Enhanced Solubility of Impurities and Enhanced Diffusion near Crystal Surfaces

J. Tersoff
IBM Research Division, Thomas J. Watson Research Center, Yorktown Heights, New York 10598
(Received 21 February 1995)

Any defect or impurity has some inherent stress, so near a free surface its energy is reduced by relief of this stress. The defect stress also couples to the intrinsic surface stress. In general, the result is to enhance impurity solubility and diffusion near the surface. With certain assumptions regarding the kinetics, the high impurity density near the surface can be frozen in as the crystal grows, permitting the growth of highly supersaturated solid solutions, e.g., high dopant concentrations in semiconductors. Calculations for carbon near the Si(001) surface illustrate the solubility enhancement.

PACS numbers: 68.35.Fx

The incorporation of impurities into a growing crystal is a classic problem in metallurgy and materials physics. This problem is of particular relevance to semiconductor technology, because of the need to achieve specific concentrations of donor or acceptor atoms on substitutional sites. More broadly, an understanding of the factors limiting incorporation is helpful in fabrication of nonequilibrium materials, an important and rapidly growing area of materials physics. Already, solutions of C in Si [1,2] and B in Si [3] have been grown with concentrations which are many orders of magnitude above the equilibrium solubility limit at the growth temperature. This raises the prospect of creating interesting and useful new materials, with increased flexibility in tailoring their properties.

Incorporation of impurities during growth is typically limited by segregation at the crystal surface. Unfortunately, a detailed treatment of the microscopic incorporation process is not feasible. Most attempts to address incorporation theoretically have therefore focused on how the equilibrium solubility depends upon the chemical potential [4,5], which is sensitive to growth conditions. Recently Rücker *et al.* [2] raised another intriguing possibility—that high concentrations of C in Si are possible, not only because of a high chemical potential for C, but also because the system can form low-energy ordered structures.

Here we show that yet another factor, the elastic interaction of the impurity with the surface, may also play a crucial role for B in Si, C in GaAs, and many other systems. For concreteness, we focus on C in Si. Detailed calculations indicate that the equilibrium solubility of C in Si is increased by several orders of magnitude for the first four layers at the Si(001) surface. Under reasonable assumptions regarding the kinetics, the carbon-enriched layers near the surface will be buried by further growth, without the opportunity for segregation of the excess carbon. This leads to high levels of substitutional C throughout the volume of the grown material.

The increased solubility results from two factors. First, the presence of a surface partially relieves the stress associated with the atomic size mismatch between C and Si. Second, there is a stress field near the surface, associated with the atomic reconstruction there. The coupling of the impurity stress to this spatially varying surface stress substantially lowers the impurity energy on certain sites. The first effect requires only the presence of a surface (a solid-vapor or solid-liquid interface) and so should be important in any system where atomic size mismatch limits the solubility. The second effect depends on the particular surface structure. However, the dimer structure studied here is relevant not only to Si and Ge (001) but also (with some modification) to (001) surfaces of GaAs and other III-V semiconductors. Thus the results here should be qualitatively applicable to such important dopants as B in Si and C in GaAs.

In addition, the rate of diffusion is controlled by the energy of a defect at the saddle point of it migration path. The saddle-point configuration is often highly strained. In that case its energy will be lowered near the surface, and hence the diffusion rate will be dramatically increased there. This has important implications both for the kinetics of impurity incorporation and as a factor limiting the ability to grow atomically sharp interfaces. This diffusion enhancement is illustrated by model calculations of the split vacancy, the saddle point for one possible mechanism of diffusion.

In equilibrium, the solubility of a substitutional impurity (i.e., the value of x for C_xSi_{1-x}) is [6]

$$x = e^{-(E-\mu)/kT}. \quad (1)$$

Here E is the energy of a substitutional C atom in Si, μ is the chemical potential for C, and k is Boltzmann's constant. This formula assumes that $x \ll 1$, which is always the case for C in Si. The concentration of C is limited by the formation of SiC precipitates, which prevents μ from rising higher than the value corresponding to a SiC reservoir.

