Ab initio global optimization of the structures of Si_nH, $n=4-10$, using parallel genetic algorithms

Ofelia Oña,[1] Victor E. Bazterra,[1,2] María C. Caputo,[2] Marta B. Ferraro,[2] and Julio C. Facelli[1,*]
[1]*Center for High Performance Computing, University of Utah, 155 South 1452 East Rm 405, Salt Lake City, UT 84112-0190*
[2]*Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pab. I (1428), Buenos Aires, Argentina*
(Received 3 August 2005; published 18 November 2005)

The results of *ab initio* global optimizations of the structures of Si_nH, $n=4-10$, atomic clusters using a parallel genetic algorithm are presented. Driving the global search with the parallel implementation of the genetic algorithm are presented and using the density functional theory as implemented in the Carr-Parinello molecular dynamics code to calculate atomic cluster energies and perform the local optimization of their structures, we have been able to demonstrate that it is possible to perform global optimizations of the structure of atomic clusters using *ab initio* methods. The results show that this approach is able to find many structures that were not previously reported in the literature. Moreover, in most cases the new structures have considerable lower energies than those previously known. The results clearly demonstrate that these calculations are now possible and in spite of their larger computational demands provide more reliable results.

DOI: 10.1103/PhysRevA.72.053205 PACS number(s): 36.40.−c

I. INTRODUCTION

The study of the structure and physical properties of atomic and molecular clusters is an extremely active area of research due to their importance in both fundamental science and applied technology. Existing experimental methods for structural determination seldom can obtain the structure of atomic clusters directly. Therefore the calculation, using theoretical structures and comparison with experimental values of their physical and optical properties, is the most common way to obtain structural information of atomic clusters.

During the last two decades a number of researchers have characterized amorphous hydrogenated silicon (α-Si:H) by different techniques and determined the physical and chemical properties of amorphous silicon films [1,2]. The role of defects in the properties of amorphous and microcrystalline silicon (α-Si and μc-Si, respectively) is very important because these materials are employed in solar cells and optoelectronic devices [3]. As the hydrogen diffuses and modifies the structure, the energy spectrum and the properties of the devices change. To understand these phenomena it is important to understand how hydrogen affects the local electronic structure and geometry in these systems. Since systems like amorphous and microcrystalline silicon are very difficult to handle computationally, understanding about the effects due to hydrogen addition may be achieved by calculations on small hydrogenated silicon clusters. These calculations critically depend on the available structures of such clusters. Prasad and co-workers [4–10] and Yang *et al.* [11] and references therein constitute a very complete bibliographic record of the state of the art in the determination of the structures of hydrogenated silicon clusters. Due to computational limitations, the authors either limited themselves to calculations using local minimizations when using density functional theory (DFT) methods for the energy calculations or used approximated methods to calculate the cluster energies when using global search techniques. Unfortunately, this has been the state of the art in the determination of atomic cluster structures [12–14], computational limitations have forced researchers to use either local optimizations or approximate energy calculations. Our recent work in Si_n clusters [15] shows the need for global optimizations, while the work on Si_nCu clusters [16] clearly demonstrates the drawbacks of using approximate energy calculations when the energy hypersurface provided by the approximate method does not closely match the one predicted from first principles. In agreement with other authors [12–14], we have concluded that it is of importance to develop the necessary computational capabilities to make possible searches of atomic cluster structures using *ab initio* methods.

In this paper we report the use of a parallel genetic algorithm (MGAC) to predict the structure of medium size hydrogen-silicon clusters, Si_nH, $n=4-10$, using DFT methods. To our knowledge this is the first calculation of this type reported in the literature.

II. METHODS

Genetic algorithms (GA's) are a family of search techniques rooted on the ideas of Darwinian biological evolution. The introductory sections of Refs. [17,18] offer a very detailed description of the progress made in the development of GA strategies to perform atomic cluster structure optimizations. The GA methods are based in the principle of survival of the fittest, considering that each string or *genome* represents a trial solution candidate of the problem and that at any generation the *genomes* or "individuals" compete with each other in the population for survival and produce offspring for the next generation by prescribed propagation rules. Operators analogous to crossover, mutation, and natural selection are employed to perform a search able to explore and learn the multidimensional parameter space and determine which

*Corresponding author. Electronic address: Julio.Facelli@utah.edu

 ©2005 The American Physical Society

regions of that space provide good solutions to the problem. One of the advantages of genetic algorithms is that they can provide not only global minima, but also information on other states with energies close to the minimum, an important property when analyzing atomic clusters.

Genetic coding of the cluster structures. In this case the genome is quite simple because there is no symmetry or periodicity constrains. In all our work we have followed previous precedence [19,20] and use a physical space genomic description of the clusters. The genome for a cluster is therefore given by the coordinates of the atoms in the cluster. The genome has a dimension of $3N$, where N is the number of atoms in the cluster.

