Materials Science and Engineering, B14 (1992) 291–296

Theory of ballistic electron emission spectroscopy of $NiSi_2$–Si(111) interfaces

M. D. Stiles

National Institute of Standards and Technology, Gaithersburg, MD 20899 (USA)

D. R. Hamann

AT&T Bell Laboratories, Murray Hill, NJ 07974 (USA)

(Received January 1, 1992)

Abstract

We discuss theoretical calculations of ballistic electron emission microscopy spectra for A- and B-type $NiSi_2$–Si(111) interfaces. The calculations are based on a first-principles computation of the transmission across the interfaces and a model calculation of the electrons tunneling from the tip. We compare the calculated spectra with existing experimental data and discuss a way of presenting experimental data that highlights the transmission process with respect to contributions from the tunneling distribution.

Ballistic electron emission microscopy (BEEM) is an experimental technique based on scanning tunneling microscopy that can be used to investigate properties of buried interfaces. In this paper we discuss the calculation of BEEM spectra for $NiSi_2$–Si(111) interfaces [1], and how the results compare with measured spectra. The spectra depend mainly on the distribution of electrons that tunnel from the tip and the transmission probability for electrons across the interface. We discuss a technique for analyzing experimental data that emphasizes the transmission process by canceling as much of the details of the tip-surface tunneling process as possible. While low transmission limits the technological potential of $NiSi_2$–Si(111) interfaces [2], they are good test systems for understanding BEEM experiments and electron transmission across coherent interfaces. A detailed understanding of the results should be of value in assessing the application of these techniques to other systems.

In a BEEM experiment [3], a scanning tunneling microscope tip is held over the thin (about 50 Å) metal overlayer, grown on the semiconductor substrate. A voltage is applied between the tip and the overlayer, with the overlayer and substrate held at the same voltage. Electrons are injected into the overlayer, some of which travel ballistically across the overlayer and are incident on the Schottky barrier at the buried interface. Those electrons that transmit across the interface are measured as a "collector" current between the tip and the substrate. The collector current, as a function of tip voltage, can be used to infer properties of the interface, such as the Schottky barrier height, and the behavior of the transmission across the interface. By scanning the tip while measuring the collector current, local variations in properties of the interface can be imaged.

$NiSi_2$ grown on Si(111) is an important test case for BEEM. The interfaces between these materials can be grown atomically abrupt and coherent, and can be grown with two interface structures depending on the growth conditions [4]. In the A-type interface, the $NiSi_2$ lattice has the same orientation as the silicon substrate lattice; in the B-type interface, the orientation is reversed. The lattice constants are matched to better than 1%, and the interface quality is quite good [5]. In a recent calculation [2], we showed that the transmission properties of these interfaces differed by a factor of about 3 for energies close to the Schottky barrier threshold. Since BEEM can probe the transmission across interfaces, it should be possible to see this difference, given the ability to grow A-type and B-type interfaces on the same substrate, thus providing a test of our understanding of both BEEM and interface transmission.

The collector current in a BEEM experiment is equal to the integral of the flux distribution incident on the buried interface times the probability that an electron in each state will be transmitted across the interface. This integral can be written as a two-

dimensional integral over wavevectors and an integral over energy:

$$I_c(V,d) = e\int_0^V \mathrm{d}E \int_{\mathrm{IBZ}} \frac{\mathrm{d}^2K}{(2\pi)^2} \times \sum_n \rho_1(n,E,\boldsymbol{K},V,d)\,T(n,E,\boldsymbol{K}) \qquad (1)$$

Here, E is the electron energy relative to the metal Fermi level, IBZ is the interface Brillouin zone, $\boldsymbol{K}$ is a wavevector in the interface Brillouin zone, n refers to all states in the overlayer with a given E and $\boldsymbol{K}$, ρ_1 is the distribution of electrons incident on the interface, V is the voltage applied between the tip and overlayer, d is the tip sample separation and T is the transmission probability across the interface.

For transmission through coherent interfaces, those with a common interface lattice net for both materials, an electron's crystal momentum parallel to the interface, as well as its energy, are conserved. These two conservation principles constrain the transmission probability T such that, if there is no state in the substrate with the same E and $\boldsymbol{K}$, the transmission probability will be zero. These kinematic constraints depend only upon the band structures of the two materials and the Schottky barrier height of the interface and they indicate which states might possibly contribute to electron transmission across the interface. The kinematic constraints for $CoSi_2$–Si(111) interfaces indicate that there is a projected gap in the $CoSi_2$ at wavevectors equal to the projection of the silicon conduction band minimum [6, 7]. This gap leads to a delay in the onset of transmission of about 0.2 V in calculated BEEM spectra and has been observed experimentally [8].

