Calculations of Silicon Self-Interstitial Defects

W.-K. Leung, R. J. Needs, and G. Rajagopal
TCM Group, Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, United Kingdom

S. Itoh and S. Ihara
Central Research Laboratory, Hitachi Ltd. (Japan), Kokubunji, Tokyo 185-8601, Japan
(Received 16 November 1998)

We report a theoretical study of self-interstitial defects in silicon using local density approximation (LDA), PW91 generalized gradient approximation (GGA), and fixed-node diffusion quantum Monte Carlo (DMC) methods. The formation energies of the stablest interstitial defects are about 3.3 eV within the LDA, 3.8 eV within the PW91-GGA, and 4.9 eV within DMC. The DMC results indicate a value for the formation + migration energy of the self-interstitial contribution to self-diffusion of about 5 eV, which is consistent with the experimental data. This confirms the importance of a proper treatment of electron correlation when studying such systems.

PACS numbers: 61.72.Bb, 66.30.–h, 71.10.–w

Silicon is the material at the heart of the microelectronics industry. A deeper understanding of this technologically important material is central to the fabrication of future generation semiconductor devices. One of the important problems in the manufacture of submicron devices is the diffusion of dopant impurity atoms during thermal processing, which limits how small they can be made. To understand these effects requires a knowledge of diffusion on the microscopic scale in situations far from equilibrium. The diffusion of impurity atoms in silicon is critically influenced by intrinsic defects such as self-interstitials and vacancies, and therefore it is of great importance to improve our understanding of the behavior of these defects.

Because of the technological importance of diffusion in silicon many experimental and theoretical studies have been carried out, including much work on silicon self-interstitials. Unfortunately, it has not been possible to detect the self-interstitials directly, although their presence has been inferred from various measurements [1]. Measurements of the self-diffusion constant or self-diffusivity of silicon at high temperatures, using radioactive isotopes of silicon as tracers, have established an Arrhenius behavior with an activation energy in the range 4.1–5.1 eV [2]. The self-diffusivity $D_{\rm SD}$ is usually written as the sum of contributions from independent diffusive mechanisms. The contribution of a particular microscopic mechanism can be written as the product of the diffusivity, D_i, and the concentration, C_i, of the relevant defect, i.e., $D_{\rm SD} = \sum_i D_i C_i$. In a recent review, Gösele *et al.* [3] gave the best current experimental estimates of the contributions to the self-diffusivity as $D_{\rm I} C_{\rm I} = 914 \exp(-4.84/k_B T)$ cm^2 s^{-1} for self-interstitials and $D_{\rm V} C_{\rm V} = 0.6 \exp(-4.03/k_B T)$ cm^2 s^{-1} for vacancies, where $k_B T$ is in units of eV. These results indicate that self-interstitial diffusion is important at higher temperatures and is dominant above 1300 K. The experimental situation regarding self-diffusion in silicon is, however, still highly controversial, especially when it comes to the individual values of D_i and C_i. Indeed, experimental data has been used to support values of the diffusivity of the silicon self-interstitial, $D_{\rm I}$, which differ by 10 orders of magnitude at the temperatures of around 800 °C at which silicon is processed [4]. Clearly self-diffusion in silicon is both technologically important and imperfectly understood, and our aim is to perform an accurate theoretical study which can help clarify the situation.

Many theoretical studies of self-interstitials in silicon have been performed. The most advanced of these have used the local density approximation (LDA) to density-functional theory (DFT) to calculate the defect formation energies and energy barriers to diffusion [5,6]. Two molecular dynamics studies using the LDA have also been reported [6,7]. There are some differences between the results of the various LDA calculations, but the consensus view is that the split-$\langle 110 \rangle$, hexagonal, and tetrahedral defects (see Fig. 1) are low in energy. Other defects, such as the split-$\langle 100 \rangle$ and bond centered interstitials, are calculated to be more than 1 eV higher in energy, and therefore we do not consider them here. In a recent LDA study [7] a new defect named the "caged" interstitial was predicted to be very low in energy, and so we have included it in our study. We have also studied the saddle point of Pandey's concerted exchange mechanism [8] for self-diffusion, which involves the exchange of neighboring atoms in the perfect lattice. The aim of our work is to perform LDA, PW91 generalized gradient approximation (PW91-GGA), and fixed-node diffusion quantum Monte Carlo (DMC) calculations to determine the formation energies of the self-interstitials in silicon and see what modifications arise from the different treatment of electron correlation. Within each of the three methods we find the split-$\langle 110 \rangle$ and hexagonal interstitials to be the stablest. However, the formation energies of the stablest defects