GA optimization. Initially a population of N_{pop} random individuals is generated. After that the GA operations of mating (crossover), mutation, and selection are used to evolve one generation into the next. The population replacement is made by employing the steady-state genetic algorithm (SSGA), which typically replaces only a proportion of individuals in each generation [21–23]. This technique is also known as elitism. Among the population, the best $N_{\mathrm{pop}}/2$ individuals—that is, 50% of the population—are copied directly into the next generation and the remaining individuals are generated by the GA operators as described below. The criteria for fitness probability, selection of the individuals, and mutation used here are those discussed in detail in Ref. [15]. As with any stochastic minimization procedure the GA should be run several times to guarantee that the resulting structures are independent of the initial population and statistically significant.

Techniques used to improve mutations and crossovers. Connected clusters are arrangements of atoms, which are located at distances such there are significant interatomic interactions among them. Therefore not every possible spatial distribution of atoms makes a valid configuration for physically plausible clusters. This can be used to greatly reduce the effective search space for the GA. In order to take advantage of this feature it is necessary to generate an initial population of compliant structures and to use only genetic operators that guarantee that the new individuals produced have physically plausible structures. In the following we describe the implementation of the genetic operators and initialization techniques used in our current implementation of MGAC.

Atom connectivity. We will consider that two atoms are connected if they are at a distance $[(d_1+d_2)/2]$ where d_1 and d_2 are characteristic interaction distances for each type atom. These distances are selected as proportional to their respective van der Walls radius. In the process of building an atomic cluster we will require that the coordinates of atom A_1 satisfy

$$\|\mathbf{r}_1-\mathbf{r}_2\|^2=\left(\frac{d_1+d_2}{2}\right)^2 \tag{1}$$

for at least one other atom A_2, where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the positions of the atoms. In this way we guarantee that the atoms in the cluster are connected and do not overlap each other. Note that this does not introduce any artificial constrains on the cluster's geometries, because all the structures of the clusters are locally optimized using a quantum mechanical method prior to inclusion in the population.

Generation of the initial population. The initial population is generated randomly. Starting with an atom at the center of coordinates, the second atom is placed along the z direction at a distance such that complies with the condition given by Eq. (1). To continue adding atoms to the cluster the initial fragment is randomly rotated and a new atom is placed along the z direction such that its distance to at least one other atom in the cluster follows Eq. (1). This procedure is continued until the desired number of atoms is achieved for the cluster. Before including the individual in the population, this construction is followed by a local optimization using the desired quantum mechanical method. This process is used as many times as necessary to generate the desired number of individuals in the initial population.

Genetic operators. In the current MGAC implementation we use a real-space representation of the genome defining each individual—i.e., atomic cluster. Therefore, the genetic operators also have to be defined such they operate in real space. The use of this approach for the determination of atomic cluster structures has been previously described [19,20], and many of these GA operators have been proposed. Unfortunately, none of them guarantees that the offspring generated by the operators are compact clusters without disconnected and/or overlapping atoms. Note that the introduction of these unphysical clusters in the population creates a great deal of problems because quite often when using them as starting structures for the local optimizations either take unordinary time to converge to plausible structures or diverge producing abnormal termination of the DFT or alternative code used for the local energy minimizations.

Recombination operators for homogeneous systems. Two well-known recombination operators were redefined to produce connected clusters. Those are based in the "cut and splice" and "two cuts and splice" operators defined by Roberts *et al.* [18]. In the case of the "cut and splice" operator the two individuals that are selected for mating—i.e., clusters A and B—are centered and rotated randomly. The atoms in the two clusters are ordered according to their position in the z direction. The index of each atom not only defines the ordering of atoms but also can be used to define an xy plane at which the atom belongs. Therefore randomly selecting an index between 1 and N, the number of atoms in the cluster, is equivalent to selecting a cut plane for the recombination of the clusters. Following the selection of the cutting plane given by m_{pos}, the atoms belonging to cluster A with indices in the interval $[1,m_{\mathrm{pos}}]$ are assigned to a fragment A_1 and those with indices $(m_{\mathrm{pos}},N]$ are assigned to a fragment A_2. In a similar manner the individuals belonging to B are assigned to the B_1 and B_2 fragments. The mating proceeds by jointing the fragments A_1 with B_2 and A_2 with B_1 to form the two new individuals under the condition that the remaining clusters are connected. This condition is accomplish by moving the two fragments along the direction defined by their center of mass until the positions of all the atoms comply with Eq. (1).

The same procedure follows for "two cuts and splice" operators where in this case is not 1 but 2 the number of cutting planes that are chosen. Therefore for each mating

cluster A and B the fragments A_1, A_2, A_3 and B_1, B_2, B_3 are generated, respectively. The mating process is done by jointing the fragments A_1 with B_2 and A_3 for the first cluster and B_1 with A_2 and B_3 for the second cluster under the condition that the remaining clusters be connected.

In both cases, the resulting clusters are locally optimized and incorporated into the new population.

Recombination operators for heterogeneous systems. In heterogeneous systems not all the possible planes for cutting are available because the selected fractions do not conserve the number of atoms of each species that define the composition of the cluster. This is because the number of atoms of each species above and below the cutting plane is different. This problem can be resolved by checking for every index if the cutting plane that they define preserves the composition of the subclusters above and below the plane for each cluster. If the composition is preserved, the index is stored and becomes eligible for the random selection process used to determine the cutting plane. Once the cutting planes have been selected, the mating process proceeds as explained above for the homogeneous clusters.

