

Theoretical Study on Structural Stability of Alloy Cages: A Case of Silicon-Doped Heterofullerenes

Xiaofeng Fan^{1,*}, Zexuan Zhu¹, Lei Liu¹, Zexiang Shen¹ and
Jer-Lai Kuo^{1,2}

¹ School of Physical and Mathematical Sciences, Nanyang Technological University,
Singapore 637371, Singapore.

² Institute of Atomic and Molecular Sciences, Academia Sinica, Taipei, 10617,
Taiwan.

Received 10 December 2009; Accepted (in revised version) 26 January 2010

Communicated by Michel A. Van Hove

Available online 12 March 2010

Abstract. Structural stability and Si-substitution pattern in fullerene cage of $C_{60-n}Si_n$ are thoroughly investigated by integrating density functional calculations with a color-bond graph (CBG) model. We find that the parameterized CBG model with genetic algorithms can efficiently scan the large configuration space of alloy and therefore identify the low-energy region within the first-principles accuracy. Low-energy (stable) structures of $C_{60-n}Si_n$ in carbon-rich region ($1 \leq n \leq 30$) were identified and the silicon atoms are found to tend to aggregate in the fullerene cage. The mixing energy of these low-energy structures is ~ 35 meV/atom and insensitive to the Si concentration. We expect that these alloy fullerene cages can be synthesized experimentally at elevated temperatures.

PACS: 71.15.Mb, 82.33.Fg, 71.15.Nc, 36.40.Qv, 73.61.Wp, 61.46.Bc

Key words: Structure of alloy cluster, density functional calculations, fullerene Alloys $C_{60-n}Si_n$ Heterofullerenes, genetic algorithms.

1 Introduction

Nano-sized clusters have gained considerable attention for their peculiar geometry-induced optical, electronic, magnetic and catalytic properties [1–4]. As a special case of

*Corresponding author. *Email addresses:* xffan@ntu.edu.sg (X. Fan), zhuzexuan@pmail.ntu.edu.sg (Z. Zhu), LiuLei@ntu.edu.sg (L. Liu), zexiang@ntu.edu.sg (Z. Shen), jlkuo@pub.iams.sinica.edu.tw (J.-L. Kuo)

nanoclusters, cage-like carbon fullerenes have rapidly become an active area of research [5–10], since the discovery of fullerene C_{60} [11–13]. Due to the stable sp^2 -hybridized carbon bonds, these cage-like molecules have been considered as a building block for new nano-structured materials with novel properties [9, 12, 14, 15]. The exohedral and endohedral doping of metal atoms on C_{60} have led to the well known superconductivity of fullerenes [16, 17].

The substitutional doping is also expected to enable tailoring and functionalizing of the fullerene molecules [18–24]. Recently, doping carbon cages with silicon has been actively explored [21, 25–28]. Although having the same number of valence electron as carbon, silicon is known to prefer sp^3 -hybridization. If silicon atoms can be substitutionally doped in the carbon cages, the reactivity of Si sites will offer a new path toward the polymerization of fullerenes [21]. So far, the successful synthesis of $C_{60-n}Si_n$ ($n=1,3$) has been reported [21, 25, 29, 30]. While from photofragmentation experiments [31] it is speculated that up to 12 silicon atoms can be substituted in a fullerene without destroying its structure, clear experimental evidence for $C_{2N-n}Si_n$ ($n > 3$) still has not been obtained [28]. Thus, the systematic theoretical study on this multi-component system will be helpful and timely to illuminate its fundamental properties, such as the stability, morphology, electronic behavior, and basic features of chemical bonding.

To study alloy clusters, the first challenge is finding their ground state and low-energy structures. For the binary alloy clusters (A_nB_m), the number of combinatorial possibilities (2^{n+m}) due to different arrangement of A- and B- type atoms increases exponentially with the size of alloy clusters [2]. This makes it almost impossible to find the lowest energy structures of alloy cluster with high-level quantum calculation. For $C_{60-n}Si_n$ cages, searching systematically the ground states in rich carbon region is still a formidable task, even though the geometric structures are assumed established. By intuitively and empirically choosing the configurations, the structural and electronic properties of some "low-energy" states with different number of substitutional atoms (silicon) have been studied in a series of density function theory calculations [27, 32–37]. However, without an effective searching method it makes the theoretical studies blindfold to some extent and less comprehensive on the binding characteristics between silicon and carbon atoms on the cage structures.

In this work, we intend to fill up the gaps in the existing theoretical framework for exploring the stability of cage-like alloy structures. Using the alloy fullerene $C_{60-n}Si_n$ as a paradigm, the complexity of the question is illustrated. A parameterized model about total energy based on first-principles calculations is established with a few randomly chosen structures. Then the initial model is used to guide a systematic search powered by genetic algorithms (GA). The low-energy structures are identified to improve the accuracy of the model. This procedure is reiterated to assure the validity of our predictions. Lastly, the selected low-energy states of $C_{60-n}Si_n$ were examined in details to understand their structural stability and electronic properties.

2 Methods

2.1 The combinatorial complexity of alloy cages and purposed method

The simplest prototype $C_{58}Si_2$ [28] is considered to be difficult to perform by first-principles calculation due to numerous isomers ($C_{60}^2 = 1770$). Billas *et al.* construct 9 isomers based on the distance between two silicon atoms and study their stability with quantum calculations [30]. We revisit this issue and find that there are just 23 symmetrically distinct configurations for $C_{58}Si_2$ thanks to the high symmetry of fullerene (I_h). However, the complexity of alloy cages quickly increases with the number of substitutional silicon atoms. We find that there are 303 symmetrically distinct configurations in $C_{57}Si_3$. For $C_{30}Si_{30}$, there are about 1.2×10^{17} (C_{60}^{30}) isomers. In total, there are about 6.3×10^{17} configurations for $0 \leq n \leq 30$. Even using the symmetry group to eliminate the duplicated configurations, there are still about 5.2×10^{15} configurations. It is obvious that this enormous task cannot be completed by the direct application of quantum calculations. It is also too large to study with even simple pair potentials, though a larger search space can be searched for computationally cheaper potentials than quantum calculations.