0031-9007/99/83(12)/2351(4)$15.00

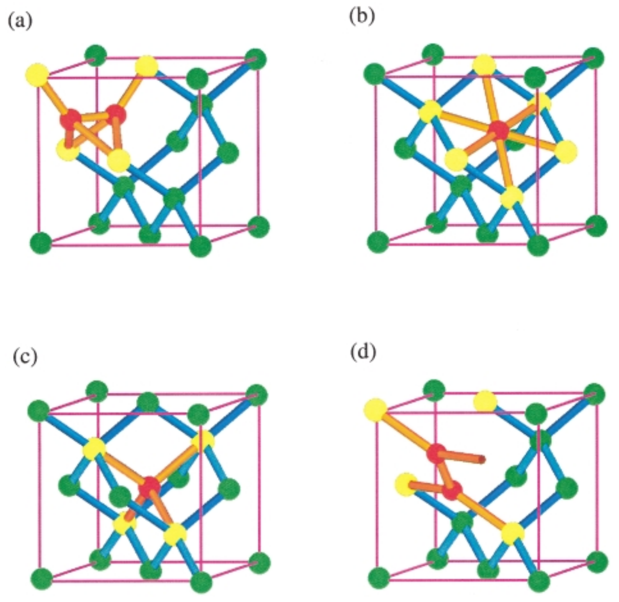

FIG. 1 (color). (a) The split-⟨110⟩, (b) hexagonal, and (c) tetrahedral interstitial defects, and (d) the saddle point of the concerted-exchange mechanism. The atom(s) forming the defect are shown in red, while the nearest neighbors to the defect atoms are shown in yellow. The bonds between the defect and nearest neighbor atoms are shown in orange.

are significantly different in the three methods. We use our DMC results to estimate the formation + migration energy for the self-interstitial contribution to self-diffusion as about 5 eV, which is in good agreement with recent experimental estimates [3].

As a preliminary to our main calculations we performed a thorough study at the LDA level [9]. The Si^{4+} ions were represented by a norm-conserving LDA pseudopotential. LDA calculations for the fully relaxed defect structures were carried out with fcc supercells for which the perfect crystalline structure contains 16, 54, and 128 atoms, respectively. The relaxed defect structures obtained with the 16, 54, and 128 atom cells were almost identical. The energies of the final structures were then recalculated using the PW91-GGA and DMC methods. We also recalculated the forces on the atoms using the PW91-GGA density functional [10], but the forces were small, demonstrating that the PW91-GGA and LDA structures are very similar. We performed extensive tests of the convergence of the calculations with respect to the size of the plane-wave basis set and the density of k-space sampling, and we estimate that the LDA and PW91-GGA defect formation energies given in Table I are converged to within 0.05 eV. (More details of the LDA and PW91-GGA calculations will be reported in a future publication [11].)