The kinematic constraints are not sufficient to calculate the detailed shape and size of BEEM spectra which depend on the transmission probabilities across the interface. These transmission probabilities must be found from a calculation which takes into account the atomic-scale details of the wavefunctions in each material and of the potential at the interface. We calculate them using an *ab-initio* method [9] developed to treat complicated interfaces such as $NiSi_2$–Si(111) [2]. BEEM measurements on coherent interfaces are sensitive to the details of the band structures of both the substrates and the overlayers, so that simple models [3, 10–12] for the band structures and the transmission probabilities cannot be used.

Following earlier models of BEEM spectroscopy, we use a modified planar-barrier tunneling model to describe the current distribution injected into the overlayer from the tunneling tip:

$$I_t(V,d) = e\int_0^V \mathrm{d}E \int_{\mathrm{IBZ}} \frac{\mathrm{d}^2K}{(2\pi)^2} \sum_n t(E_\perp, V, d) \qquad (2)$$

where t is a Wentzel–Kramers–Brillouin approximation to the tunneling probability, which depends on the energy minus the kinetic energy of motion parallel to the interface: $E_\perp = E - K^2/2m$. (We use atomic units; $\hbar$ is unity, masses are in electron masses, distances are in Bohr radii and energies are in hartrees.) In addition, we make a correction for the applied voltage V across the barrier of height V_b:

$$t(E_\perp, V, d) = t_0 \exp\left(-2d[2m(V_b - E_\perp)]^{1/2} \times \frac{1-[1-V/(V_b-E_\perp)]^{3/2}}{\frac{3}{2}V/(V_b-E_\perp)} \right) \qquad (3)$$

The factor t_0 is an unknown scale factor included because of the many unknown factors in the actual tunneling distribution, such as the atomic-scale details of the overlayer surface electronic structure, image potential effects, the detailed shape of the tunneling tip and tip densities of states.

The main difficulty in comparing calculated BEEM spectra with the existing spectra is a lack of knowledge about the tunneling distribution. Without more information it is necessary to treat both t_0 and d as free parameters. In this case it is possible to generate a family of spectra depending on these values. It is possible to remove much of the uncertainty in the collector current due to uncertainties in the tunneling distribution by measuring the spectra in the constant-height mode and analyzing the data in a way discussed below. However, it is still necessary to constrain the tip distribution as much as possible to allow a unique comparison. In a constant-height measurement, the tip and the collector currents vary as a function of voltage for fixed tip-to-sample separation. Such measurements are more difficult to do but easier to interpret than constant-current measurements in which the tip-to-sample separation and the collector current vary for fixed tip current. They are more difficult because the current varies over a wider range but easier to interpret because the tip-to-sample separation is not a function of voltage. Below, we present calculations for spectra measured using both a constant-height technique, and a constant-current technique.

When scattering in the overlayer and substrate are not important, the tunneling and interface distributions are the same. The calculations discussed below assume that this is the case. Whether or not this assumption is correct in a particular experimental configuration is not clear *a priori*. Hopefully the comparison between

theory and experiment can be helpful in determining whether it is. However, unless the tunneling distribution can be constrained experimentally, it will be impossible to tell whether disagreement is due to uncertainty in the tunneling distribution or the scattering in the overlayer.

In Fig. 1, we show the collector current and tip current as a function of tip voltage for spectra calculated in constant-height mode for $NiSi_2$–Si(111) A- and B-type interfaces. The different curves are calculated assuming different tip-to-sample separations d and then scaled, by adjusting t_0, to give a tip current of 1 nA at a tip voltage of 2.0 V. The tip currents depend only on the overlayer and are the same for the two interfaces considered. These calculations should be compared with the constant-height measurements such as those of Hasegawa and coworkers [13]. To make a complete comparison, it would be desirable to compare both the collector current and the tip current. The latter could be used to choose a single combination of t_0 and d and to restrict the comparison with the collector current.