To overcome the overmany configurations within the first-principles accuracy, a model about total energy with proper parameters is needed. From the view of chemical bonds, it is normally anticipated that the stability of the chemical bonds decides the stability of the molecule structure. The rules of chemical bonding, such as the octet rule, are very important for analyzing the stability of the chemical structure by intuition. Moreover, the stability of the alloy configurations can be studied by counting the number of different bonds. Actually, we have proposed that bond counting rules (BCR) can be considered as an effective guide in identifying the low-energy configurations in cage-like $C_{12}B_6N_6$ [38]. However, such a rule is so simple that it cannot quantitatively evaluate the stability of an alloy structure.

The same chemical bonds are well known to have different strength intensity and bond length depending on the chemical environments. For example, different chemical environments (such as the nearest-neighbor bond type, the angle between bonds) can induce different amount of charge transfer and thereby result in different local structural relaxation. It is therefore reasonable to assume that total energy should be correlated with the chemical bonds and its nearby local chemical environments. Thus, we generalize the idea of BCR to a new method named color-bond graph (CBG) theory. In CBG, the energy of the alloy configurations is recast in a general form:

$$E(\sigma) = E_0 + \sum_i \alpha_1^i C_1^i(\sigma) + \sum_{i,k} \alpha_{2,k}^i B_{2,k}^i(\sigma) + \sum_{i,k} \alpha_{3,k}^i A_{3,k}^i(\sigma) + \sum_{i,k} \alpha_{4,k}^i F_{4,k}^i(\sigma) + \dots, \quad (2.1)$$

where the parameters α 's represent the interaction energies, C_1 's are the atomic energy of vertex i ($i=1$ to N), B_2 's, A_3 's and F_4 's are the generalized correlation functions for different types of "bonds" to include contributions from two-vertex (the B_2 's), three-vertex (the A_3 's) terms and four-vertex (the F_4 's). In this work, the vertices are sorted into 2 colors

(carbon and silicon) and we include all two-vertex correlation functions, one three-vertex and one four-vertex correlation functions. In total, there are about 40 parameters (α 's) to fit to the first principle calculations.

After the effective model "Hamiltonian" with fitted parameters is obtained, we can tackle the selection of low-energy structures by different algorithms methods, such as Monte Carlo methods [39] and GA [40, 41]. The latter mimics the idea of Darwinian biological evolution and has been demonstrated to be a powerful paradigm for global optimization [42–44]. Inspired by the power of natural evolution, GA conducts global search by maintaining a population of candidate solutions and evolving the population by subjecting the solutions to selection, crossover, and mutation [41]. Firstly, an initial population of candidate configurations (randomly chosen $C_{60-n}Si_n$ with different concentration of silicon atoms) is created to ensure a uniform sampling throughout the search space. Each candidate configuration is represented by a string ('chromosome') of m integers ('genes') with each integer indicating the localization of a silicon atom at the corresponding vertex. The length of the chromosome, m , is determined by the concentration of Si atoms. Secondly, the fitness of each chromosome is then evaluated using Eq. (2.1). Two parent chromosomes are selected based on the fitness values so that the lower the energy of a chromosome the more likely it will be selected. Thirdly, new chromosomes, generated from previous ones by crossover and mutations, are then used to generate the next generation. Using uniform crossover, two new child chromosomes are generated from parent 1 and parent 2. Then the child chromosomes undergo random mutations. Finally, the new child chromosomes are added to the new population and the best chromosome from previous generation is guaranteed to survive in the new population.

2.2 Quantum chemistry method

The initial structures of Si-doped fullerenes are derived from the geometry of C_{60} . Preliminary investigations are first carried out using the method of modified neglect of differential overlap (MNDO) which is a semi-empirical method based on the neglect of differential diatomic overlap integral approximation for the quantum calculation of molecular electronic structure. The vibrational analyses after geometry optimization are engaged to ensure the optimized structures are true minimum in their energy landscapes. Previous investigations on carbon cage systems have shown electron correlation is important for the stability of structure [45, 46]. We employ B3LYP (Becke, three-parameter, Lee-Yang-Parr) exchange-correlation functional [47–49] with 3-21G basis to refine the MNDO results. A sufficiently complete basis set is important to deal with a practical system by the first-principle theoretical calculations. With the limited computational resource, it is compulsory to deal with the limited basis set for the large-system calculation. In order to get the reliable information about the structural and electronic properties, a small effective basis set, such as the composite basis set [50], is usually needed to test the effectiveness. To ensure that 3-21G basis can deliver reliable relative energies, we have examined all symmetrically distinct Si_2C_{58} cages with a larger basis (6-31G*). While the

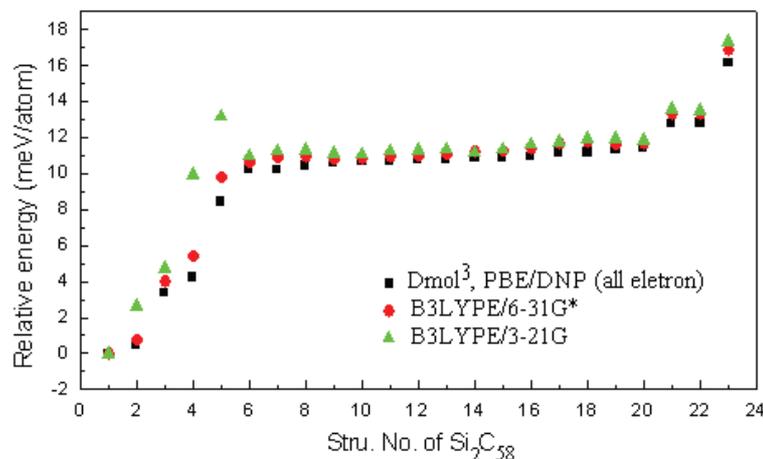


Figure 1: Relative energies of the 23 symmetrically distinct configurations of Si_2C_{58} with different basis sets and functionals.