Mutation operators. The operators "atom replacement," "twisting," and "permutation" mutation were implemented. For the cases of "atom replacement" and "permutation" the implementation was done following the procedure presented in [19]. In the case of "twisting," the individual selected for mutation is centered, randomly rotated and its atoms are indexed as described above according to their position along the z direction. A randomly chosen index m_{pos} is used to assign the atoms with indices $(m_{\mathrm{pos}},N]$ to A' and those to the interval $[1,m_{\mathrm{pos}}]$ to the fragment A''. The two fragments are independently and randomly rotated and thereafter connected to produce a new structure that after local minimization is incorporated into the population. The use of the "atom replacement" operator can create structures that are slightly not connected, but its use in general is appropriated for keeping a high diversity on the population avoiding the premature convergence of the algorithm.

The new operators introduce here are essentially similar to those that had been previously used in the literature [19,20], but with the incorporation of the notion of connectivity between atoms to restrict the search to connected clusters and therefore speeding up the process of local optimization. We have not tested the effectiveness of the new operators, except for the fact that they were able to produce converged populations in a reasonable number of generations.

Density functional theory. The use of DFT to predict the structure of atomic clusters using different exchange correlation functionals is well documented [14,24–30], and it is not necessary to discuss its merits. It is universally accepted that DFT is the preferred method for atomic cluster structure calculations and that other less computationally demanding approximations are used due to the lack of the computing capabilities needed for DFT methods [14]. In this work we used for the local optimizations the DFT methods as they have been implemented in the wave plane code CPMD (Carr Parinello molecular dynamics) [31]. Within the CPMD code we tested different exchange correlation functionals and pseudopotentials, comparing the results with those from highly correlated calculations (see results section). Based on the tests we decided to use in the local CPMD optimizations the (Becke's exchange correlation and Lee, Yang and Parr exchange functional) [32,33] exchange correlation functional and the Goedecker *et al.* [35] pseudopotential [34,35]. The calculations for the clusters with four to seven Si atoms were done using an energy cutoff of 32 Ry, while the calculations for the larger clusters where done using a cutoff of 16 Ry. In the results section we show that the cutoff of 16 Ry is enough to obtain reliable results while leading to important savings of computational resources with respect to the 32 -Ry cutoff. To avoid the edge effects due to the finite size of the cell introduced by the artificial periodic boundary conditions used by the CPMD code, the cell length was defined as corresponding to the largest dimension of the cluster plus 8 Å on each side.

Computational aspects. The MGAC package has been written in C++, MPI, and using GALib [36] for the GA implementation; this makes it very portable as well as easy to maintain and upgrade.

The evaluation of the objective function for our GA—i.e., the local energy minimization using the DFT method—requires the largest amount of computer time and it is the limiting factor on the problem size that can be solved using the MGAC method. To alleviate this problem we have used a global parallelization scheme of the genetic algorithm. Such parallelization schemes rely on the simultaneous evaluation of the fitness of the individuals belonging to the same population in every generation using our adaptive parallel genetic algorithm (APGA) [37] strategy. The APGA was designed to perform efficiently on heterogeneous cluster of computers and provides a great degree of adaptability and performance in distributed systems. However, this level of parallelization only allows making the execution time approximately independent of population size. In the ideal case when the number of processors is equal to the number of individuals that are evaluated per generation, the computer time per generation is equal to the evaluation time for one individual. Therefore, the achievable best performance when using this strategy is given by the time needed for the local energy minimization of one cluster using the DFT method. This time rapidly increases with the size of the atomic cluster. To further reduce the total execution time it is necessary to execute the local energy minimization in parallel. This was achieved by using a second level of parallelization, where the evaluation of each individual is done in parallel too. This was accomplished using the parallel capabilities of the CPMD code [31]. Table I shows the running and timing parameters from the calculation of the structures of all the isomers studied here with the MGAC-CPMD approach. It is apparent from Table I that our method shows great scalability properties and it can take advantage of modern parallel computer systems.

III. RESULTS AND DISCUSSION

CPMD validation process. To validate the suitability of using the different combinations of exchange correlation functionals and pseudopotentials available in the CPMD

TABLE I. Computational and timing parameters used in the calculation of the structures of Si_nH atomic clusters using the MGAC method for global search and the CPMD method for the local optimization and energy calculations. Note that calculations for $n=4-7$ were done using an energy cutoff of 32 Ry, while for $n=8-10$ an energy cutoff of 16 Ry was used.

	Individuals in the population	Number of computer processors			Elapsed time per generation[a]	Number of generations
		Level 1	Level 2	Total		
Si_4H	10	4	10	40	1.5	10
Si_5H	10	5	15	75	2	15
Si_6H	15	5	20	100	2	15
Si_7H	15	7	25	175	2	20
Si_8H	20	5	25	125	2	30
Si_9H	30	5	35	175	1.5	55
$Si_{10}H$	30	5	43	215	1.5	50

[a]The elapsed time is given in hours. The total execution time of the search is the time per generation times the number of generations needed to achieve convergence.

program we have used a series of high level *ab initio* calculations in H, H_2, Si, Si_2, SiH (silylidyne), SiH_2 (silicon dihydride), SiH_3 (silyl), and SiH_4 (silane) as benchmarks [38–42]. For each system we calculated the optimized geometry, the binding energy, and the vibrational frequencies for all the valid combinations of exchange correlation functional and pseudo potential available in the CPMD code.