However, epitaxial growth is an inherently nonequilibrium process. As a manageable model, the details of surface structure and kinetics are generally assumed to matter only in that they determine the effective potential μ for the impurity. For example, during growth the nucleation of SiC at the surface may be kinetically suppressed [2]. By raising μ, the amount of C incorporated during growth

 0031-9007/95/74(25)/5080(4)$06.00 © 1995 The American Physical Society

can then be increased beyond the equilibrium limit. Most studies of impurity solubility have focused on this role of μ in controlling the concentration [4,5].

However, here we are concerned with another effect which operates *in addition to,* and independently of, any change in the chemical potential. The presence of a surface breaks the translational symmetry of the crystal, so that sites in different layers are no longer equivalent. Thus the energy of a substitutional C atom on site n, E_n, is different for each inequivalent site, and Eq. (1) must be modified accordingly. The concentration of C on site n is then approximately [6]

$$x_n = e^{-(E_n - \mu)/kT}. \tag{2}$$

Thus to understand the limits of C incorporation, we must calculate the energy of substitutional C at all distinct sites near the Si surface.

The calculations here use an empirical many-body potential [7] to the model the interactions between atoms in the system. This method has been shown to have *predictive* accuracy for substitutional C in bulk Si [8]. Since C in Si is isovalent, electronic interactions between defect and surface (which are neglected in this classical model) are not expected to be important beyond the first atomic layer. In any case, since we are concerned primarily with demonstrating a general qualitative effect, the quantitative accuracy of the method is not a primary issue. The Si(001) surface is described using an 8×8 supercell, to minimize impurity interactions, with at least 20 layers being relaxed until all forces are less than 10^{-3} eV/Å (2×10^{-5} a.u.).

To distinguish the role of surface reconstruction, we calculate E_n for C atoms at both the dimerized Si(001)-(2×1) surface and a hypothetical unreconstructed (1×1) Si(001) surface. The results are shown in Figs. 1 and 2. (Results are not shown for C in the surface layer, layer 1, since the large contribution from the chemical change in the dangling bond makes the classical model less reliable there.)

For the unreconstructed surface, Fig. 1, the energy of the substitutional C is well described by

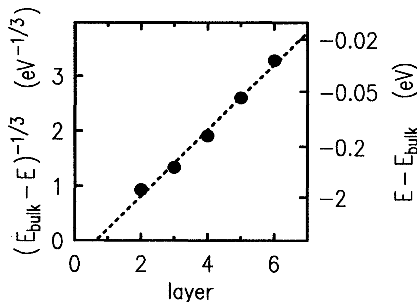

FIG. 1. Energy of a substitutional C atom near an unreconstructed Si(001) surface, relative to the energy of C in Si bulk. Surface is layer 1. Vertical axis is chosen to test Eq. (3); dotted line is the least squares fit.

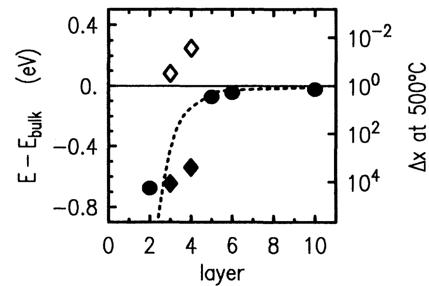

FIG. 2. Energy of a substitutional C atom near a reconstructed Si(001)-(2×1) surface, relative to bulk energy. (The calculated bulk energy is 1.6 eV, relative to a SiC reservoir [8].) Dotted line is reproduced from Fig. 1 for comparison. Filled and open diamonds correspond to α and β sites, respectively. Δx is the factor by which solubility is enhanced at each site, according to Eq. (2).

$$E \approx E_{\text{bulk}} - C(z - z_0)^{-3}, \tag{3}$$

consistent with the elastic interaction of the tensile C impurity with a compressive "image" impurity. The image plane z_0 lies about one third of a layer spacing outside the surface layer.