3-21G basis set gives larger relative energy than those from 6-31G* basis set for low-energy metastable structures, it is obvious from Fig. 1 that the relative energies are not very sensitive to the basis set. Furthermore, we have also carried out calculations with PBE (Perdew-Burke-Ernzerhof) exchange-correlation functional with the largest basis set DNP (which is the double-numeric quality basis set with polarization functions) using DMol³ program [51,52] and the results indicate little dependence on the choice of functional as shown in Fig. 1. Here, the analysis of the structural stability is present with the B3LYP/3-21G method, and the electronic properties of the low-energy configurations are analyzed by the B3LYP/6-31G* method. All first-principles calculations reported in this work are performed with the GAUSSIAN 03 package [53] and DMol³ program [51,52].

3 Results and discussion

3.1 Parent structures and isomers with low Si concentration

The structures of C_{60} and Si_{60} have different characteristics. Strong hybridized quasi- sp^2 bonds of carbon stabilize the structure of C_{60} with high binding energy (tabulated in Table 1). On the other hand, silicon atom prefers sp^3 hybridization which leads to the distorted fullerene cage (as shown in Fig. 2) and smaller binding energy. Furthermore, other forms of Si_{60} , such as stuffed fullerene (SF) $\text{Si}_{20}\text{@Si}_{40}$, are possible to be more stable than Si_{60} fullerene cage. In this work, we will focus on the Si-substitution in carbon-rich region ($1 \leq n \leq 30$) in which we expect the fullerene structure would be maintained.

The fullerene cage structure has very high symmetry (I_h) and all sites in C_{60} are equivalent. For C_{59}Si_1 , there is only one unique isomer. The relaxed structure has an energy gain about 8.59 eV and is 0.09 eV/atom smaller than the binding energy of C_{60} . The

modifications of the bond lengths are mainly confined to Si-C bonds and elongation of two Si-C bonds elevates the Si atom as shown in Fig. 2. For $C_{58}Si_2$, the structures of all 23 symmetrically distinct configurations are fully optimized. The three most stable structures are found to have both Si atoms occupying in the same ring (hexagon or pentagon) shown in Fig. 2. In the most stable isomer $C_{58}Si_2$ -A, two Si atoms are arranged on parasites (the third nearest-neighbor positions) of a hexagonal ring which is consistent with the previous study and their experimental findings for disilabenzene complexes [21]. It should be noted here that the energetic difference (~ 4 meV/atom) among three most stable structures are very small thus making the confirmation of the most stable structure non-trivial [28].

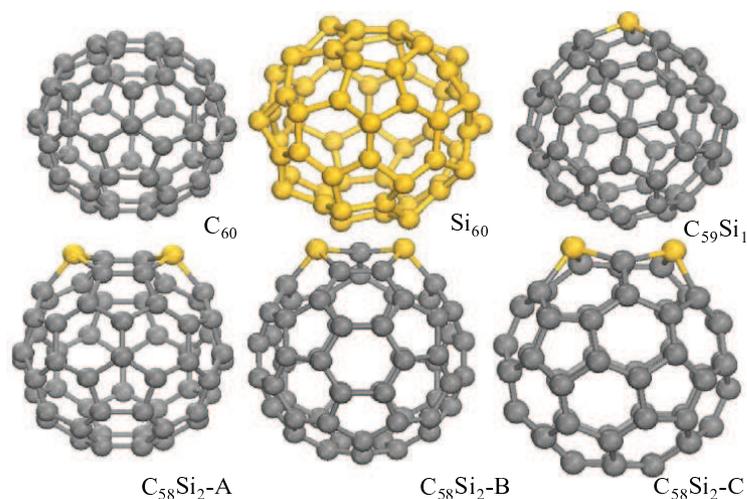


Figure 2: Structures of C_{60} , Si_{60} , $C_{59}Si_1$ and the low-energy isomers ($C_{58}Si_2$ -A, $C_{58}Si_2$ -B and $C_{58}Si_2$ -C) of $C_{58}Si_2$. Black spheres for C atoms, yellow spheres for Si atoms.

For further evaluation of the relative stability of alloy cage, we will use the notion of formation energy (ΔE_{form}) defined as

$$\Delta E_{\text{form}}(\sigma, C_{60-n}Si_n) = \left[E(\sigma, C_{60-n}Si_n) - \frac{60-n}{60} E(C_{60}) - \frac{n}{60} E(Si_{60}) \right] / 60, \quad (3.1)$$

where $E(C_{60})$ and $E(Si_{60})$ are the energy of C_{60} and Si_{60} fullerene cage. The formation energy is often referred as mixing energy which corresponds to the change of energy when constructing the alloy cage from the parent cages. A negative value of ΔE_{form} means the alloy cages are stable with respect to pure parent cages. The formation energies of $C_{59}Si_1$ and some isomers of $C_{58}Si_2$ are tabulated in Table 1. The positive mixing energies of these isomers indicate that these isomers are metastable structures and will not be thermodynamically stable at low temperature.

Table 1: Calculated energy, Binding energy, Formation energy and Relative energy of the parents with cage structures (C_{60} and Si_{60}), $C_{59}Si_1$ and three low-energy isomers of $C_{58}Si_2$.

