Bull. Mater. Sci., Vol. 26, No. 1, January 2003, pp. 127–130. © Indian Academy of Sciences.

# Application of genetic algorithms to hydrogenated silicon clusters

**N CHAKRABORTI[†] and R PRASAD***

Department of Physics, Indian Institute of Technology, Kanpur 208 016, India
[†]Department of Metallurgical and Materials Engineering, Indian Institute of Technology, Kharagpur 721 302, India

**Abstract. We discuss the application of biologically inspired genetic algorithms to determine the ground state structures of a number of Si–H clusters. The total energy of a given configuration of a cluster has been obtained by using a non-orthogonal tight-binding model and the energy minimization has been carried out by using genetic algorithms and their recent variant differential evolution. Our results for ground state structures and cohesive energies for Si–H clusters are in good agreement with the earlier work conducted using the simulated annealing technique. We find that the results obtained by genetic algorithms turn out to be comparable and often better than the results obtained by the simulated annealing technique.**

**Keywords. Hydrogenated silicon; genetic algorithms; differential evolution; *ab initio* calculation.**

## 1. Introduction

In some of our recent articles (Gupte and Prasad 1998a, b; Prasad and Chakraborti 1998; Chakraborti *et al* 1999; 2001a, 2002), we have reported successful ground state calculations for a number of Si–H clusters, by coupling the biologically inspired genetic algorithms (Mitchell 1998) with a tight-binding formulation (Ohno *et al* 1999). Since hydrogenated silicon is an important opto-electronic material, such studies are expected to have a significant impact in the field of computational material design. Our usage of the tight-binding methodology has drastically reduced the computing time as compared to the first principle *ab initio* technique (Car and Parrinello 1985; Balamurugan and Prasad 2001). The results obtained by tight binding method, on the other hand, are quite comparable to what we have recently obtained through Car–Parrinello molecular dynamics (Car and Parrinello 1985; Balamurugan and Prasad 2001). In our earlier work a non-orthogonal tight-binding formulation was coupled with a simulated annealing strategy (Gupte and Prasad 1998a, b). Subsequent adaptation of genetic algorithms has rendered the search for the ground state energy minimum, a far more efficient process. A brief overview of the methodology along with some of the salient features of the results obtained is presented in this article.

## 2. Genetic fundamentals

Genetic algorithms are state-of-the-art computing techniques, which tend to mimic the Darwinian principle of survival of the fittest, often in a context far fetched from biology of any kind. Being a highly *robust* optimizer, they are now being vigorously applied in a wide variety of problems, some of which are of extremely complex and computing intensive (Wang and Ho 1997). In the most common form of genetic algorithms, viz. the simple genetic algorithms (SGA), the variables are mapped on to corresponding binaries, generally using a linear mapping formula (Chakraborti *et al* 2001b). The binary variables are juxtaposed to form an individual, while a collection of individuals forms a population. Each member of the population, therefore, carries a potential solution of the problem and the initial population is generated randomly. A number of genetic operators, crossover and mutation for example, then act on the population, roughly emulating the similar processes in the natural world. The better individuals for the next generation are selected based upon their respective fitness values—a parameter that quantifies the proximity of the individual to the actual solution. In the context of an optimization problem, a pseudo code for genetic algorithms can be given as: initialize random binary population; do {calculation of fitness; reproduction; crossover; mutation} while (termination criterion not satisfied).

The major difference between genetic algorithms and most of the traditional search routines lies in the implementation of the population concept. In contrast to the traditional methods where only one updated solution is available at any particular iteration, genetic algorithms rely upon a number of possible solutions, each being a member of the population at a particular generation—the term synonymous to iteration in the classical sense. For a successful application of genetic algorithms, the diversity among the population members need to be ensured till the solution converges. Unless there is enough diversity in the gene pool, the relatively strong individuals tend to make increasing number of copies of themselves in the

*Author for correspondence

forthcoming generations and ultimately the stagnant population leads to a premature convergence. A good genetic algorithm generally uses a safeguard against it. A fitness-sharing concept was used in our earlier work on Si–H clusters (Chakraborti *et al* 2002) where each individual shared its cohesive energy values with the population members in the neighbourhood, measured by a prescribed Euclidean distance. This prevented uninhibited growth of any individual, since the weaker members in the vicinity always lowered its fitness.