The results for all those systems, when using energy cutoff of 32 Ry and 16 Ry, are summarized in Table II. It is apparent that the quality of the results, as determined by their rms with respect of the experimental values, depends mainly on the accuracy of the pseudopotential (PP) used. Methods with improved functionals cannot compensate for PP deficiencies. It is observed that the BLYP-GO combination gives one of the best performances for all four geometrical characteristics of the systems. For bond distances BLYP-GO has a rms of 0.025 Å where the best performance was for MP2FC/6-311G* with a rms of 0.008 Å. In the cases of bond angles, BLYP-GO has a rms of 0.58° where the best performance was for MP4/6-311G* with a rms of 0.56°. For binding energy, BLYP-GO has a rms of 0.088 eV where the best performance was for BLYP/6-311G* with a rms of 0.045 eV. Finally for normal modes frequencies, BLYP-GO has a rms of 59 cm^{-1} where the best performance was for

TABLE II. rms between experimental and calculated geometry parameters, binding energy, and vibrational frequencies for selected Si_nH_m systems with $E_{cut}=32$ Ry and 16 Ry.

Method	rms (Re) (Å)	Δangle (deg)	rms (Eb) [eV]	rms (Freq) [cm^{-1}]	rms (Re) (Å)	rms (angle) [deg]	rms (Eb) [eV]	rms (Freq) [cm^{-1}]
CPMD	Energy cutoff 32 Ry				Energy cutoff 16 Ry			
BLYP/GO	0.025	0.58	0.088	59	0.046	0.53	0.04	102
LDA/GO	0.030	1.23	0.298	92	0.052	1.50	0.22	162
BLYP/MT	0.134	1.46	0.340	270	0.147	1.52	0.54	313
BP/MT	0.122	1.78	0.281	245	0.136	1.78	0.25	294
HCTH/MT	0.129	1.79	0.573	263	0.140	1.72	0.54	283
LDA/MT	0.111	2.34	0.159	248	0.125	2.48	0.31	301
OLYP/MT	0.116	1.70	0.287	223	0.129	1.66	0.34	267
PBE/MT	0.120	1.89	0.338	239	0.133	1.90	3.82	287
revPBE /MT	0.123	1.70	0.363	238	0.135	1.77	3.85	286
All electron calculations								
BLYP/6-311G*	0.038	0.88	0.045	38				
B3LYP/6-311G*	0.020	0.89	0.062	35				
MP2FC/6-311G*	0.008	0.67	0.382	80				
MP4/6-311G*	0.030	0.56	0.319	41				
QCISD(T)/6-311G*	0.032	0.57	0.311	38				
CCSD(T)/6-311G*	0.033	0.57	0.306	41				

TABLE III. Calculated binding energies, average Si-Si bond lengths, Si-H bond lengths, and α- and β-energy gaps for the Si_nH clusters found by the MGAC-CPMD method. Their structures are given in Fig. 1. The italicized entries correspond to those found by the MGAC-CPMD method that were not previously reported in the literature [4–11].

		Bond lengths (Å)			
Isomer	E_b (eV)	Si-Si	Si-H	α gap (eV)	β gap (eV)
Si_4H-*a*[a,b]	2.742	2.344	1.511	1.674	1.971
Si_5H-*a*[b]	2.901	2.389	1.491	1.197	2.126
Si_5H-*b*[a,b]	2.869	2.425	1.501	0.672	2.701
Si_5H-*c*[a,b]	2.847	2.497	1.730	1.038	2.186
Si_6H-*a*[b]	3.006	2.480	1.506	0.953	2.212
Si_6H-*b*[a]	2.982	2.416	1.502	1.344	1.543
Si_6H-c	*2.962*	*2.484*	*1.498*	*1.567*	*1.085*
Si_7H-a	*3.032*	*2.522*	*1.425*	*1.602*	*1.611*
Si_7H-*b*[a,b]	3.024	2.513	1.507	1.071	2.554
Si_7H-c	*2.994*	*2.518*	*1.519*	*1.818*	*1.075*
Si_7H-d	*2.991*	*2.467*	*1.499*	*1.584*	*1.524*
Si_7H-e	*2.979*	*2.475*	*1.503*	*0.814*	*1.398*
Si_8H-*a* [b]	3.056	2.472	1.529	0.672	1.687
Si_8H-*b* [a]	3.050	2.525	1.531	1.192	1.171
Si_8H-c	*3.044*	*2.514*	*1.537*	*0.832*	*1.483*
Si_8H-d	*3.042*	*2.517*	*1.530*	*1.346*	*1.122*
Si_8H-e	*3.042*	*2.517*	*1.530*	*1.341*	*1.124*
Si_9H-*a*[b]	3.100	2.488	1.530	0.970	1.360
Si_9H-b	*3.099*	*2.496*	*1.529*	*0.973*	*1.367*
Si_9H-c	*3.099*	*2.508*	*1.528*	*1.289*	*1.308*
Si_9H-d	*3.098*	*2.495*	*1.526*	*1.117*	*1.155*
Si_9H-e	*3.098*	*2.501*	*1.528*	*1.085*	*1.446*
$Si_{10}H$-*a*[a]	3.173	2.515	1.533	1.408	1.366
$Si_{10}H$-b	*3.172*	*2.486*	*1.524*	*1.165*	*1.466*
$Si_{10}H$-c	*3.171*	*2.504*	*1.522*	*1.257*	*1.455*
$Si_{10}H$-d	*3.166*	*2.506*	*1.517*	*1.397*	*0.979*

[a]Reported by Prasad and co-workers *et al.* [4–10].
[b]Reported by Yang *et al.* [11].