Isomer stru.	Energy (Hartree)	Binding energy (eV/atom)	Formation energy (meV/atom)	Relative energy (eV/60-atoms)
C_{60}	-2273.5216	8.684	0.000	—
Si_{60}	-17278.9600	4.321	0.000	—
$C_{59}Si_1$	-2523.5679	8.591	20.104	—
$C_{58}Si_2$ -A	-2773.6392	8.509	28.872	0.000
$C_{58}Si_2$ -B	-2773.6334	8.507	31.506	0.158
$C_{58}Si_2$ -C	-2773.6287	8.505	33.619	0.284

3.2 Searching low-energy isomers with high Si concentrations

Following the increased concentration, existence of numerous isomers makes the systematic search for low-energy structures a challenging task. For each concentration n ($C_{60-n}Si_n$), there are $60!/[n!(60-n)!]$ configurations. Here we rely on the CBG theory to predict the low-energy structures. In order to obtain the parameters (α 's) in Eq. (2.1), we randomly choose 6-8 configurations for each concentration and there are about 200 configurations in total chosen in the range of concentration $2 \leq n \leq 30$. Then the structures of all these configurations are optimized by first-principles methods and the formation energies are extracted from the calculated total energies. By using the formation energies for the fitting of Eq. (2.1), the parameters (α 's) are obtained. The ability of the parameterized model (CBG) to capture the trends in energy of $C_{60-n}Si_n$ is confirmed in Fig. 3. The root-mean-squared (RMS) deviation of the model which is fitted from the extracted formation energy is ~ 8 meV/atom. Compared to the range of ~ 280 meV/atom (estimated from random-chosen structures), the ratio is just 2.8%. This means we can use the simple model with fitted parameters to predict the low-energy structures for each concentration.

The GA method is integrated with CBG model for searching the low-energy structures in the space of the enormous configurations. For each concentration, at least 5 low-energy isomers are identified and we find that the formation energies of these configurations are in the region of ~ 25 -75 meV/atom. To evaluate the validity of CBG model in the region of low formation energies, all the ~ 140 configurations are examined by first-principles calculations. It is found that RMS deviation of the predicted values by the model still is ~ 8 meV/atom. But the ratio of RMS deviation to the formation energy of the low energy structure is about 16% and this error is obviously large for the small energy region (25-75 meV/atom) of the low energy structures.

To reassure our description on ground-state structures, we use the data from low-formation energy region (~ 140 configurations) to improve Eq. (2.1). A training set of 80 configurations, shown as blue diamonds in Fig. 4, is chosen to re-parameterize Eq. (2.1). The RMS deviation of the fitting has gone down to ~ 3.4 meV/atom. The formation energies of other ~ 60 configurations (used as test set and shown as magenta diamonds

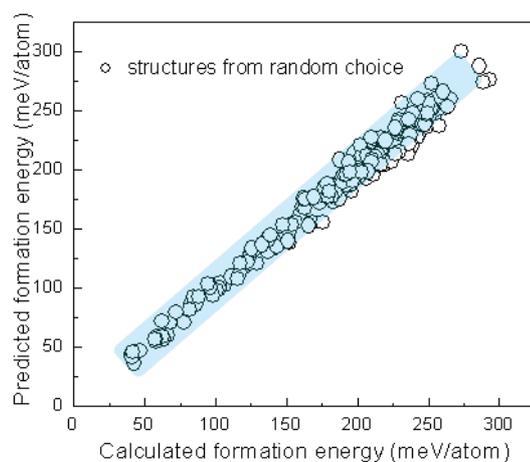


Figure 3: The effectiveness of CBG model for 200 randomly selected configurations: the x -axis is the DFT-calculated formation energies and the y -axis is the predicted formation energies from CBG model.

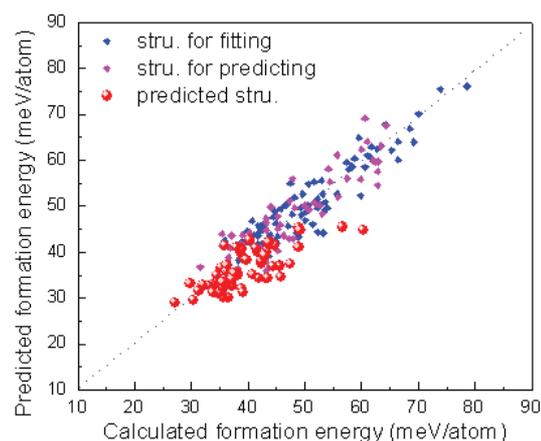


Figure 4: The effectiveness of CBG in predicting in low-energy region. The same convention as Fig. 3 is adapted here. The blue diamonds (~ 80 configurations) are used to re-parameterize Eq. (2.1) and the magenta diamonds (~ 60 configurations) are used as test set to verify the effectiveness of the CBG. Their RMS errors are 3.4 and 4.0 meV/atom respectively. Red sphere are the energies of new low-energy structure predicted with the revised CBG model.

in Fig. 4) are found to have a RMS deviation of ~ 4.0 meV/atom. Obviously, the error is reduced after the model is re-parameterized in the region of low energy. Thus, we will use this re-parameterized CBG model to explore the low-energy structures further, with GA method. There are 2 low-energy isomers identified for each concentration (in total 58 configurations). Then the structures of these low-energy isomers are optimized by first-principles methods and the formation energies are obtained from the calculated total energies. As shown by the red sphere in Fig. 4 and Fig. 5, the formation energies