Genetic algorithms, as indicated before, require mapping of all the system variables in a binary format. In case of a large number of variables, this leads to processing of large arrays containing 1s and 0s, which often substantially retards the computational speed. In addition, binary arithmetic has an implicit disadvantage called Hamming Cliff problem (Deb and Agrawal 1995), which sometimes adversely affects the performance of SGA. When the search reaches a Hamming Cliff, any small change in the variable values in the real space requires a very large change in the corresponding binary. This decelerates the progress of the search and sometimes, fine convergence becomes impossible in a near optimal scenario. Such problems can be easily averted if, instead of the binary equivalents, the genetic operations can be performed on the real-coded variables themselves. Some of the recent versions of genetic algorithms precisely tend to attempt that. We have successfully applied one such algorithm, the recently proposed differential evolution (DE) (Price and Storn 1997), for the present problem. Like most other genetic algorithms, differential evolution can also be used for optimizing any function with a number of constraints. Being an evolutionary optimization technique, DE works with a population, resorts to natural selection, based upon the fitness of its individuals and creates a new and hopefully improved, generation by doing crossover and mutation. All such operations however, need to be redefined in case of DE, as the variable space is real, not binary. Further details of DE are provided elsewhere (Chakraborti *et al* 2001a).

## 3. Non-orthogonal tight-binding approximation

The tight binding approximation is now becoming increasingly popular for studying covalently bonded materials (Ohno *et al* 1999). It assumes the system to be consisting of ionic cores and electron gas, and attempts to calculate the total energy functional for the entire system ($E^{\text{total}}$) by adding up the one particle eigenvalues and the individual pair potential terms, such that

$$E^{\text{total}} = U_0 + E^{\text{el}} + E^{\text{pair}}, \tag{1}$$

where the constant, $U_0$, shifts the cohesive energy as needed, $E^{\text{el}}$ denotes the energy associated with the occupied eigenvalues of the electronic system and $E^{\text{pair}}$ the sum of pair potential terms arising due to repulsion between the ionic cores.

Denoting the occupancy of the $k$th eigenstate as $g_k$, and $N^{\text{occ}}$ as the number of occupied orbitals, the electronic contribution to the total energy is expressed as

$$E^{\text{el}} = \sum_{k=1}^{N^{\text{occ}}} g_k \varepsilon_k. \tag{2}$$

Furthermore, summing up the pair potential terms related to repulsion between the ionic cores, $E^{\text{pair}}$ is obtained as

$$E^{\text{pair}} = \sum_{i<j} \phi(r_{ij}). \tag{3}$$

Utilizing this basic definition for total energy, the wave functions of these eigenstates are given in terms of non-orthogonal basis as

$$|\psi_n\rangle = \sum_i C_i^n |\phi^i\rangle, \tag{4}$$

where $|\phi^i\rangle$ are the basis functions. In non-orthogonal tight binding theory employed here the basis functions are localized on each atom resembling its atomic orbital and spherical harmonic functions ($Y_{\text{im}}$) are used to describe their angular parts. The characteristic equation is then expressed as

$$\sum_i (H_{ij} - \varepsilon_n S_{ij}) C_i^n = 0, \tag{5}$$

where $H_{ij}$ denotes the Hamiltonian matrix elements between the $i$th and $j$th orbitals, such that

$$H_{ij} = \langle \phi^i \mid H \mid \phi^j \rangle. \tag{6}$$

The overlap matrix elements between them are expressed as

$$S_{ij} = \langle \phi^i \parallel \phi^j \rangle. \tag{7}$$

Further details of calculating the Hamiltonian and the overlap elements are provided elsewhere (Gupte 1998).