B3LYP/6-311G* with a rms of 35 cm^{-1}. Similar results are obtained when using a smaller cutoff of 16 Ry; again, the best performance for all geometrical characteristics of the systems is the BLYP-GO combination. While the rms values for the smaller cutoff are not as good as those obtained using 32 Ry, they are quite comparable to those obtained using accurate highly correlated methods such as MP2, MP4, QCISD, and CCSD. Following these observations the calculations reported here for the Si_nH clusters were performed using the BLYP-GO. For clusters with $n=4-7$ we used a cutoff of 32 Ry, while for the larger ones $n=8-10$ a cutoff of 16 Ry was used.

Structures of Si_nH clusters. In Table III we present the calculated binding energies, average Si-Si bond lengths, Si-H bond lengths, and α- and β-energy gaps for all the unique Si_nH clusters found by the MGAC-CPMD method. The corresponding structures of the clusters are given in Fig. 1 and their Cartesian coordinates have been deposited in AIP's EPAPS [43]. Comparing the structures of the clusters in Fig. 1 with the ones reported by Prasad and co-workers [4–10] and Yang *et al.* [11] it is apparent that the MGAC-CPMD approach is able to find a number of structures that were not found previously. The entries corresponding to these structures have been italicized in Table III. In many cases the new structures correspond to lower-energy structures than those from the literature. The inspection of the calculated vibrational frequencies of all the structures in Fig. 1, deposited as supplementary information in AIP's EPAPS [43], revels that all the calculated frequencies are real, indicating that the structures are true minima in the energy hypersurface and not meta stable configurations.

Most of the structures show one frequency around 2200 cm^{-1} corresponding to the Si-H stretching vibration mode. The experimental value is 2042 cm^{-1} [42] for SiH. This stretching vibrational mode does not appear in Si_5H-*c*

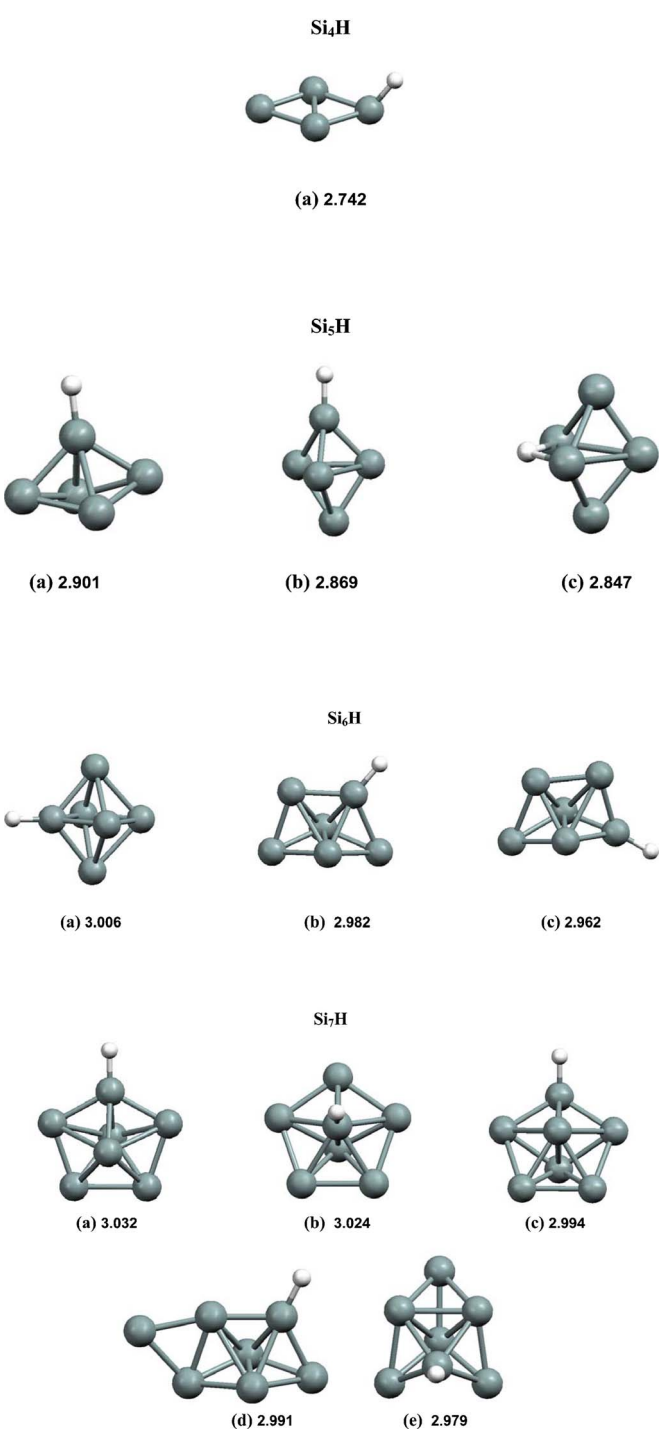

FIG. 1. (Color online). Best isomers found by MGAC-CPMD calculations. The binding energy given under the schema are in eV.

because this isomer shows the hydrogen bounded to two silicon atoms.