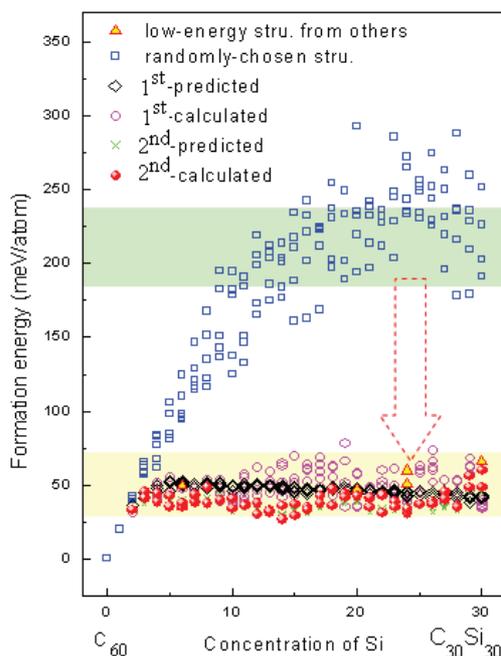


Figure 5: Formation energies as the function of the concentration of Si. Unlike the randomly selected isomers, we found that formation energy of low-energy configurations shows little dependency on the concentration of Si.

of the new low-energy isomers are found to locate at the positions which are just a little lower than that of the configurations searched in first time. This means the process for searching the low-energy structure has been converged and the obtained configurations can be considered to be the ground states of alloy fullerenes $C_{60-n}Si_n$.

The concentration dependence of all the formation energies calculated by first-principles method and predicted by the CBG model is shown in Fig. 5. The formation energies of these low energy isomers obtained finally by the model are found to be lower than that of the configurations chosen by the intuitive and empirical methods in literature. This shows the uncertainty of the empirical methods and the validity of the CBG model. At the same time, it is also demonstrated the validity of the combination of CBG model with genetic algorithm which can predict the low-energy ground states quickly by starting from the randomly-chosen configurations. From Fig. 5, it is also found that the formation energies of low-energy configurations show very little dependency on the concentrations of Si. This means the alloys with cage structures in each concentration ($2 \leq n \leq 30$) have same structural stabilities based on the stabilities of the structures of two parents (fullerene cage C_{60} and Si_{60}).

In the literature, an upper limit of 12-Si doping in the alloy cage is speculated based on mass spectrometry and photofragmentation studies [31, 54]. In another work, the upper limit is expended to 24-Si dopant by the analysis of dynamical instability [27]. The

formation energies from our calculation reveal no special features for $n = 12$ or 24 and therefore It is considered that the maxima number of Si atom to reside in the fullerene will depend on the dynamical stability of the alloy cage as the concentration of silicon increases.

3.3 Properties of the low-energy configurations

Ten lowest-energy states of some concentrations are chosen to analyze their properties in detail. The calculated energies, binding energies and formation energies are tabulated in Table 2. The binding energies are between that of the pure cage C_{60} and Si_{60} . This means the structures of alloy cages are relatively stable like the parent cage structure. All these low-energy structures have positive formation energies at the temperature 0 K. From the energetic point of view, all the alloy cages are metastable structures and tend to separate into pure parent cage structures at low temperature. Fortunately, the values of positive formation energies of these alloy structures are not very large (~ 35 meV). The thermal energy induced by several hundred degree of temperature can counteract the effect of positive formation energy. In other words, our calculations suggest that in the range of concentrations of silicon ($1 \leq n \leq 30$), the alloy cages are likely to be produced by recent experiment techniques, such as laser ablation or vaporization.

Table 2: Energy of ten most-stable cage alloy structures chosen for different concentration of silicon (structures are shown in Fig. 6).

Structure	Energy (Hartree)	Binding energy (eV/atom)	Formation energy (meV/atom)
$C_{55}Si_5$	-3523.8958	8.284	35.827
$C_{54}Si_6$	-3773.9877	8.212	35.247
$C_{50}Si_{10}$	-4774.3518	7.922	34.556
$C_{48}Si_{12}$	-5274.5425	7.781	30.278
$C_{46}Si_{14}$	-5774.7309	7.639	27.061
$C_{42}Si_{18}$	-6775.0738	7.339	35.986
$C_{38}Si_{22}$	-7775.4464	7.053	31.400
$C_{36}Si_{24}$	-8275.6284	6.908	31.095
$C_{34}Si_{26}$	-8775.7966	6.756	37.038
$C_{30}Si_{30}$	-9776.1649	6.468	35.827

The structures of ten lowest-energy states with different concentrations are depicted in Fig. 6. An obvious trend is all the low-energy alloy configurations are at phase-separated states and the chemical ordering rule of bulk SiC don't exist in fullerene cage with 60 atoms. In bulk SiC, the existent of chemical ordering means sp^3 -hybridized Si-C bond are more stable than the average of Si-Si and C-C bonds. But this notion is not true for quasi- sp^2 hybridized Si-C bonds in the fullerene cage. This may be ascribed to two important reasons. Firstly, we consider the strain effect of the different sp^3 -hybridized bonds in bulk Si_xC_{1-x} and that of the different sp^2 -hybridized bonds in heterofullerenes

