PHYSICAL REVIEW B VOLUME 58, NUMBER 16 15 OCTOBER 1998-II

Alloy decomposition during growth due to mobility differences

P. Venezuela
University of Wisconsin, Madison, Wisconsin 53706

J. Tersoff
IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598

(Received 29 June 1998)

During growth of an alloy, any evolution of the surface morphology can lead to nonuniform composition. This occurs because there is always some difference in the mobility of the respective atoms. We examine in particular the much-studied case of a surface under stress. In this case there are two mechanisms driving decomposition—the different atomic sizes and the different mobilities. Depending on the system, the effect of mobility difference can dominate or even cancel the well-known effect of atomic size difference. [S0163-1829(98)04740-7]

Many technological applications of alloys require a uniform composition. This is particularly true of semiconductor devices, where small variations in local composition can scatter carriers and shift threshold voltages. Some alloys are inherently unstable against spinodal decomposition; but composition fluctuations are observed even in nominally stable alloys. Thus there has been intense interest in the fact that local variations in stress at a nonplanar surface during growth can lead to significant inhomogeneities in alloy composition.[1–3] The underlying mechanism is simple: the larger atoms incorporate preferentially in the regions that are (relatively) more tensile, and the small atoms in the more compressive regions.

Here we show that atomic size difference is not the only mechanism, or even necessarily the dominant one, leading to alloy decomposition during growth. Differences in atomic mobility on the surface can lead to substantial decomposition, even when the atomic sizes are identical. Conversely, mobility differences can cancel the decomposition which would otherwise result from atomic size differences. We demonstrate these important effects by analyzing the growth of an alloy surface under stress, both for step-flow growth and in a continuum model. We find that alloy decomposition is present even if the alloy components have the same size, as for $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{As}$ grown on a misfitting substrate.

In fact, the mechanism of differential mobility does not require strain at all. Any morphological evolution of an alloy surface will lead to decomposition. Potential examples include surface faceting, and step bunching due to step-edge diffusion barriers. We illustrate such strain-free decomposition by analyzing the decay of surface roughness for an unstrained alloy.

A particular interesting and important case is that of an AB alloy on a substrate of pure B, e.g., SiGe on Si, or $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{As}$ on GaAs. Typically the mobilities of A and B may differ by an order of magnitude or more. In that case, if A is the more mobile species (as for SiGe on Si), the mobility difference roughly doubles the degree of alloy decomposition, relative to the case of equal mobilities. However, when A is the less mobile species, as for SiGe on Ge, the decomposition can be almost entirely suppressed.

We begin with a linear instability analysis for the case of step-flow growth, followed by the corresponding analysis for a continuum model. As far as possible, we use consistent approximations in the two cases. The results are sufficiently similar that we can then discuss both cases together, drawing general conclusions about the relative importance of mobility and size differences in alloy decomposition. We make certain simplifications in both cases: we include only the terms that are of lowest order in the surface modulation (as in any instability analysis); we also neglect terms of order ε (the strain) relative to unity; and we neglect the energy of mixing of the alloy (though we include the entropy of mixing). For simplicity, and to make the results more transparent, we also neglect the additional stresses that are caused by the decomposition. These have been treated previously both for continuum and step-flow models.[2–4] They are negligible for sufficiently fast growth or small atomic size difference, and in any case they do not affect our general conclusions.

For the step-flow case we follow the approach of Ref. 3. On a vicinal surface (a staircase of terraces and atomic steps), step motion arises from attachment or detachment of adatoms, which diffuse over the terraces. For simplicity we do not consider the possible dependence of adatom mobility on the local strain or terrace composition, and we treat all terraces and steps as equivalent.