The structures of the split-⟨110⟩, hexagonal, and tetrahedral interstitial defects and of the saddle point of the concerted-exchange mechanism are illustrated in Fig. 1. The caged defect is not illustrated as it is similar to the split-⟨110⟩ interstitial. To obtain our caged structure we started from the atomic coordinates of Clark and Ackland [7] and relaxed the structure. The structure relaxed towards the split-⟨110⟩ interstitial and is therefore not a stable defect. We decided to investigate the structure obtained after partial relaxation, which is closer to the split-⟨110⟩ interstitial than the caged structure of Ref. [7]. One of the intriguing features of the various defect structures is the wide range of interatomic bonding they exhibit. The bond length for the perfect crystal is 2.35 Å. In the split-⟨110⟩ structure the two atoms forming the defect are fourfold coordinated, but two of the surrounding atoms are fivefold coordinated. The hexagonal interstitial is sixfold coordinated with bonds of length 2.36 Å joining it to six neighbors, which are therefore fivefold coordinated. The tetrahedral interstitial is fourfold coordinated and has bonds of length 2.44 Å joining it to its four neighbors, which are therefore fivefold coordinated. The charge accumulation in these bonds is small, implying they are weak. In Pandey's concerted-exchange mechanism two nearest neighbor atoms interchange via a complicated three-dimensional path which allows the atoms to avoid large energy barriers [8]. At the saddle point of the exchange two bonds are broken so that the exchanging atoms and two other atoms are threefold coordinated. At the saddle point the exchanging atoms are joined by a very short bond of length 2.15 Å, which has a large accumulation of electronic charge at its center and is presumably very strong. The degree and nature of the relaxations of the atoms surrounding the various defect sites also vary, being small for the tetrahedral interstitial but significant for the other defects.

Our LDA formation energy for the split-⟨110⟩ interstitial is in good agreement with the result of Blöchl *et al.* [6]. We find the hexagonal interstitial to be low in energy in agreement with earlier LDA results, and we find it to be essentially degenerate with the split-⟨110⟩ interstitial, with DFT formation energies of 3.31 eV (LDA) and 3.80 eV (PW91-GGA). In agreement with earlier LDA work we find the tetrahedral interstitial to be unstable to small displacements of the interstitial atom, while the split-⟨110⟩ and hexagonal interstitial defects are locally stable. The split-⟨110⟩ and hexagonal interstitials are therefore low-energy locally stable defects with DFT formation energies of about 3.3 eV (LDA) and 3.8 eV (PW91-GGA).

TABLE I. LDA, PW91-GGA, and DMC formation energies in eV of the self-interstitial defects and the saddle point of the concerted-exchange mechanism.

Defect	LDA	GGA	DMC[a]	DMC[b]
Split-⟨110⟩	3.31	3.84	4.96(24)	4.96(28)
Hexagonal	3.31	3.80	4.70(24)	4.82(28)
"Caged"	3.34	3.85	5.26(24)	5.17(28)
Tetrahedral	3.43	4.07	5.50(24)	5.40(28)
Concerted exchange	4.45	4.80	5.85(23)	5.78(27)

[a] The 16 atom supercell.
[b] The 54 atom supercell.

We have also calculated the energy barriers to diffusive jumps between the low-energy structures within DFT, finding a path for split-⟨110⟩–hexagonal diffusion with a barrier of 0.15 eV (LDA) and 0.20 eV (PW91-GGA), and a barrier for hexagonal-hexagonal diffusive jumps of 0.03 eV (LDA) and 0.18 eV (PW91-GGA).

Our LDA value for the activation energy of Pandey's concerted-exchange mechanism of 4.45 eV is in good agreement with earlier calculations [8]. Although the LDA activation energy and the PW91-GGA activation energy of 4.80 eV for Pandey's concerted exchange are within the experimental range for self-diffusion of 4.1–5.1 eV [2], our results indicate that the activation energies for interstitial mediated self-diffusion are lower. The prediction of two low-energy interstitial defects connected by low-energy paths might be the explanation of the large prefactor in the self-diffusivity [2,3]. However, our DFT results give the sum of the formation and migration energies for self-interstitials as 3.5 eV (LDA) and 4.0 eV (PW91-GGA), which are considerably smaller than the activation energy for self-interstitial diffusion deduced from experimental measurements of 4.84 eV [3].