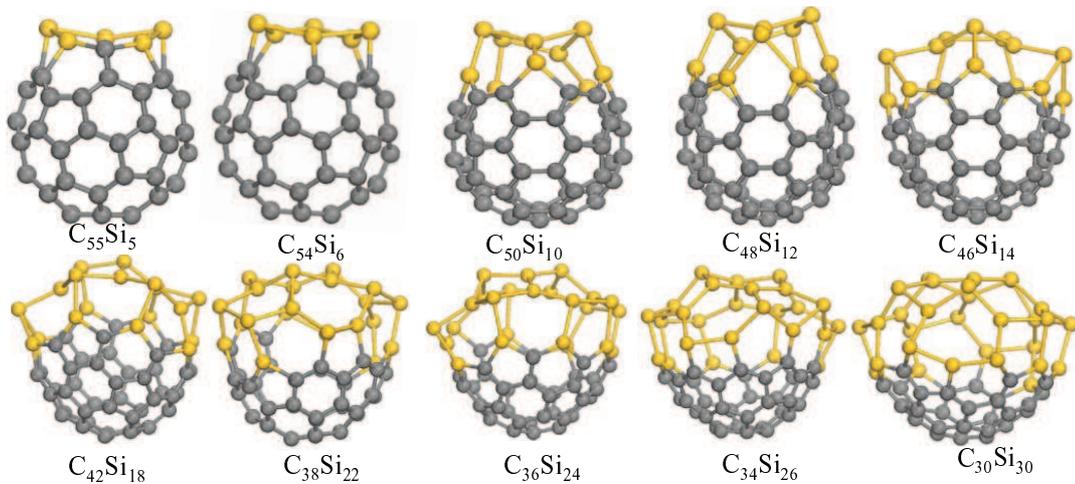


Figure 6: First-principles optimized structures of the ten most-stable cage alloy structures with different concentration of silicon predicted by CBG model.

$C_{60-n}Si_n$. Due to the bond-length difference of Si-C, C-C and Si-Si bonds, these highly-directional covalent bonds will result in the large strain energy in the Si-C alloy systems. In order to reduce the strain energy, the alloy phases will separate to annihilate the Si-C bonds, except that the binding energy of Si-C is large enough that the chemical energy gain can counteract the strain energy during the formation of Si-C bonds. Actually, it is the strain energy that makes the formation of alloy difficult in most semiconductor alloy, such as BeZnO, CdZnO and ZnSO alloys [55–57], under thermodynamic equilibrium condition. For bulk Si_xC_{1-x} alloys, the chemical ordering exists only for those with the 50% Si concentration as a result of the elimination of Si-Si and C-C bonds in the four-coordinated crystal structures. However, for heterofullerenes $C_{60-n}Si_n$, the Si-Si (or C-C) bond is impossible to be eliminated completely due to the 5-number ring in the fullerene cages. Secondly, the tendency of the sp^3 -hybridization of Si bonds and the large size of Si atoms need to be considered. The two factors will result in the distortion and thereby the strain energy of cages. Therefore, aggregation of Si atoms in the cage structure can reduce strain energy by changing the position in a larger local space of the cage structure.

Electronic properties of these cage-like cluster $C_{60-n}Si_n$ ($0 \leq n \leq 30$) are studied by examining their highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) orbital. As shown in Fig. 7, we can find the gaps between HOMO and LUMO quickly reduce to a value that is near the gap of Si_{60} , following the increase of the concentration of Si. This trend is consistent with the observation that HOMO/LUMO orbitals have main contribution from the Si atoms of the cage. This means the gaps of alloy clusters are more similar to that of Si_{60} . We should also note here that the HOMO-LUMO gap is sensitive to the substitutional positions. For an example, the gaps of three low-energy structures of Si_2C_{58} are 1.63, 2.21 and 2.13 eV, respectively.

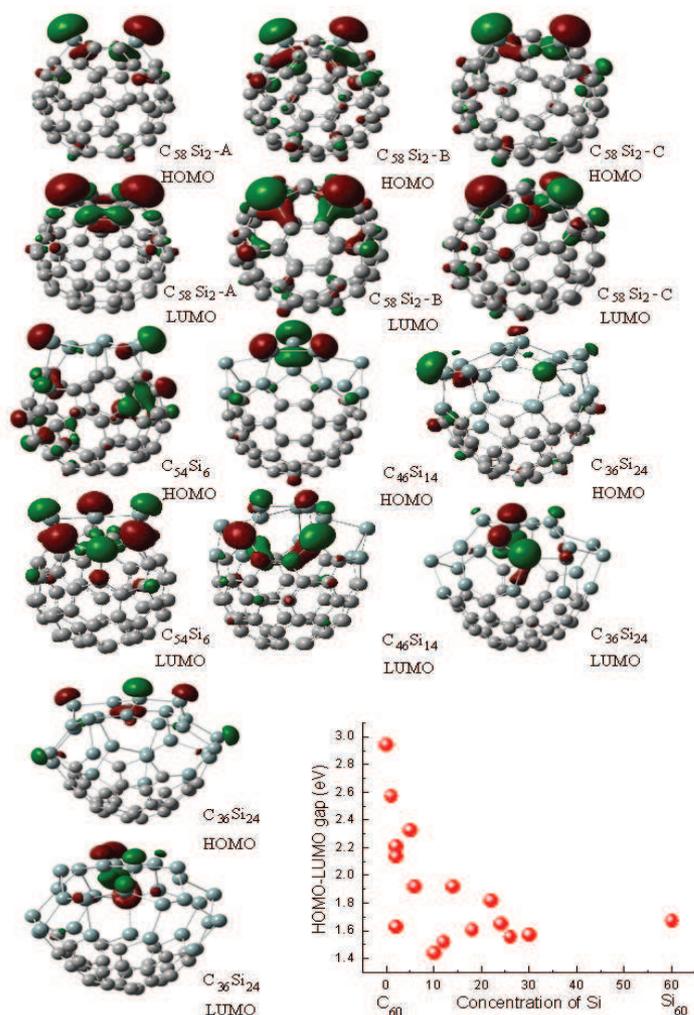


Figure 7: Molecular orbitals (HOMO and LUMO) of $C_{58}Si_2$ -A, $C_{58}Si_2$ -B, $C_{58}Si_2$ -C, $C_{54}Si_6$, $C_{46}Si_{14}$, $C_{36}Si_{24}$, and $C_{30}Si_{30}$. It is obvious that the spatial distribution of both HOMO and LUMO are concentrated on Si atoms. This spatial preference can explain that HOMO-LUMO gaps of stable low-energy configurations (shown in the bottom) quickly reduce to a value very close to Si_{60} .