Energies of C near the 2×1 surface are given in Fig. 2. This case is considerably more complicated, because in layers 3, 4, 7, 8, ..., there are two inequivalent sites. As discussed by Kelires and Tersoff [9], the sites beneath the surface dimers are under compressive stress. Thus they are favorable sites for a smaller atom such as C. We shall call these the α sites. The β sites, between dimers, are under tensile stress, and so are highly unfavorable sites for C.

In general, there should also be a coupling of the impurity to the overall surface stress, decreasing with depth as z^{-1}. If the impurity and surface are both under tension, for example, this should give a repulsion, in addition to the attractive image force. However, this interaction is so weak for the 1×1 surface that it does not show up in Fig. 1. For the 2×1 surface, the site-to-site variations in the stress dwarf the effect of the average surface stress.

Figure 2 also shows the resulting change in solubility, relative to the bulk, at a typical growth temperature [1] of 500 °C. The solubility is calculated from the energies according to Eq. (2). Remarkably, down to the fourth layer the solubility is enhanced by roughly 10^4 relative to the bulk. (In layers 3 and 4, this must be reduced by a factor of 2, because only half the sites have low energy.)

The equilibrium enhancement of solubility is dramatic but affects only a few layers at the surface. The real significance of these results lies in their implications for the growth of thicker layers. Diffusion is known to be rapid on the surface and negligibly slow in the bulk, under typical growth conditions. As a simple model for the kinetics, suppose that diffusion among the first N

layers is rapid, so that these layers are in equilibrium with the surface, while diffusion to layers deeper than N is negligible. Suppose that the first M layers near the surface exhibit enhanced impurity solubility. Then if $M \geq N$, the first N layers are in equilibrium at the enhanced concentration. When, with further growth, the Nth layer becomes the $N + 1$st, it becomes inaccessible to diffusion, so the enhanced concentration is frozen in. For C at Si(001), we have seen that $M = 4$. Therefore, if $N \leq 4$ (as seems almost certainly true), the entire thickness grown will reflect the enhanced concentration characteristic of the surface layers.

The highest C concentrations which have been achieved experimentally are much more than 10^4 times higher than the equilibrium solubility at the growth temperature [1,2]. Thus, the present effect alone is not sufficient to explain the experiments. The ability to achieve elevated values of the C chemical potential is also essential, and specific strategies for suppressing nucleation of SiC have been discussed [2]. Still, whatever the chemical potential, the interaction of the defects stress with the surface should permit growth of C-Si alloys with a carbon content 4 orders of magnitude larger than would otherwise be possible.

The mechanism described here has another possible consequence, besides enhanced solubility, If the C has sufficient mobility within layers 3 and 4, as suggested by results below, then it will lie on α sites, with negligible β-site occupancy. This ordering should not be lost as the layers are buried deeper, where mobility is negligible. If the 2×1 intralayer ordering of C exhibits the appropriate phase relation between successive layers, the result will be exactly the same "bulk" ordering as in SiGe [9,10] and should be observable by x-ray diffraction.

The results above apply qualitatively to any substitutional atom having a large size mismatch with the host. This includes several important dopants, such as B and Sb, with two differences. The smaller the size mismatch, the weaker the effect, and when the substitutional atom is too large rather than too small, as for Sb in Si, it will prefer the β sites rather than the α sites.

Finally, it is interesting to speculate about the role of different surface orientations or growth conditions. For the unreconstructed (001) surface, the large solubility enhancement did not extend as deep as for the 2×1, increasing the likelihood that impurities could still segregate rather than be frozen in. Thus changes in the surface structure, e.g., in growth with surfactants or in liquid-phase epitaxy versus molecular-beam epitaxy, could drastically alter the achievable doping levels. Similarly, growth on a different surface, such as (111) or (110), could result in very different levels of incorporation.