The collector currents are seen to be more sensitive to the tip-to-sample separation than the tip currents are. The increased sensitivity comes from a combination of two factors. The first factor is that for (111) interfaces the silicon conduction band minima are away from the interface Brillouin zone center, $\boldsymbol{K}=0$. The other factor is that the tunneling distribution becomes more and more concentrated toward the zone center as the tip-to-sample separation increases. Thus, for a given current, the fraction that can transmit across the interface decreases as the tip-to-sample separation increases. This effect becomes less important above the energy of the silicon L valley minimum, one of which is at the center of the interface Brillouin zone, $\boldsymbol{K}=0$. This energy is at 1.66 eV above the Fermi level for the A-type interface and at 1.81 for the B-type interface and is the source of the rapid onset seen in the collector currents at the corresponding voltages. The onset appears to occur at slightly lower voltages owing to unrelated changes in a different part of the Brillouin zone.

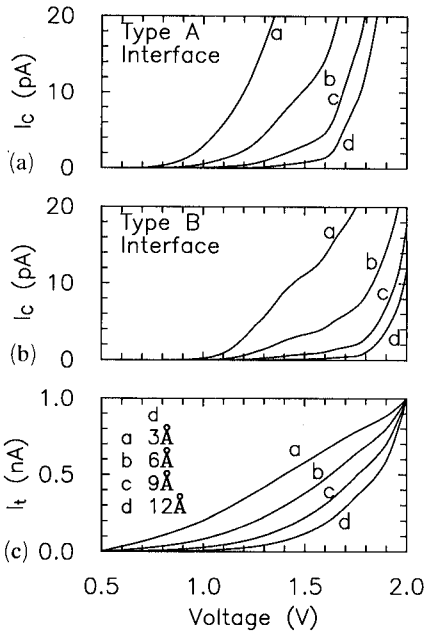

Fig. 1. Constant-height BEEM spectra for $NiSi_2$–Si(111): (a) the collector currents I_c for A-type interfaces; (b) the collector currents I_c for B-type interfaces; (c) the tip current I_t as a function of tip voltage for a series of tunneling distributions, characterized by the tip-to-sample separation d as labeled (as discussed in the text the tunneling distribution is scaled to give a tip current of 1 nA at a tip voltage of 2.0 V).

While the comparison that can be made with the experimental data [13] is limited owing to the uncertainty in the tunneling current, several features can be compared. First, the experimental data seem to show the difference in threshold expected from the difference in Schottky barrier heights for the two interfaces, but this determination is somewhat uncertain because of the slow onset seen both experimentally and in the calculations. Second, the A-type interface shows a higher collector current as is expected from the calculation, but this difference could be due to many other factors, such as different overlayer thicknesses or defect densities. Third, if the collector currents are made to agree close to threshold, then the experimental curves increase much more slowly with increasing voltage than the calculated curves do. This difference is consistent with an energy-dependent inelastic scattering rate that reduces the ballistic transport of electrons across the metal overlayer more and more as the electron energy increases. Finally, the data show an onset that could be due to the L valley minimum, but the onset voltage is different. This difference is not unexpected since we calculate the band structures using the local-density approximation, which does not reproduce band structures exactly. Experimentally the energy of the L valley minimum has been measured to be 2.1 eV [14] and 2.4 eV [15] above the valence band maximum, which would put the expected location of the thresholds at slightly higher voltages than are seen in the BEEM spectra. The apparent thresholds at this minimum are not separated by the difference in the Schottky barrier heights for the two interfaces. This could possibly be due to the additional structure seen in the calculated curves for the B-type interface that is due to changes in the $NiSi_2$ band structure. It would be useful to have the spectra measured over a larger voltage range to see whether there is a threshold where it would be expected to be based on the measurements of the L valley minimum. This is likely to be difficult

owing to the large range of currents that would be involved.

If the onset at higher voltages is due to the L valley minimum in the silicon, then it is likely that the interface is well ordered and that the scattering from disorder in the overlayer is not a dominant effect. Disorder in the bulk would tend to smooth out the parallel momentum distribution of the current incident on the interface. Thus the collector current would tend to be stronger than expected because there would be more current with a parallel momentum close to that of the silicon conduction band minimum. Disorder at the interface would not change the incident distribution but would allow states with different parallel momenta to couple. In general, it would be difficult to tell the difference between the two scattering mechanisms. One way would be the overlayer thickness dependence of the effect. The other is applicable only to special cases, such as that of $CoSi_2$–Si(111) [6–8], where there is a projected gap in the metal band structure corresponding to the semiconductor conduction band minimum. In such cases, bulk scattering would not change the delayed onset that is seen, but scattering at the interface would.