The Si_4H structure corresponds to four silicon atoms forming a flat rhombus and the hydrogen bonded to a silicon atom. From the comparison of the Si-Si bond lengths with those found in Ref. [13] for Si_4 we see that they are reduced in 0.08 Å due to the addition of the hydrogen. The MGAC-CPMD structure for Si_4H is in agreement with all previous reports of the stable structure for this cluster.

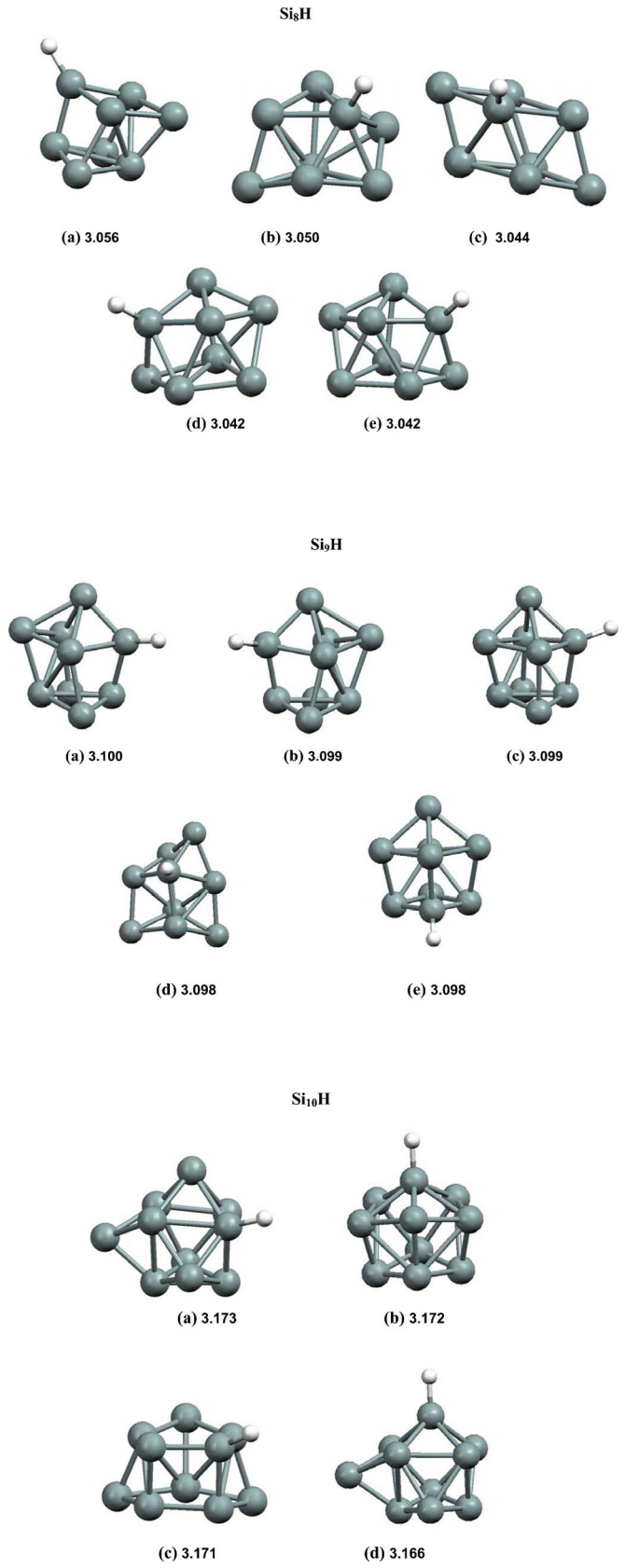

FIG. 1. (*Continued*).

For Si_5H we found three different stable isomers displayed in Fig. 1: Si_5H-*a* matches the new structure reported by Yang *et al.* [11] employing full electron DFT calculations. The four nonhydrogen silicon bonded atoms define a dihedral of 142°. The Si-Si average bond is 2.389 Å and the Si-H bond is 1.491 Å. Si_5H-*b* and Si_5H-*c* correspond to those reported by Balamurugan and Prasad [10], but in our case structure Si_5H-*c*, which presents a hydrogen atom bonded to two silicon atoms, is not the most stable in agreement with Yang *et al.* [11]. The Si-Si distances coincide with those previously reported in [13] for Si_5, but the bond lengths are slightly different, they change in about 0.06 Å and the angles (bonding and dihedral) change by approximately 5°, because of the presence of hydrogen in the cluster.