This means there should be a distribution of gaps in experimental sample due to the co-existence of different substitutional patterns.

4 Conclusions

The size, shape and combinatorial atom arrangement make the theoretical simulation of the alloy clusters complicated. Especially, the combinatorial possibilities can result

in a very large configuration space (2^{n+m}). Using the alloy fullerene $C_{60-n}Si_n$ as a paradigm, we attempt to tackle the large alloy configuration space ($\sim 1.15 \times 10^{18}$) by the first-principles calculation with GA. Based on the CBG theory, a model "Hamiltonian" has been built to search the minimums of the total energy of the alloy cage. From first-principles calculations, the parameters in the model are established. Then the validity of the parameterized model "Hamiltonian" is verified. With the GA method, the model can efficiently guide GA to locate the low-energy region for cage structures. By optimizing the parameters in the model further itinerantly, the low-energy stable states can be obtained for each concentration.

For fullerene cage $C_{60-n}Si_n$, all the alloy cages are studied systematically in the C-rich region. In all concentrations ($2 < n \leq 30$), silicon atoms have the tendency of aggregation and the chemical ordering doesn't exist in the fullerene cage of 60 atoms. Very small positive mixing energy (~ 35 meV/atom) can be easily counteracted by the thermal effect and therefore we expect that the alloy cages can be produced experimentally.

Acknowledgments

This work was supported from Academia Sinica and Nanyang Technological University.

References

- [1] Klabunde, K. J. *Free Atoms, Clusters, and Nanoscale Particles*; Academic: New York, 1994.
- [2] Jellinek, J.; Krissinel, E. B. *Theory of Atomic and Molecular Clusters*; Springer: Berlin, 1999.
- [3] Andrews, M. P.; O'Brien, S. C. *J. Phys. Chem.* 1992, 96, 8233.
- [4] Jellinek, J. *Faraday Discuss.* 2008, 138, 11.
- [5] Cox, D. M.; Behal, S.; Disko, M.; Gorun, S. M.; Greaney, M.; Hsu, C. S.; Kollin, E. B.; Millar, J.; Robbins, J.; Robbins, W.; Shewood, R. D.; Tindall, P. J. *Am. Chem. Soc.* 1991, 113, 2940.
- [6] Allemand, P.-M.; Khemani, K. C.; Koch, A.; Wudl, F.; Holczer, K.; Donovan, S.; Gruner, G.; Thompson, J. D. *Science* 1991, 253, 301.
- [7] Haddon, R. C.; Hebard, A. F.; Rosseinsky, M. J.; Murphy, D. W.; Duclos, S. J.; Lyons, K. B.; Miller, B.; Rosamilia, J. M.; Fleming, R. M.; Kortan, A. R.; Giarum, S. H.; Makhija, A. V.; Muller, A. J.; Eick, R. H.; Zahurak, S. M.; Tycko, R.; Dabbagh, G.; Thiel, F. A. *Nature* 1991, 350, 320.
- [8] Schön, J. H.; Kloc, C.; Haddon, R. C.; Batlogg, B. *Science* 2000, 288, 656.
- [9] Yang, W. L.; Brouet, V.; Zhou, X. J.; Choi, H. J.; Louie, S. G.; Cohen, M. L.; Kellar, S. A.; Bogdanov, P. V.; Lanzara, A.; Goldoni, A.; Parmigiani, F.; Hussain, Z.; Shen, Z.-X. *Science* 2003, 300, 303.
- [10] Dinnebier, R. E.; Gunnarsson, O.; Brumm, H.; Koch, E.; Stephens, P. W.; Huq, A.; Jansen, M. *Science* 2002, 296, 109.
- [11] Kroto, H. W.; Heath, J. R.; O'Brien, S. C.; Curl, R. F.; Smalley, R. E. *Nature* 1985, 318, 162.
- [12] Krätschmer, W.; Lamb, L. D.; Fostiropoulos, K.; Huffman, D. R. *Nature* 1990, 347, 354.
- [13] Taylor, R.; Hare, J. P.; Abdul-Sada, A. K.; Kroto, H. W. *J. Chem. Soc., Chem. Commun.* 1990, 1423.