We now turn to a related but rather different issue, that of diffusion. In equilibrium, the diffusion rate is proportional to the number of diffusing defects and to their rate of motion. These depend on the formation and migration energies E_f and E_m of the defects, respectively, as $e^{-E_f/kT}$ and $e^{-E_m/kT}$. The migration energy is $E_m = E_s - E_f$, where E_s is the energy of the defect at the saddle point along its migration path. Thus the total diffusion rate may be written

$$D \propto e^{-E_s/kT}. \tag{4}$$

Unfortunately, there are numerous obstacles to an accurate calculation of the saddle-point energy E_s. The actual diffusion mechanism has been controversial, even for self-diffusion in pure bulk Si, because several possible mechanisms have comparable saddle-point energies [11]. Impurity diffusion is even more complex. The symmetry is drastically reduced near the reconstructed surface, so one loses the guidance normally provided by symmetry in determining the position of the saddle point; and a much larger unit cell is required than in the bulk. As a result, *ab initio* calculations of subsurface defects and diffusion have not been attempted (although a few *tour-de-force* calculations of adatom diffusion have been reported [12–14]).

Despite these difficulties, it is important to estimate the extent to which diffusion may be enhanced near a surface. We therefore focus on one particular process, vacancy-mediated self-diffusion, to illustrate this effect. The saddle point for vacancy diffusion is the "split vacancy" [15], where atoms at two neighboring crystal sites are replaced by a single atom in the middle. The energy of the split vacancy must therefore be calculated at each possible site near the Si(001)-(2×1) surface.

The empirical interatomic potential used here gives reasonable defect energies [7,16]. However, in the present context only the defect stress is important. In an *ab initio* calculation [15] of the split vacancy in the bulk, while defect stress was not reported *per se*, the six neighbors of the central atom were found to relax inward by 0.28 Å. The result here is 0.26 Å, indicating a tensile stress almost identical to the *ab initio* result. Thus the energy lowering calculated here should be relatively accurate, despite the limitations of the method.

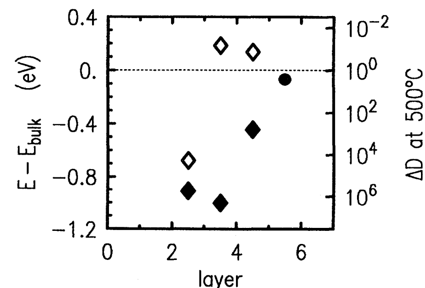

FIG. 3. Energy of split vacancy near the Si(001)-(2×1) surface, relative to energy in bulk. Filled and open diamonds refer to inequivalent sites in the same plane. Note that the split vacancy lies *between* layers, so there are two inequivalent sites except between layers 1 and 2, 5 and 6, etc.

The results for the energy of the split vacancy are shown in Fig. 3. The change in the diffusion rate, according to Eq. (4), is also shown. The predicted enhancement of the diffusion is extremely large, roughly 10^6 for the first four layers. This is probably enough to drastically affect heteroepitaxial growth.

The energies on some sites, though, actually go up. An examination of the geometry shows that all atoms in the first three layers exhibit enhanced diffusion, as do half the atoms in the fourth layer. However, the fourth-layer sites between dimers are even less accessible to vacancy diffusion than bulk sites.

A comparable enhancement seem likely for other diffusion mechanisms. For example, the saddle point for Si interstitial self-diffusion is believed to be the bond-centered configuration, which has a large compressive stress [17]. C diffusion in Si may also be via a compressive bond-centered interstitial [8].