For Si_6H-*a* the Si-H bond length is 1.506 Å, the hydrogen and four silicon atoms are in a plane, and the other two silicon atoms are above and below it, lying in an orthogonal plane formed by four silicon atoms. The Si-Si average bond is 2.480 Å. The corresponding Si_6 isomer [13] shows an average bond length of 2.434 Å and a dihedral angle of 24° instead of a flat four silicon plane. Si_6H-*b* coincides with the isomer reported by Balamurugan and Prasad [10]. In a near energy range, higher by 0.02 eV, we found that isomer Si_6H-*c* differs from the Si_6H-*b* one in the silicon site to which the hydrogen is bonded.

Si_7H-*a* and Si_7H-*c* are new isomers, with structures very similar to Si_7H-*b*, reported by Balamurugan and Prasad [10]. A five-silicon ring characterizes them, and they differ in the silicon site bonded to the hydrogen atom. Si_7H-*d* and Si_7H-*e* are also new isomers reported here. For isomers (a)–(c) and (e), the silicon structure was previously found in Ref. [13], but the average Si-Si distance have been reduced by 0.03 Å.

We found five stable configurations for Si_8H: isomer Si_8H-*a* was reported by Yang *et al.* [11], isomer Si_8H-*b* was reported by Balamurugan and Prasad [10], and isomers Si_8H-*c*, Si_8H-*d*, and Si_8H-*e* are new structures produced by MGAC-CPMD. The Si-H bond length is 1.529 Å for isomer (a), 1.531 Å for (b), 1.537 Å for (c), and 1.530 Å for (d) and (e), which are mirror images of the same isomer. The corresponding Si-Si average bond lengths are 2.472 Å, 2.525 Å, 2.514 Å, and 2.517 Å, respectively. Comparing these structures with the Si_8 [13] we observe that the addition of hydrogen increases the symmetry of these isomers.

We found five isomers for Si_9H; structures Si_9H-*a* and Si_9H-*b* are mirror structures of the same isomer. Yang *et al.* also reported the ground state Si_9H-*a* [11]. The average Si-Si bond lengths are of the order of 2.50 Å, and Si-H runs between 1.526 Å and 1.529 Å. From the comparison of these structures with those of Si_9 in Ref. [13] we conclude that the Si-Si bond length is shortened in $\approx$0.05 Å and all the Si_9H isomers might be associated to one of those Si_9.

We found four stable isomers for $Si_{10}H$; the ground state corresponds to that reported by Balamurugan and Prasad [10]; the Si-Si-average bond length is 2.515 Å for isomer $Si_{10}H$-*a*, 2.486 Å for isomer $Si_{10}H$-*b*, and 2.505 Å for $Si_{10}H$-*c* and $Si_{10}H$-*d*. The corresponding Si-H bond lengths are 1.533 Å, 1.524 Å, 1.522 Å, and 1.517 Å, respectively. From the comparison of these isomers with those reported in Ref. [13] we observe that for $Si_{10}H$-*a*, $Si_{10}H$-*c*, and $Si_{10}H$-*d* there is a Si_{11} for which one silicon atom seems replaced by an hydrogen on these $Si_{10}H$ compounds.

Our results show good agreement with previous searches of the configurations of small SiH clusters, but consistently we found increasing numbers of new isomers for the larger ones. This is a clear indication that global searches become increasingly important for larger clusters. On the question of the Si-H binding we found that for the most stable isomer the hydrogen is always bonded to only one Si atom and it is on the outside of the cluster. This result follows the electronegativity argument recently provided by Yang *et al.* [11].

IV. CONCLUSIONS

This paper clearly demonstrates that it is possible, using current computational resources, to perform global searches of the structures of medium size atomic clusters using first-principles (DFT) methods. The results clearly show that this procedure is more reliable than previous approximate methods based on either approximate energy calculations and/or more constrained searches. The ability of the MGAC-CPMD method to find better isomers than those previously found is a clear justification for its use in spite of the additional computational resource needed for the calculations. Note the importance of performing global searches greatly increases with the size of the cluster. Moreover, the architecture of the MGAC-CPMD software allows for taking advantage of large computer clusters making these large calculations feasible when using parallel computer systems available at major research centers. The comparison of our DFT results with those for which either experimental data or high-level calculations exists shows that the quality of the results depends mainly on the accuracy of the pseudopotential used. Methods with improved functionals cannot compensate for PP deficiencies. It is observed that the BLYP-GO combination gives one of the best performances for the Si_nH clusters studied here.

ACKNOWLEDGMENTS

Financial support for this research by the University of Buenos Aires (Grant No. UBACYT-X035) and the Argentinean CONICET (Grant No. PIP5119) is gratefully acknowledged. The software for this work used the GAlib genetic algorithm package, written by Matthew Wall at the Massachusetts Institute of Technology. The Center for High Performance Computing provided computer resources in the Arches cluster for this project. The Arches cluster was partially funded by NIH-National Center for Research Resource (Grant No. 1S10RR017214-01).