## 4. Computational

There are actually infinite atomic arrangements possible in this solution domain and the task of genetic algorithms was to locate the configuration at ground state containing the minimum energy. The atomic coordinates were taken as the genetic variables and a Cartesian coordinate system was used for simplicity. The search for the ground state was conducted in a cubic space of 125 cubic Å. The population size was adjusted on the basis of number of variables and a scheme for adjusting the mutation constant and crossover probability was evolved through a systematic trial and error.

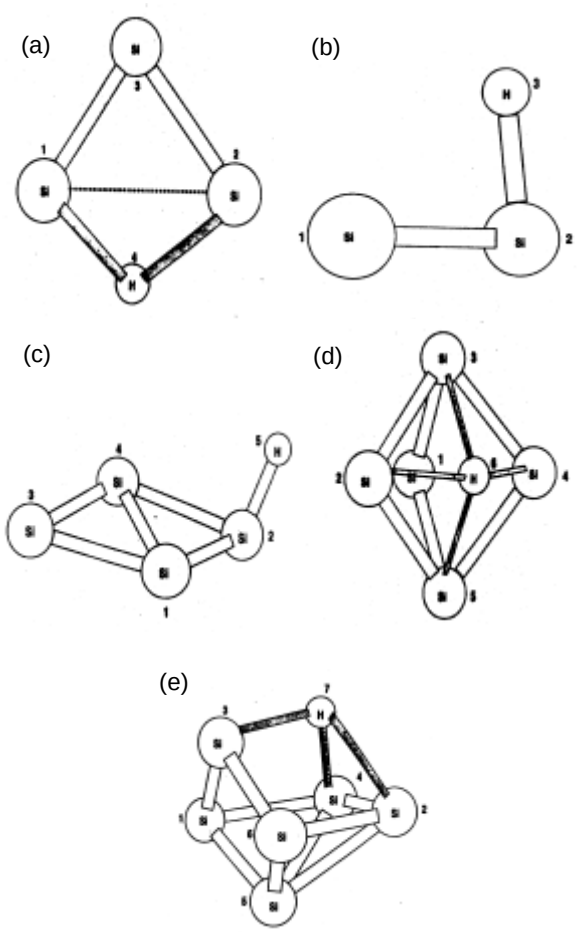


**Figure 1.** Ground state geometry of (**a**) $Si_2H$, (**b**) $Si_3H$, (**c**) $Si_4H$, (**d**) $Si_5H$ cluster and (**e**) $Si_6H$ cluster.

All the calculations were performed in a local area network of a number of silicon graphics workstations of SG 200 origin series.

## 5. Results and discussion

The ground state geometries of $Si_nH$ clusters using the genetic algorithms from $n = 2$ to $n = 6$ are shown in figure 1. The results obtained by simulated annealing (Gupte and Prasad 1998a, b), genetic algorithms (Chakraborti *et al* 1999, 2002) and differential evolution (Chakraborti *et al* 2001a) are in excellent agreement with what has been calculated recently by using Car–Parrinello technique (Balamurugan and Prasad 2001). It is interesting to note that hydrogen is bonded with two silicon atoms in $Si_3H$ and $Si_5H$ although its valence is one. Such Si–H–Si bridge type bonds are thought to be present in amorphous hydrogenated silicon and play an important role in explaining Staebler–Wronski effect.

Further details of each of the clusters are provided below.

*$Si_2H$*: There are two competing structures for $Si_2H$, one is symmetric and the other is asymmetric. Our initial calculations using differential evolution (Chakraborti *et al* 2001a) obtained this structure as planar, asymmetric as shown in figure 1(a). However, subsequent calculations (Balamurugan and Prasad 2001) using Car–Parrinello molecular dynamics gives symmetric structure as ground state.

*$Si_3H$*: In $Si_3H$ structure two Si atoms are situated symmetrically with respect to the third. The lone hydrogen atom is also situated equidistantly from two Si atoms, forming a planar structure shown in figure 1(b). Note that in this case H is bonded to two silicon atoms.

*$Si_4H$*: The ground state structure of this cluster is quite similar to that of $Si_4$ reported earlier (Menon and Subbaswamy 1994). The presence of hydrogen causes some distortion in the structure. The essential geometric features of $Si_4$ are however retained, as shown in figure 1(c).