The adatom density on a terrace obeys the diffusion equation, subject to boundary conditions of equilibrium at the steps. (We do not include any step-edge diffusion barriers.[5]) Solving this equation we find the current of adatoms. The discontinuity in this current at a step gives the net attachment or detachment rate, and hence the step velocity. The velocity v_m of step m can then be decomposed into contributions of the respective species ν, writing $v_m=\Sigma_\nu v_{\nu m}$, where

$$v_{\nu m}=\frac{F_\nu}{2}(x_{m+1}-x_{m-1}) + D_\nu\left(\frac{\eta_{\nu,m+1}-\eta_{\nu m}}{x_{m+1}-x_m}-\frac{\eta_{\nu m}-\eta_{\nu,m-1}}{x_m-x_{m-1}}\right). \tag{1}$$

Here F_ν and D_ν are the incident flux and adatom mobility of the ν component; $\eta_{\nu m}$ is the adatom density in equilibrium with step m; and x_m is the step position. Thus, the relative composition of each atomic component at the steps is given by $c_{\nu m}=v_{\nu m}/v_m$.

The adatom density at step m is given by

$$\eta_{\nu m}=\exp\left(\frac{\mu_{\nu m}-E_\nu^{\mathrm{ad}}}{k_BT}\right). \tag{2}$$

Here T is the temperature, E_ν^{ad} is the adatom formation energy, and $\mu_{\nu m}$ is the chemical potential of element ν at the step m. If the energy of mixing is unimportant (e.g., negligible tendency towards spinodal decomposition) then

$$\mu_{\nu m}=\tfrac{1}{2}M\varepsilon_m^2+M\varepsilon_m c_{-\nu m}\Delta\varepsilon_\nu+k_BT\ln c_{\nu m}. \tag{3}$$

Here M is the product of the atomic volume and an elastic constant (we assume that $M_\nu=M_{-\nu}=M$); $\Delta\varepsilon_\nu=(a_{-\nu}-a_\nu)/a_0$, where a_ν and a_0 are the lattice parameters of the ν component and of the substrate, respectively; the subscript $-\nu$ denotes the component other than ν; and the last term of Eq. (3) comes from the entropy of mixing. The strain at step m is given by[3,6]

$$\varepsilon_m=\overline{\varepsilon}+\beta h\sum_{n\neq m}\frac{\varepsilon_n}{x_m-x_n}, \tag{4}$$

where $\overline{\varepsilon}$ is the misfit strain, β is a ratio of elastic constants, and h the step height.

Consider a small periodic deviation ξ_ν from ideal flow of equally spaced steps:

$$x_{\nu m}=L(\overline{c}_\nu m+F_\nu t)+\mathrm{Re}(\xi_\nu e^{ikm}). \tag{5}$$

Here $x_{\nu m}=\int v_{\nu m}dt$, and $x_m=\Sigma_\nu x_{\nu m}$; $\overline{c}_\nu$ are the average compositions, equal to F_ν/F; $F=\Sigma_\nu F_\nu$ is the total flux; and the complex coefficient ξ_ν gives the amplitude and phase for a perturbation of wavelength λ, where $k=2\pi L/\lambda$, L being the average step spacing.

Substituting Eqs. (2)–(5) into (1), and linearizing in small amplitude ξ_ν, we obtain the following equation for the temporal evolution of the perturbation:

$$\dot{\xi}_\nu=(Q_\nu+iF_\nu\zeta)(\xi_\nu+\xi_{-\nu})-\frac{iR_\nu}{FL}(\overline{c}_\nu\dot{\xi}_{-\nu}-\overline{c}_{-\nu}\dot{\xi}_\nu), \tag{6}$$

where $\dot{\xi}_\nu=d\xi_\nu/dt$, $\zeta=\sin k$ and

$$Q_\nu=\overline{c}_\nu D_\nu\overline{\eta}_\nu\overline{\varepsilon}(\overline{\varepsilon}+\overline{c}_{-\nu}\Delta\varepsilon_\nu)\frac{2\beta hM}{k_BTL^3}\left(\pi k-\frac{1}{2}k^2\right)$$
$$\times(1-\cos k),$$

$$R_\nu=Q_\nu\frac{\Delta\varepsilon_\nu}{\overline{\varepsilon}}L\frac{2\pi-2k}{k(2\pi-k)}.$$

Here $\overline{\eta}_\nu$ is the adatom density for a pure and unstrained element ν.