We conclude that the DFT results for self-interstitial diffusion in silicon do not afford a satisfactory explanation of the experimental temperature dependence of the self-diffusivity. There are a number of possible reasons for this discrepancy. First, it would be more appropriate to compare our calculated results with measurements of the equilibrium self-diffusion constant at low temperatures, but unfortunately such experimental data do not exist. A further source of difficulty in comparing with experiment could be due to the self-diffusion being slowed by trapping of self-interstitials, for instance, by the carbon impurities which are normally present in silicon. There is also a technical reason for questioning the accuracy of the DFT for the self-diffusion problem. The wide range of coordination numbers (from 3 to 6) and the occurrence of weak and strong bonds in the defect structures present a very severe test of the accuracy of electronic structure methods. The LDA and PW91-GGA functionals are believed to be fairly reliable for describing changes in energy due to small displacements of the atoms in materials such as silicon, and consequently they give accurate phonon frequencies and structures, but they are not reliable for the energy differences between structures with very different interatomic bonding.

Given the doubts about the accuracy of the DFT functionals and the technological importance of the problem, we have recalculated the energies of the defect structures using the DMC method. The DMC method [12] is a stochastic method for solving the many-body Schrödinger equation. It is now well established that the DMC method can give an excellent description of electron correlation in the ground state. We performed DMC calculations for the 16 and 54 atom fcc structures obtained from our LDA study. Relaxation of atomic positions in solids within DMC has not yet been demonstrated, and we therefore used the structures obtained from our LDA study. We used the same pseudopotential as in our LDA study, and the nonlocal energy [13] was evaluated using the "locality approximation" [14]. Our guiding wave functions were of the Slater-Jastrow type with the single-particle orbitals obtained from LDA calculations with a basis set cutoff of 18 Ry calculated at the L point of the Brillouin zone. This gives a much better representation of the charge density and hence smaller finite size effects than the Γ point, while still allowing the use of real Bloch wave functions [15]. The "Coulomb finite size effects" [16] were found to be small, and we corrected for the residual "single-particle finite size effects" [16] using the results of our LDA study. We used localized spherically symmetric functions for the single-body part of the Jastrow term (denoted by χ), each consisting of a polynomial containing 16 parameters. For the perfect crystal we used the same χ function for each atom. The atoms in the defect structures are not all equivalent, and it was found beneficial to allow the χ functions on inequivalent atoms to differ. The parallel- and antiparallel-spin two-body part of the Jastrow term were constrained to obey the cusp conditions [17] and each contained eight parameters [18]. The guiding wave function for the perfect structure contained 32 parameters, while those for the defect structures contained 64 parameters. The optimal values were obtained by minimizing the variance of the energy [18,19].

First we calculated the cohesive energy of silicon within the variational Monte Carlo (VMC) and DMC methods by performing calculations for the atomic ground state and for the perfect solid using the 54-atom simulation cell. For the solid we included a correction for the zero-point energy of 0.061 eV per atom, and a finite size correction of 0.077 eV per atom obtained from the results of VMC and DMC calculations on cells containing up to 250 atoms [20]. The resulting VMC and DMC cohesive energies of 4.48(1) eV and 4.63(2) eV per atom, respectively, are in excellent agreement with the experimental value of 4.62(8) eV per atom.

The results of our DMC calculations for the defect formation energies are shown in Table I. First we note that the DMC results for the 16- and 54-atom simulation cells are consistent, indicating that the residual finite size effects are small. The clearest conclusion is that the DMC formation energies are roughly 1 eV larger than the PW91-GGA values and 1.5 eV larger than the LDA values. Within DMC the hexagonal interstitial has the lowest formation energy, while the split-⟨110⟩ interstitial is slightly higher in energy. The caged interstitial has a slightly larger formation energy than the split-⟨110⟩ interstitial, which is expected because it is a distortion of the split-⟨110⟩ interstitial, while the tetrahedral interstitial has a significantly higher energy. The saddle point of the concerted exchange has an energy which is too high to explain self-diffusion in silicon.

Since mapping the energy barriers to diffusion is computationally prohibitive within DMC we estimate the migration energy as follows. The tetrahedral interstitial is a saddle point of a possible diffusion path between neighboring hexagonal sites. The DMC formation energy of 5.4 eV for the tetrahedral interstitial is therefore an upper bound to the formation + migration energy of the hexagonal interstitial. The true formation + migration energy is expected to be less than this, and we can use the following argument to obtain a crude estimate of it. Within the LDA we have found a diffusion path for the hexagonal interstitial with a barrier of only 25% of the tetrahedral-hexagonal energy difference. Applying the same percentage reduction to the DMC barrier then gives an estimate of the formation + migration energy of the hexagonal interstitial of 5 eV. This estimate is in good agreement with the experimental activation energy for self-interstitial diffusion of 4.84 eV [3].