In Fig. 2, we show the collector current and tip-to-sample separation as a function of tip voltage for spectra calculated in constant-current mode for

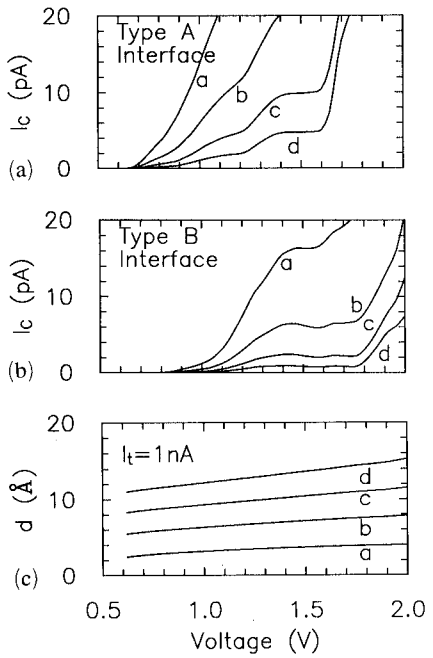

Fig. 2. Constant-current BEEM spectra for $NiSi_2$–Si(111): (a) the corresponding collector currents I_c for A-type interfaces; (b) the corresponding collector currents for B-type interfaces; (c) the tip-to-sample separation d as a function of tip voltage for a series of tunneling distributions with the scale factor t_0 adjusted to give a tip current of 1 nA as discussed in the text.

$NiSi_2$–Si(111) A- and B-type interfaces. The different curves are calculated assuming different tip-to-sample separations d at a particular voltage and then scaled, by adjusting t_0, to give a tip current of 1 nA. The tip-to-sample separations depend only on the overlayer and are the same for the two interfaces considered. These calculations should be compared with the constant-current measurements like those of Fernandez *et al.* [10]. While in principle, the tip-to-sample separation as a function of voltage could be used to pin down the tunneling distribution, in practice this is difficult. Since the tunneling probability depends exponentially on the tip-to-sample separation, small uncertainties in the separation translate into large uncertainties in the distribution. In addition, the absolute value of the separation is not well known; only relative changes can be accurately measured. The measured changes in tip-to-sample separation [10] are consistent with those shown here. Unfortunately they are consistent with a range of separations so they do not strongly restrict the tunneling distribution.

Any structure in the collector current tends to be more prominent in constant-current calculations than in constant-height calculations. The constant-current collector current is comparable with the collector current divided by the tip current in a constant-height measurement. One example of the increased sensitivity is the large difference in the onset above the silicon L valley minimum between the two interfaces. The difference is due to the difference in the density of states in the $NiSi_2$ at the different energies relevant to the onset for the two interfaces. There is a reduced density of states at the center of the interface Brillouin zone, $\boldsymbol{K}=0$, in the $NiSi_2$ at the higher energies relevant to the B-type interface.

The measured collector currents are of the same magnitude as the calculated currents, but they do not show any of the structure found in the calculated spectra. One possible explanation of the difference is that there is both strong elastic scattering and strong inelastic scattering in the metal overlayer. Strong elastic scattering would increase the flux incident on the interface in large parallel momentum states and greatly increase the collector current. Simple calculations (not shown) in which the tip distribution is randomized are consistent with this expectation; the collector current increases by about an order of magnitude. Consistency with the measured collector currents then requires strong inelastic scattering to reduce the calculated collector currents to the measured values. If this explanation is correct, the collector current should be very sensitive to the overlayer thicknesses in these measurements.

This comparison with experiment highlights the need to use experimental data to restrict the tunneling

distribution as much as possible. This is best done in a constant-height measurement because the tip current as a function of voltage at constant height is more sensitive to the details of the tunneling distribution than the tip-to-sample separation as a function of voltage at constant current. The tip current as a function of voltage should constrain the model for the tunneling distribution to one value of the tip-to-sample separation d and one value of the scale factor t_0.