Equation (6) has one trivial static solution plus the solution of interest, $\xi_\nu\propto r_\nu\exp(rt)$. Here $r=\Sigma_\nu r_\nu$ is the growth rate of the instability, with $r_\nu=Q_\nu+i\zeta F_\nu$.

The morphological perturbation $u_m=\mathrm{Re}(\xi e^{ikm})$ is the deviation of the step position x_m from that under ideal step flow, $L(m+Ft)$

$$u_m=Ce^{Qt}\cos(km+\phi), \tag{7}$$

where $Q=\Sigma_\nu Q_\nu$, C is the amplitude at $t=0$, and

$$\phi=\tan^{-1}\left[\frac{F\zeta\cos(Ft\zeta)+Q\sin(Ft\zeta)}{Q\cos(Ft\zeta)-F\zeta\sin(Ft\zeta)}\right].$$

The absolute decomposition $c_{\nu m}-\overline{c}_\nu$ grows in proportion to the morphological perturbation u_m, as

$$c_{\nu m}-\overline{c}_\nu=u_m\frac{2\beta hM\overline{G}}{k_BTFL^4}\left(\pi k-\frac{1}{2}k^2\right)(1-\cos k)$$
$$\times[\overline{c}_\nu\overline{c}_{-\nu}\overline{\varepsilon}^2(f_\nu^{\mathrm{mob}}+f_\nu^{\mathrm{size}}+f_\nu^{\mathrm{cross}})], \tag{8}$$

where

$$f_\nu^{\mathrm{mob}}=\frac{\Delta G_\nu}{\overline{G}}, \tag{9}$$

$$f_\nu^{\mathrm{size}}=\frac{\Delta\varepsilon_\nu}{\overline{\varepsilon}}, \tag{10}$$

$$f_\nu^{\mathrm{cross}}=\frac{\Delta\varepsilon_\nu}{\overline{\varepsilon}}\frac{\Delta G_\nu}{\overline{G}}\left(\frac{1}{2}-\overline{c}_\nu\right). \tag{11}$$

Here $G_\nu=\overline{\eta}_\nu D_\nu$, $\Delta G_\nu=G_\nu-G_{-\nu}$, and $\overline{G}=(G_\nu+G_{-\nu})/2$.

Equations (8)–(11) are our central results, describing the respective contributions of atomic size difference and mobility difference in the decomposition. However, before discussing the implications of these results, we repeat this analysis for the case of a continuum model of growth. The similarity of the two sets of results is important in establishing the generality of our conclusions.

In a continuum model, one considers a small-amplitude sinusoidal surface profile given by

$$H_\nu(x,t)=F_\nu t+\xi_\nu(t)\cos(kx), \tag{12}$$

where $H_\nu(x,t)$ is the ν-element contribution to the surface height $H(x,t)=\Sigma_\nu H_\nu(x,t)$. (Note that sometimes we use the same symbols in the step-flow and continuum analyses, to represent related but slightly different quantities with different dimensions.)

As in Refs. 2–4, the chemical potential in the surface for the component ν (relative to that for a flat surface) is

$$\mu_\nu(x)=\gamma V_0\kappa+\tfrac{1}{2}M\varepsilon^2+M\overline{c}_{-\nu}\varepsilon\Delta\varepsilon_\nu+k_BT\ln c_\nu, \tag{13}$$

where κ, ε, and c_ν are the curvature, the strain and the composition of the surface at point x, γ is the surface energy and V_0 is the atomic volume of the substrate.

