having equal number of hydrogen above and below the graphene sheet because randomness does not allow a perfect cancellation of strain energies at short-length scales. One normally thinks of the random state as a high-temperature solution phase. It being energetically unfavorable is generally an indication of a miscibility gap.



FIG. 3. Chemisorption energy of hydrogen onto graphene. Circles indicate the structures used in the structure expansion, some of the most stable structures are indicated. Solid line indicates a completely random distribution of hydrogen as computed with Eqs. (4) and (5). Pure graphene (structure 6560) and structure 0820 have been taken as reference points.

As only structures 6560 and 0820 are stable it follows that during hydrogenation islands of structure 0820 must form, which gradually grow at the expense of structure 6560. In pure 6560 and 0820 strains associated with sp^2 and sp^3 bonding are perfectly accommodated, but at the edges of structure 0820 this cannot be the case. Therefore, the perimeter of the 0820 islands must be associated with excess strain energy. This perimeter energy plays the same role as the interfacial energy in the theory of nucleation so that chemisorption



FIG. 2. (Color) Most stable hydrogenated graphene layers as determined with the cluster expansion.

TABLE I.	Total energy	E_{tot} and	number of a	carbon atom	s with hy	drogen at	oove ($\#C^{H}),$	hydrogen	below ($\#C_H$),	and	without l	nydrogen
(#C) for stru	ctures Q use	d for com	puting ECIs.	. Labelling o	of structur	res accord	ling to	Eq. 11						

				E _{tot}					E _{tot}
Q	$\#C^{H}$	$\#C_H$	# <i>C</i>	(eV/cell)	Q	$\#C^{H}$	$\#C_H$	# <i>C</i>	(eV/cell)
0000	8	0	0	-86.564	1102	1	6	1	-93.429
0001	7	1	0	-95.957	1112	2	4	2	-93.215
0002	7	0	1	-89.835	1640	4	0	4	-80.190
0008	6	0	2	-88.987	1646	3	0	5	-78.860
0010	6	2	0	-99.190	1700	2	0	6	-77.39
0012	6	2	0	-98.802	1932	2	2	4	-84.435
0024	6	0	2	-89.383	1968	2	0	2	-84.542
0080	4	0	4	-84.949	2024	2	0	6	-78.951
0084	6	2	0	-99.416	2168	2	0	6	-78.867
0087	6	1	1	-93.429	2186	1	0	7	-75.857
0091	5	3	0	-101.189	3280	0	8	0	-86.564
0140	4	2	2	-93.215	3281	0	7	1	-89.835
0168	6	0	2	-88.097	3284	0	6	2	-88.987
0218	5	0	3	-86.278	3292	0	6	2	-89.383
0328	4	4	0	-102.066	3320	0	4	4	-84.949
0656	4	0	4	-84.719	3364	0	6	2	-88.097
0728	2	0	6	-79.636	3389	0	5	3	-86.278
0820	4	4	0	-102.893	3608	0	4	4	-84.719
0821	4	3	1	-97.242	3644	0	2	6	-79.636
0823	3	5	0	-101.189	4100	0	4	4	-80.190
0826	3	4	1	-97.242	4103	0	3	5	-78.860
0840	4	4	0	-101.676	4130	0	2	6	-77.392
0850	2	6	0	-99.190	4264	0	4	4	-84.542
0948	4	2	2	-93.539	4292	0	2	6	-78.951
0984	4	4	0	-102.447	4364	0	2	6	-78.867
1012	2	6	0	-98.802	4373	0	1	7	-75.857
1024	2	4	2	-93.539	6560	0	0	8	-74.141
1093	1	7	0	- 95.957	1084	2	6	0	-99.416

must be accompanied by classical nucleation and growth phenomena.

The energy released by chemisorption at zero temperature can be roughly estimated by calculating the energy of the H_2 molecules in a large cell where inter H_2 interactions vanish. The zero temperature chemisorption energy which neglects all zero point librations and vibrations is about 0.2 eV per hydrogen atom relative to H_2 gas in near vacuum. This value agrees well with the value reported in Fig. 2 of Ref. 8.

The ECIs can be used to compute the most stable absorption states, at arbitrary compositions and temperatures, as well as the phase diagram, provided configurational aspects only suffice. The configurational partition function was solved with the cluster variation method^{2,27} employing the two maximal clusters depicted in Fig. 1. The calculations revealed that at any hydrogen coverage, the situation where equal numbers of hydrogen are above and below the graphene sheet always gives the lowest free energy. In retrospect, this is reasonable because such a distribution does not induce a macroscopic curvature in the graphene sheet. As discussed above, Fig. 3 indicates a miscibility gap in this system. The maximum of the miscibility gap occurs at a coverage of about 0.27 hydrogen atom/carbon atom and a temperature of 3240 K. Naturally, the neglect of vibrational contributions on the surface and of translational, librational, and vibrational contributions in the H₂ gas phase affects these numbers. At zero temperature, considering configurational aspects only, the chemisorbed structure (0820) is 0.2 eV per C and H atom more stable than graphene and H_2 gas. At finite temperature the neglected contributions to the free energy would stabilize the H₂ gas phase the most, so that the chemisorbed structure will become already unstable at a rather low temperature. Therefore, the miscibility gap computed here between graphene and structure 0820 in reality in likelihood evolves into a two-phase region between graphene and hydrogen gas. This means that, with the possible exception at very high H₂ pressures, random chemisorbed structures with only part of the carbon sites bonded with hydrogen do not occur at any temperature.