If the collector current is then also measured in constant-height mode, presenting the data as the derivative of the collector current divided by the derivative of the tip current as a function of voltage [1] highlights the contribution of the transmission across the interface to the measurement. This ratio is proportional to an average of the transmission probability over parallel momentum at the energy of the applied voltage eV:

$$\frac{\mathrm{d}I_c}{\mathrm{d}V}\Big/\frac{\mathrm{d}I_t}{\mathrm{d}V}\approx\frac{2}{3}\int_{\mathrm{IBZ}}\frac{\mathrm{d}^2K}{(2\pi)^2}\sum_n W(\boldsymbol{K}, V, d)\,T(n, eV, \boldsymbol{K}) \qquad (4)$$

where W is a normalized weighting function that depends on the tip-to-sample separation but only weakly on the voltage for low voltages:

$$W(\boldsymbol{K}, V, d)=t\left(eV-\frac{K^2}{2m}, V, d\right) \Big/ \int_{\mathrm{IBZ}}\frac{\mathrm{d}^2K'}{(2\pi)^2}\sum_{n'} t\left(eV-\frac{K'^2}{2m}, V, d\right) \qquad (5)$$

Any unknown scale factors in the tunneling distribution, such as t_0, cancel out in this expression, leaving a weighting factor that is roughly gaussian in the parallel wavevector and dependent on the tip-to-sample separation.

The currents depend on the voltage both through the upper cut-off of the energy integration and because the barrier is skewed by the voltage. Differentiating the tip current with respect to voltage gives a contribution from both dependences:

$$\frac{\mathrm{d}I_t(V, d)}{\mathrm{d}V}=e\int_{\mathrm{IBZ}}\frac{\mathrm{d}^2K}{(2\pi)}\sum_n t\left(eV-\frac{K^2}{2m}, V, d\right) + e\int_0^V \mathrm{d}E\int_{\mathrm{IBZ}}\frac{\mathrm{d}^2K}{(2\pi)^2}\sum_n\frac{\mathrm{d}t(E_\perp, V, d)}{\mathrm{d}V} \qquad (6)$$

The first term is the contribution from the upper cut-off which removes the energy integral from eqn. (2) and sets the energy E equal to the voltage eV. The other term is roughly half the first term, because to a good approximation the tunneling probability depends on the energy and voltage in the combination $E+eV/2$:

$$t(E_\perp, V, d)\approx t\left(E_\perp+\frac{eV}{2}, d\right) \qquad (7)$$

If this relation is used in the second term, the derivative with respect to voltage can be changed into a derivative with respect to energy and the integration can be carried out. Because of the factor of $\frac{1}{2}$ in the approximation for the tunneling probability the contribution from the second term is approximately half the contribution from the first term. On the contrary, differentiating the collector current with respect to voltage gives a result that is dominated by the contribution from the upper cut-off. The energy dependence of the transmission probability reduces the contribution from the voltage dependence of the tunneling, thus reducing the relative contribution of the second term. Taking the ratio of the leading contributions to these two derivatives gives the approximate expression eqn. (4) which is an average of the transmission probability over the interface Brillouin zone at the energy of the applied voltage eV.

The weighting factor W in eqn. (5) still depends on the tip-to-sample separation. For systems with substrates such as Si(111), where the conduction band minima are not at the zone center, this dependence can be used as an additional experimental variable that can be used to restrict the comparison with theory. Measurements of the tip current as a function of voltage at several tip-to-sample separations would not only constrain the parameters in the tunneling model but could also test whether the model was correct. Then, measuring the collector current at several tip-to-sample separations would be a very stringent test of the transmission calculation.

Since the comparison of calculations with existing experimental data indicates that inelastic scattering and in some cases elastic scattering are important, it would be useful to carry out measurements for several thicknesses of the metal overlayer. The changes in the BEEM spectra should indicate how important scattering is for different systems. The BEEM spectra will be changed by this scattering because the distribution ρ_1, of electrons incident on the interface, will be changed from the distribution ρ_t tunneling from the tip. Elastic scattering, due to defects in the overlayer, will not change the electron energy, but will just redistribute the electrons between different n and $\boldsymbol{K}$. This will effectively change the weighting function W in eqn. (4). For systems with the conduction band minima away from the zone center, such as those considered here, strong elastic scattering will increase the average transmission close to threshold and reduce the sensitivity on the tip-to-sample separation. Inelastic scattering, whether due

to phonons or electron–hole pairs, is likely to decrease the collector current because some of the electrons will be scattered into states below the Schottky barrier. The effect of these scattering mechanisms on ρ_I should increase in proportion to the overlayer thickness.