- [14] Paillard, V.; Melinon, P.; Dupuis, V.; Perez, J. P.; Perez, A.; Champagnon, B. *Phys. Rev. Lett.* 1993, 71, 4170.
- [15] Xie, R.-H.; Bryant, G. W.; Zhao, J.; Smith, V. H.; Jr.; Carlo, A. D.; Pecchia, A. *Phys. Rev. Lett.* 2003, 90, 206602.
- [16] Johnson, R. D.; Vries, M. S. d.; Salem, J.; Bethune, D. S.; Yannoni, C. S. *Nature* 1992, 355, 239.
- [17] Dresselhaus, M. S.; Dresselhaus, G.; Eklund, P. C. *Science of fullerenes and carbon nanotubes*; Academic: New York, 1996.
- [18] Taylor, R.; Walton, D. R. M. *Nature* 1993, 363, 685.
- [19] Hummelen, J. C.; Knight, B.; Pavlovich, J.; González, R.; Wudl, F. *Science* 1995, 269, 1554.
- [20] Tast, F.; Malinowski, N.; Frank, S.; Heinebrodt, M.; Billas, I. M. L.; Martin, T. P. *Phys. Rev. Lett.* 1996, 77, 3529.
- [21] Ray, C.; Pellarin, M.; Lermé, J. L.; Vialle, J. L.; Broyer, M.; Blase, X.; Mélinon, P.; Kéghélian, P.; Perez, A. *Phys. Rev. Lett.* 1998, 80, 5365.
- [22] Hultman, L.; Stafström, S.; Czigány, Z.; Neidhardt, J.; Hellgren, N.; Brunell, I. F.; Suenaga, K.; Colliex, C. *Phys. Rev. Lett.* 2001, 87, 225503.
- [23] Schultz, D.; Droppa, R.; Jr.; Alvarez, F.; Santos, M. C. d. *Phys. Rev. Lett.* 2003, 90, 015501.
- [24] Otero, G.; Biddau, G.; Sánchez-Sánchez, C.; Caillard, R.; López, M. F.; Rogero, C.; Palomares, F. J.; Cabello, N.; Basanta, M. A.; Ortega, J.; Mendez, J.; Echavarren, A. M.; Pérez, R.; Gómez-Lor, B.; Martín-Gago, J. A. *Nature* 2008, 454, 865.
- [25] Kimura, T.; Sugai, T.; Shinohara, H. *Chem. Phys. Lett.* 1996, 256, 269.
- [26] Fye, J. L.; Jarrold, M. F. J. *Phys. Chem. A* 1997, 101, 1836.
- [27] Matsubara, M.; Kortus, J.; Parlebas, J.-C.; Massobrio, C. *Phys. Rev. Lett.* 2006, 96, 155502.
- [28] Mélinon, P.; Masenelli, B.; Tournus, F.; Perez, A. *Nature Materials* 2007, 6, 479.
- [29] Pellarin, M.; Ray, C.; Mélinon, P.; Lermé, J.; Vialle, J. L.; Kéghélian, P.; Perez, A.; Broyer, M. *Chem. Phys. Lett.* 1997, 277, 96.
- [30] Billas, I. M. L.; Massobrio, C.; Boero, M.; Parrinello, M.; Branz, W.; Tast, F.; Malinowski, N.; Heinebrodt, M.; Martin, T. P. *J. Chem. Phys.* 1999, 111, 6787.
- [31] Pellarin, M.; Ray, C.; Lermé, J.; Vialle, J. L.; Broyer, M.; Blase, X.; Kéghélian, P.; Mélinon, P.; Perez, A. *J. Chem. Phys.* 1999, 110, 6927.
- [32] Billas, I. M. L.; Branz, W.; Malinowski, N.; Tast, F.; Heinebrodt, M.; Martin, T. P.; Massobrio, C.; Boero, M.; Parrinello, M. *Nanostruct. Mater.* 1999, 12, 1071.
- [33] Menon, M. J. *Chem. Phys.* 2001, 114, 7731.
- [34] Fu, C.-C.; Weissmann, M.; Machado, M.; Ordejón, P. *Phys. Rev. B* 2001, 63, 085411.
- [35] Matsubara, M.; Massobrio, C. *J. Phys. Chem. A* 2005, 109, 4415.
- [36] Marcos, P. A.; Alonso, J. A.; Lopèz, M. J. *J. Chem. Phys.* 2005, 123, 204323.
- [37] Matsubara, M.; Massobrio, C. *J. Chem. Phys.* 2005, 122, 084304.
- [38] Fan, X. F.; Zhu, Z. X.; Shen, Z. X.; Kuo, J.-L. *J. Phys. Chem. C* 2008, 112, 15691.
- [39] Landau, D. P.; Binder, K. *A guide to monte carlo simulations in statistical physics*; Cambridge University Press: Cambridge, 2005.
- [40] Michalewicz, Z.; Fogel, D. B. *How to solve it: modern heuristics*; Springer: Berlin, 2000.
- [41] Holland, J. H. *Adaptation in natural artificial systems*, 2nd edition; MIT Press, 1992.
- [42] Pardalos, P. M.; Shalloway, D.; Xue, G. *J. Global Optim.* 1994, 4, 117.
- [43] Deaven, D.; Ho, K. M. *Phys. Rev. Lett.* 1995, 75, 288.
- [44] Hobday, S.; Smith, R. J. *Chem. Soc., Faraday Trans.* 1997, 93, 3919.
- [45] Martin, J. M. L.; Taylor, P. R. *J. Phys. Chem.* 1994, 98, 6105.
- [46] Kar, T.; Cuma, M.; Scheiner, S. *J. Phys. Chem. A* 1998, 102, 10134.
- [47] Becke, A. D. *J. Chem. Phys.* 1992, 96, 2155.

- [48] Becke, A. D. J. Chem. Phys. 1993, 98, 5648.
- [49] Cohen, A.; Handy, N. C. Chem. Phys. Lett. 2000, 316, 160.
- [50] Zhang, R. Q.; Wong, N. B.; Lee, S. T.; Zhu, R. S.; Han, K. L. Chem. Phys. Lett. 2000, 319, 213.
- [51] Delley, B. J. Chem. Phys. 1990, 92, 508.
- [52] Delley, B. J. Chem. Phys. 2000, 119, 7756.
- [53] Frisch, G. W. T. M. J.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A.; Gaussian, Inc.: Wallingford, CT, 2004.
- [54] Pellarin, M.; Ray, C.; Lermé, J.; Vialle, J. L.; Broyer, M.; Blase, X.; Kéghélian, P.; Mélinon, P.; Perez, A. Eur. Phys. J. D 1999, 9, 49.
- [55] Fan, X. F.; Zhu, Z. X.; Ong, Y.-S.; Lu, Y. M.; Shen, Z. X.; Kuo, J.-L. Appl. Phys. Lett. 2007, 91, 121121.
- [56] Fan, X. F.; Sun, H. D.; Shen, Z. X.; Jer-LaiKuo; Lu, Y. M. J. Phys.: Condens. Matter 2008, 20, 235221.
- [57] Fan, X. F.; Shen, Z. X.; Lu, Y. M.; Kuo, J.-L. New Journal of Physics 2009, 11, 093008.