[1] H. Fritzsche, Annu. Rev. Mater. Sci. **31**, 47 (2001).
[2] R. O. Jones, B. W. Clare, and P. J. Jennings, Phys. Rev. B **64**, 125203 (2001).
[3] A. Hiraki *et al.*, Sol. Energy Mater. **2**, 125 (1979).
[4] N. Chakraborti, P. S. De, and R. Prasad, Z. Metallkd. **90**, 508 (1999).
[5] G. R. Gupte and R. Prasad, Int. J. Mod. Phys. B **12**, 1607 (1998).
[6] G. R. Gupte and R. Prasad, Int. J. Mod. Phys. B **12**, 1737 (1998).
[7] N. Chakraborti, P. S. De, and R. Prasad, Mater. Lett. **55**, 20 (2002).
[8] D. Balamurugan and R. Prasad, Bull. Mater. Sci. **26**, 123 (2003).
[9] R. Prasad, Bull. Mater. Sci. **26**, 117 (2003).
[10] D. Balamurugan and R. Prasad, Phys. Rev. B **64**, 205406 (2001).
[11] J. C. Yang *et al.*, J. Phys. Chem. A **109**, 5717 (2005).
[12] J. Zhao and R.-H. Xie, J. Comput. Theor. Nanosci. **1**, 117 (2004).
[13] W. Ekardt, *Metal Clusters*, Wiley Series in Theoretical Chemistry (Wiley, Chichester, 1999).
[14] P. Ballone and W. Andreoni, in *Metal Clusters*, edited by W. Ekardt (Wiley, Chichester, 1999), p. 71.
[15] V. E. Bazterra *et al.*, Phys. Rev. A **69**, 053202 (2004).
[16] O. Oña *et al.*, J. Mol. Struct.: THEOCHEM **681**, 149 (2004).
[17] M. Iwamatsu, J. Chem. Phys. **112**, 10976 (2000).
[18] C. Roberts, R. L. Johnston, and N. T. Wilson, Theor. Chem. Acc. **104**, 123 (2000).
[19] H. M. Cartwright and L. M. Sztandera, *Soft Computing Approaches in Chemistry*, (Springer-Verlag, Heidelberg, 2003).
[20] R. L. Johnson and C. Robers, in *Soft Computing Approaches in Chemistry*, edited by H. M. Cartwright and L. M. Sztandera (Springer-Verlag, Heidelberg, 2003), Vol. 120, p. 161.
[21] G. Syswerda, in *Proceedings of the Third International Conference on Genetic Algorithms, Fairfax, VA, 1989*, edited by J. David Schaffer, (Morgan Kaufman Publishers Inc., San Francisco 1989), pp. 2–9.
[22] D. Whitley and J. Kauth, in Proceeding of the Rocky Mountain Conference on Artificial Intelligence, Denver, Colorado, 1998, pp. 118–130.
[23] D. Whitley in *Proceedings of the Third International Conference on Genetic Algorithms, Fairfax, VA, 1989*, edited by J. David Schaffer, (Morgan Kaufman Publishers Inc., San Francisco 1989) pp. 116–121.
[24] Y. Xiao and D. E. Williams, Chem. Phys. Lett. **215**, 17 (1993).
[25] C. Y. Xiao *et al.*, Int. J. Quantum Chem. **94**, 416 (2004).
[26] V. V. Ovcharenko *et al.*, J. Chem. Phys. **114**, 9028 (2001).
[27] C. Xiao *et al.*, J. Phys. Chem. A **106**, 11380 (2002).
[28] C. Xiao, F. Hagelberg, and W. A. Lester, Phys. Rev. B **66**, 075425 (2002).
[29] C. Xiao *et al.*, J. Mol. Struct.: THEOCHEM **549**, 181 (2001).
[30] C. Xiao and F. Hagelberg, J. Mol. Struct.: THEOCHEM **529**, 241 (2000).
[31] M. Parrinello and W. Andreoni, computer code CPMD V3.9, http://www.cpmd.org/.
[32] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[33] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[34] C. Hartwigsen, S. Goedecker, and J. Hutter, Phys. Rev. B **58**, 3641 (1998).
[35] S. Goedecker, M. Teter, and J. Hutter, Phys. Rev. B **54**, 1703 (1996).
[36] M. Wall, GAlib library, http://lancet.mit.edu/ga/.
[37] V. E. Bazterra *et al.*, J. Parallel Distrib. Comput. **65**, 48 (2005).
[38] K. P. Huber and G. Herzberg, *Constants of Diatomic Molecules*, Vol. IV of *Molecular Spectra and Molecular Structure* (Van Nostrand Reinhold, New York, 1979).
[39] J. D. Cox, D. D. Wagman, and V. A. Medvedev, *CODATA Key Values for Thermodynamics* (Hemisphere, New York, 1989).
[40] L. V. Gurvich, I. V. Veyts, and C. B. Alcock, *Thermodynamic Properties of Individual Substances* (Hemisphere, New York, 1989).
[41] M. E. Jacox, "Vibrational and Electronic Energy Levels of Polyatomic Transient Molecules," J. Phys. Chem. Ref. Data Monogr. 3 (1994).
[42] NIST Standard Reference Database, Computational Chemistry Comparison and Benchmark DataBase, 2004.
[43] See EPAPS Document No. PLRAAN-72-139510 for the Cartessian coordinates of all the clusters studied here and their calculated vibrational frequencies. This document can be reached via a direct link in the online article's HTML reference section or via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html).