An additional factor that can complicate the comparison between the experimental data and the calculated transmission probabilities is scattering in the semiconductor substrate [12]. It is possible for electrons to scatter back into the overlayer after they have successfully transmitted across the interface into the substrate. It should be possible by changing the doping of the substrate and the temperature to minimize this effect.

We have shown that detailed calculations of the BEEM spectra can be carried out for atomically abrupt coherent interfaces such as $NiSi_2$–Si(111), based on a first-principles calculation of the band structures and the transmission probabilities across the interface. We have discussed the comparison of the calculated spectra with the existing experimental data and discussed what we would envisage as an ideal BEEM experiment for further comparison. The largest uncertainty in the comparison is in the tunneling distribution; measuring the spectra in constant-height mode and reporting the tip current as a function of voltage would greatly restrict this uncertainty. Reporting the experimental data in terms of the ratio of the derivative of the collector current to the derivative of the tip current would further reduce the dependence on the uncertain tunneling distribution and highlight the contributions from the interface transmission. The $NiSi_2$–Si(111) interfaces are a good test system for developing the techniques for analyzing BEEM spectra. Further efforts along these lines would increase our understanding of BEEM investigations of buried interfaces, and electron transmission across interfaces.

References

1 M. D. Stiles and D. R. Hamann, *Phys. Rev. Lett., 66* (1991) 3179.
2 M. D. Stiles and D. R. Hamann, *Phys. Rev. B, 40* (1989) 1349.
3 W. J. Kaiser and L. D. Bell, *Phys. Rev. Lett., 60* (1988) 1406.
L. D. Bell and W. J. Kaiser, *Phys. Rev. Lett., 61* (1988) 2368.
L. D. Bell, M. H. Hecht, W. J. Kaiser and L. C. Davis, *Phys. Rev. Lett., 64* (1990) 2679.
M. H. Hecht, L. D. Bell, W. J. Kaiser and L. C. Davis, *Phys. Rev. B, 42* (1990) 7663.
4 R. T. Tung, J. M. Gibson and J. M. Poate, *Phys. Rev. Lett., 50* (1983) 429.
5 D. Cherns, G. R. Anstis, J. L. Hutchinson and J. C. H. Spence, *Philos. Mag. A, 46* (1982) 849.
E. Vlieg, A. E. M. J. Fischer, J. F. van der Veen, B. N. Dev and G. Materlik, *Surf. Sci., 178* (1986) 36.
E. J. van Loenen, J. W. M. Frenken, J. F. van der Veen and S. Valeri, *Phys. Rev. Lett., 54* (1985) 827.
6 M. D. Stiles and D. R. Hamann, *J. Vac. Sci. Technol. B, 9* (1991) 2394.
7 L. F. Mattheiss and D. R. Hamann, *Phys. Rev. B, 37* (1988) 10623.
8 W. J. Kaiser, M. H. Hecht, R. W. Fathauer, L. D. Bell, E. Y. Lee and L. C. Davis, *Phys. Rev. B, 44* (1991) 6546.
9 M. D. Stiles and D. R. Hamann, *Phys. Rev. B, 38* (1988) 2021.
10 A. Fernandez, H. D. Hallen, T. Huang, R. A. Buhrman and J. Silcox, *J. Vac. Sci. Technol. B, 9* (1991) 590.
A. Fernandez, H. D. Hallen, T. Huang, R. A. Buhrman and J. Silcox, personal communication.
11 M. Prietsch and R. Ludeke, *Phys. Rev. Lett., 66* (1991) 2511.
M. Prietsch and R. Ludeke, *Surf. Sci., 251–252* (1991) 413.
12 E. Y. Lee and L. J. Schowalter, *J. Appl. Phys., 65* (1989) 4903.
E. Y. Lee and L. J. Schowalter, *J. Appl. Phys.*, (1992), in the press.
L. J. Schowalter and E. Y. Lee, *Phys. Rev. B, 43* (1991) 9308.
13 Y. Hasegawa, Y. Kuk, R. T. Tung, P. J. Silverman and T. Sakurai, *J. Vac. Sci. Technol. B, 9* (1991) 578.
Y. Hasegawa and Y. Kuk, personal communication.
14 R. Hulthen and N. G. Nilsson, *Solid State Commun., 18* (1976) 1341.
15 D. Straub, L. Ley and F. J. Himpsel, *Phys. Rev. Lett., 54* (1985) 142.