*$Si_5H$*: In this case three Si atoms are situated on the same plane as the lone hydrogen atom as shown in figure 1(d). The coordinates of hydrogen atom predicted by Car–Parrinello method is slightly different from what has been obtained by GAs and DE, and is shown here for the sake of accuracy. The structure is symmetric along this plane with Si atoms. This structure is quite deviant from the reported structure of $Si_5$ (Menon and Subbaswamy 1994), and in fact it is more like the reported structure of $Si_6$ (Menon and Subbaswamy 1994).

*$Si_6H$*: The ground state structure of $Si_6H$ appears to be a slightly distorted bicapped tetrahedron. The distortion is caused by the presence of hydrogen in the lattice, which shifts the Si atoms in its immediate neighbourhood. The structure is shown in figure 1(f). This structure is quite different from that obtained from the Car–Parrinello method.

This study has suggested that hydrogen can bond with two silicon atoms forming a bridge-like Si–H–Si bond. In case of a closed compact silicon cluster, hydrogen bonds to the cluster from outside. Our results show that the addition of single hydrogen can cause large changes in the electronic structure of a silicon cluster but the geometry is not much affected. Such observations may have some important consequences in the material selection for the opto-electronic devices like light sensors, thin film transistors, light emitting diodes etc where hydrogenated amorphous silicon appears to be an important material (Paesler *et al* 1989).

## 6. Conclusions

During the last decade there has been a very rapid development in the study of small clusters. This has been partly due to the growing importance of these systems in applications like catalysis and due to possibilities of developing nano-electronic devices. Also some clusters are found to be very stable and referred to as magic clusters. There have been attempts to design novel materials using these clusters as basic building blocks. These materials may have very different properties compared to the naturally occurring solids. This study has demonstrated that biologically inspired genetic algorithms and differential evolution can be efficiently used for computing the ground state configurations of the small, hydrogenated silicon clusters. These techniques can be easily tried out for a large number of related problems and their increasing presence in materials research is expected to be seen in a very near future.

## Acknowledgement

Partial financial support provided by the Department of Science and Technology, New Delhi, through grant number SP/S2/M-51/96, is acknowledged.

## References

Balamurugan D and Prasad R 2001 *Phys. Rev.* **B64** 205406
Car R and Parrinello M 1985 *Phys. Rev. Lett.* **55** 2471
Chakraborti N, De P S and Prasad R 1999 *Z. Metalkde* **90** 508
Chakraborti N, Misra K, Bhatt P and Prasad R 2001a *J. Phase Equilibria* **22** 525
Chakraborti N, Kumar R and Jain D 2001b *Appl. Math. Modell.* **25** 287
Chakraborti N, De P S and Prasad R 2002 *Mater. Lett.* **55** 20
Deb K and Agrawal R B 1995 *Complex Systems* **9** 115
Gupte G R 1998 *Molecular dynamics studies of small hydrogenated silicon clusters and hydrogenated amorphous silicon,* Ph D dissertation, Indian Institute of Technology, Kanpur
Gupte G R and Prasad R 1998a *Int. J. Mod. Phys.* **B12** 1607
Gupte G R and Prasad R 1998b *Int. J. Mod. Phys.* **B12** 1737
Menon M and Subbaswamy K R 1994 *Phys. Rev.* **B50** 11577
Mitchell M 1998 *An introduction to genetic algorithms* (Prentice-Hall)
Ohno K, Esfarjani K and Kawazoe Y 1999 *Computational materials science from ab initio to Monte Carlo methods* (New York: Springer-Verlag)
Paesler M, Agarwal S C and Zallen M 1989 *Amorphous and liquid semiconductors* (Amsterdam: North-Holland)
Prasad R and Chakraborti N 1998 *Metals, Mater. & Process.* **10** 203
Price K and Storn R 1997 *Dr Dobbs Journal* **22** 18
Wang C Z and Ho K M 1997 *J. Phase Equilibria* **18** 516