V. CONCLUSIONS

It has been demonstrated that an efficient cluster expansion technique can be applied to chemisorption and physisorption at surfaces and that important physical insights can be gleaned from such an analysis. Even for chemisorption at graphene, where strain effects are extremely large, one can obtain a cluster expansion that identifies the most stable surface structures. In the case of hydrogen onto graphene it was shown that a fully covered structure with alternating above and below graphene-plane hydrogen attachment is the most stable. This can be readily rationalized on the basis of strains associated with the sp^2 to sp^3 conversion for the carbon hybridization. Moreover, such an analysis shows that partial coverage is not favored. Therefore, chemisorbed regions must form islands where coverage is complete, which, as a consequence, must exhibit nucleation and growth phenomena also. A cluster variation analysis indicates that these features will remain at high temperature. After submitting this manuscript the authors learned that a cluster expansion gave

- *Electronic address: marcel@imr.edu; URL: http:// www-lab.imr.edu/~marcel
- [†]Electronic address: kawazoe@imr.edu; URL: http:// www-lab.imr.edu/~wkawazoe
- ¹J.S. Arellano, L.M. Molina, A. Rubio, and J.A. Alonso, J. Chem. Phys. **112**, 8114 (2000).
- ²D. de Fontaine: in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1979) Vol. 34, p. 73.
- ³M.H.F. Sluiter, Y. Watanabe, D. de Fontaine, and Y. Kawazoe, Phys. Rev. B **53**, 6137 (1996).
- ⁴D. de Fontaine: in *Solid State Physics*, edited by H. Ehrenreich and D. Turnbull (Academic, New York 1994), Vol. 47, p. 80.
- ⁵D.B. Laks, L.G. Ferreira, S. Froyen, and A. Zunger, Phys. Rev. B 46, 12 587 (1992).
- ⁶Y. Ye, C.C. Ahn, C. Witham, B. Fultz, J. Liu, A.G. Rinzler, D. Colbert, K.A. Smith, and R.E. Smalley, Appl. Phys. Lett. **74**, 2307 (1999).
- ⁷M.G. Nijkamp, J.E.M.J. Raaymakers, A.J. van Dillen, and K.P. de Jong, Appl. Phys. A: Mater. Sci. Process. **72**, 619 (2001).
- ⁸J.S. Arellano, L.M. Molina, A. Rubio, M.J. López, and J.A. Alonso, J. Chem. Phys. **117**, 2281 (2002).
- ⁹M. Shiraishi, T. Takenobu, and M. Ata, Chem. Phys. Lett. **367**, 633 (2003).
- ¹⁰M. Volpe and F. Cleri, Chem. Phys. Lett. **371**, 476 (2003).
- ¹¹A. van de Walle and G. Ceder, Rev. Mod. Phys. 74, 11 (2002).
- ¹²M. Sluiter, D. de Fontaine, X.Q. Guo, R. Podloucky, and A.J. Freeman, Phys. Rev. B 42, 10 460 (1990).

a good representation of lithium adatoms onto molybdenum (112) surfaces, as one would expect on the basis of this study.²⁸

ACKNOWLEDGMENTS

This work was performed under the interuniversity cooperative research program of the Laboratory for Advanced Materials, Institute for Materials Research, Tohoku University. The authors gratefully acknowledge the Center for Computational Materials Science at the Institute for Materials Research for allocations on the Hitachi SR8000 supercomputer system. The authors wish to thank Alain Pasturel for bringing to our attention to Ref. 28.

- ¹³G. Kresse and J. Furthmüller, Comput. Mater. Sci. 6, 15 (1996).
- ¹⁴G.D. Garbulsky and G. Ceder, Phys. Rev. B **51**, 67 (1995).
- ¹⁵G. Ceder, G.D. Garbulsky, and P.D. Tepesch, Phys. Rev. B **51**, 11 257 (1995).
- ¹⁶J.W.D. Connolly and A.R. Williams, Phys. Rev. B 27, 5169 (1983).
- ¹⁷S.-H. Wei, S.B. Zhang, and A. Zunger, Phys. Rev. B **59**, R2478 (1999).
- ¹⁸C. Wolverton and A. Zunger, Phys. Rev. B 50, 10 548 (1994).
- ¹⁹M. Asta, D. de Fontaine, and M. van Schilfgaarde, J. Mater. Res. 8, 2554 (1993).
- ²⁰M. Sluiter and Y. Kawazoe, Mater. Trans., JIM **42**, 2201 (2001).
- ²¹G. Kresse and J. Furthmüller, Phys. Rev. B 54, 11 169 (1996).
- ²²J.P. Perdew and Y. Wang, Phys. Rev. B **45**, 13 244 (1992).
- ²³G. Kresse and J. Hafner, J. Phys.: Condens. Matter 6, 8245 (1994).
- ²⁴H.J. Monkhorst and J.D. Pack, Phys. Rev. B 13, 5188 (1976).
- ²⁵M. Methfessel and A.T. Paxton, Phys. Rev. B 40, 3616 (1989).
- ²⁶D.A. Vul and D. de Fontaine, in Material Theory and Modeling, edited by J. Broughton, P.D. Bristowe, J.M. Newsam, Mater. Res. Soc. Symp. Proc. 291 (Materials Research Society, Pittsburgh, 1994), p. 401; D. de Fontaine, in *Solid State Physics*, edited by H. Ehrenreich and D. Turnbull (Academic, New York, 1994), Vol. 47 p. 84.
- ²⁷R. Kikuchi, Phys. Rev. **81**, 988 (1951).
- ²⁸R. Drautz, R. Singer, and M. Fähnle, Phys. Rev. B 67, 035418 (2003).