In summary, we have demonstrated the importance of a proper treatment of electron correlation when calculating defect formation energies in silicon. The formation energies are sensitive to the description of electron correlation. The LDA and PW91-GGA functionals do not provide a satisfactory explanation of the experimental temperature dependence of the self-interstitial contribution to the self-diffusivity because they predict formation + migration energies which are too small. The larger defect formation energies found in our DMC calculations indicate a possible resolution of this problem. This may be an important step in improving our understanding of self-diffusion in silicon.

We thank Paul Kent and Randy Hood for help with the calculations. Financial support was provided by the Engineering and Physical Sciences Research Council (UK), Hitachi Europe Ltd., and Hitachi Japan Ltd. The DMC calculations were performed on the Hitachi SR2201 at the University of Cambridge High Performance Computing Facility and at Hitachi's Central Research Laboratory.

[1] P. M. Fahey, P. B. Griffin, and J. D. Plummer, Rev. Mod. Phys. **61**, 289 (1989).
[2] W. Frank, U. Gösele, H. Mehrer, and A. Seeger, in *Diffusion in Crystalline Solids,* edited by G. E. Murch and A. S. Nowick (Academic, Orlando, FL, 1985), p. 64.
[3] U. Gösele, A. Plössl, and T. Y. Tan, in *Process Physics and Modeling in Semiconductor Technology,* edited by G. R. Srinivasan, C. S. Murthy, and S. T. Dunham (Electrochemical Society, Pennington, NJ, 1996), p. 309.
[4] D. Eaglesham, Phys. World **8**, 41 (1995).
[5] Y. Bar-Yam and J. D. Joannopoulos, Phys. Rev. B **30**, 1844 (1984).
[6] P. E. Blöchl *et al.,* Phys. Rev. Lett. **70**, 2435 (1993).
[7] S. J. Clark and G. J. Ackland, Phys. Rev. B **56**, 47 (1997).
[8] K. C. Pandey, Phys. Rev. Lett. **57**, 2287 (1986); E. Kaxiras and K. C. Pandey, Phys. Rev. B **47**, 1659 (1993).
[9] We used the LDA of J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981), who parametrized the DMC data of D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
[10] J. P. Perdew, in *Electronic Structure of Solids '91,* edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991), p. 11.
[11] R. J. Needs (unpublished).
[12] D. Ceperley, G. Chester, and M. Kalos, Phys. Rev. B **16**, 3081 (1977); B. L. Hammond, W. A. Lester, Jr., and P. J. Reynolds, *Monte Carlo Methods in ab initio quantum Chemistry* (World Scientific, Singapore, 1994).
[13] S. Fahy, X. W. Wang, and S. G. Louie, Phys. Rev. B **42**, 3503 (1990).
[14] M. M. Hurley and P. A. Christiansen, J. Chem. Phys. **86**, 1069 (1987); B. L. Hammond *et al., ibid.* **87**, 1130 (1987); L. Mitáš, E. L. Shirley, and D. M. Ceperley, *ibid.* **95**, 3467 (1991).
[15] G. Rajagopal *et al.,* Phys. Rev. Lett. **73**, 1959 (1994); G. Rajagopal *et al.,* Phys. Rev. B **51**, 10 591 (1995).
[16] A. J. Williamson *et al.,* Phys. Rev. B **55**, R4851 (1997).
[17] T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957).
[18] A. J. Williamson *et al.,* Phys. Rev. B **53**, 9640 (1996).
[19] C. J. Umrigar *et al.,* Phys. Rev. Lett. **60**, 1719 (1988).
[20] P. R. C. Kent *et al.,* Phys. Rev. B **59**, 1917 (1999).