Thus, for a variety of diffusion mechanisms, it is reasonable to expect that diffusion among the first four layers will be greatly enhanced. This satisfies the condition discussed above, for the surface layers with high impurity concentration to be buried without segregation, allowing growth of material which is supersaturated throughout. Enhanced diffusion in the third and fourth layers is also just what is needed to explain the observed ordering of semiconductor alloy layers as resulting from the different site stresses [9,10]. And surface-enhanced diffusion will tend to make heterointerfaces less abrupt.

Impurity incorporation and diffusion near a surface are complex problems. A complete picture will require detailed calculations using more accurate *ab initio* methods, as well as a better understanding of the growth kinetics. Nevertheless, the model calculations reported here show convincingly that impurities and other defects have a strong elastic interaction with the surface. This interaction can drastically enhance both impurity incorporation and diffusion during growth.

I am grateful to R.M. Tromp and A. Zangwill for stimulating this work and for valuable discussions.

[1] S.S. Iyer *et al.*, Appl. Phys. Lett. **60**, 356 (1992).

[2] H. Rücker, M. Methfessel, E. Bugiel, and H.J. Osten, Phys. Rev. Lett. **72**, 3578 (1994).

[3] B.S. Meyerson *et al.*, Appl. Phys. Lett. **50**, 113 (1987); B.S. Meyerson (private communication).

[4] C.G. Van de Walle, D.B. Laks, G.F. Neumark, and S.T. Pantelides, Phys. Rev. B **47**, 9425 (1993).

[5] J.E. Northrup and B.S. Zhang, Phys. Rev. B **47**, 6791 (1993).

[6] Equation (1) is obtained by minimizing the free energy with respect to x, including energy and configurational entropy for small x; see, e.g., C. Kittel and H. Kroemer, *Thermal Physics* (Freeman, San Francisco, 1980), 2nd ed. [Equation (86) on page 80 should be rederived using Eq. (80) rather than Eq. (83) to avoid the erroneous factor of e^{-1}.] Vibrational entropy is neglected here. For spatially inhomogeneous systems, the local approximation of Eq. (2) may be improved by taking into account the concentrations on neighboring sties; see, e.g., J. Tersoff, Phys. Rev. B **42**, 10 965 (1990).

[7] For background on the method, see J. Tersoff, Phys. Rev. Lett. **56**, 632 (1986); Phys. Rev. B **37**, 6991 (1988); **38**, 9902 (1988); **39**, 5566 (1989).

[8] J. Tersoff, Phys. Rev. Lett. **64**, 1757 (1990). For further support of the conclusions of this reference, see A. Dal Pino, Jr., A.M. Rappe, and J.D. Joannopoulos, Phys. Rev. B **47**, 12 554 (1993).

[9] P.C. Kelires and J. Tersoff, Phys. Rev. Lett. **63**, 1164 (1989).

[10] F.K. LeGoues, V.P. Kesan, S.S. Iyer, J. Tersoff, and R. Tromp, Phys. Rev. Lett. **64**, 2038 (1990).

[11] R. Car, P.J. Kelly, A. Oshiyama, and S.T. Pantelides, Phys. Rev. Lett. **52**, 1814 (1984); P.E. Blöchl, Phys. Rev. Lett. **70**, 2435 (1993), and references therein.

[12] P.J. Feibelman, Phys. Rev. Lett. **65**, 729 (1990).

[13] G. Brocks, P.J. Kelly, and R. Car, Surf. Sci. **269–270B**, 860 (1992).

[14] R. Stumpf and M. Scheffler, Phys. Rev. Lett. **72**, 254 (1994).

[15] P.J. Kelly, R. Car, and S.T. Pantelides, Mater. Sci. Forum **10–12**, 115 (1986).

[16] The *shape* of the energy barrier along the diffusion path is not as well described, because of the somewhat abrupt cutoff of the potential with distance. In fact, the split vacancy is a stable configuration with this potential, rather than a saddle point.

[17] A. Antonelli and J. Bernholc, Phys. Rev. B **40**, 10 643 (1989).