Solving the elastic problem gives $\varepsilon=\overline{\varepsilon}[1-\beta\pi\xi k\cos(kx)]$. Thus for the surface profile given by Eq. (12) we will have

$$\mu_\nu(x)=\tfrac{1}{2}M\overline{\varepsilon^2}+k_BT\ln c_\nu + \xi k\cos(kx)(\gamma V_0k-\beta\pi M\overline{\varepsilon^2}-M\overline{c}_{-\nu}\Delta\varepsilon_\nu\overline{\varepsilon}). \tag{14}$$

The diffusion current on the surface is $J_\nu=-\overline{c}_\nu G_\nu(d\mu_\nu/dx)$,[4] where G_ν is the ν-element surface mobility. Thus the evolution of each component of the surface height is determined by the equation

$$\frac{dH_\nu}{dt}=\overline{c}_\nu G_\nu\frac{d^2\mu_\nu}{dx^2}+F_\nu. \tag{15}$$

Substituting Eqs. (12) and (14) into Eq. (15) we have the following equation for the amplitude of the perturbation: $\dot{\xi}_\nu=r_\nu(\xi_\nu+\xi_{-\nu})$, where

$$r_\nu=-k^3\overline{c}_\nu G_\nu(\gamma V_0k-\beta\pi M\overline{\varepsilon^2}-M\overline{c}_{-\nu}\Delta\varepsilon_\nu\overline{\varepsilon}). \tag{16}$$

As discussed above, we have taken $F_\nu \gg k^2G_\nu k_BT$. But our qualitative results should remain valid in the more general case.

As in the step-flow analysis there is one trivial static solution plus the solution of interest, which is given by $\xi_\nu \propto r_\nu\exp(rt)$. Now the morphological instability is

$$\Delta H(x,t)=H(x,t)-Ft=Cr\exp(rt)\cos(kx),$$

and the absolute decomposition is

$$c_\nu(x,t)-\overline{c}_\nu=\Delta H(x,t)\frac{k^3\overline{G}}{F}\overline{c}_\nu\overline{c}_{-\nu}\left\{\frac{\Delta G_\nu}{\overline{G}}(\beta\pi M\overline{\varepsilon^2}-\gamma kV_0) + M\overline{\varepsilon}\left[\Delta\varepsilon_\nu+\Delta\varepsilon_\nu\frac{\Delta G_\nu}{\overline{G}}\left(\frac{1}{2}-\overline{c}_\nu\right)\right]\right\}. \tag{17}$$

The instability is maximum for $k_{\max}$ defined by $dr/dk|_{k=k_{\max}}=0$, and

$$k_{\max}=\frac{3M\overline{\varepsilon}}{4\gamma V_0}\left[\beta\pi\overline{\varepsilon}+\frac{\overline{c}_{-\nu}\overline{c}_\nu\Delta G_\nu\Delta\varepsilon_\nu}{\overline{G}+\Delta G_\nu(\overline{c}_\nu-1/2)}\right]. \tag{18}$$

The absolute decomposition at $k_{\max}$ can be written as

$$c_\nu(x,t)|_{k=k_{\max}}-\overline{c}_\nu=\Delta H(x,t)\frac{k_{\max}^3M\overline{G}}{F}\overline{c}_\nu\overline{c}_{-\nu}\overline{\varepsilon}^2 \times(f_\nu^{\text{mob}}+f_\nu^{\text{size}}+f_\nu^{\text{cross}}), \tag{19}$$

where

$$f_\nu^{\text{mob}}=\frac{\beta\pi}{4}\frac{\Delta G_\nu}{\overline{G}}, \tag{20}$$

$$f_\nu^{\text{size}}=\frac{\Delta\varepsilon_\nu}{\overline{\varepsilon}}, \tag{21}$$

$$f_\nu^{\text{cross}}=\frac{\Delta\varepsilon_\nu}{\overline{\varepsilon}}\frac{\Delta G_\nu}{\overline{G}}\left[\left(\frac{1}{2}-\overline{c}_\nu\right)-\frac{3}{4}\left(\frac{\overline{c}_\nu\overline{c}_{-\nu}}{(\overline{G}/\Delta G_\nu)+(\overline{c}_\nu-1/2)}\right)\right]. \tag{22}$$

In Eqs. (8)–(11) and (19)–(22), the terms f_ν^{mob} and f_ν^{size} reflect decomposition caused by differences in adatom mobilities and sizes, respectively. There is also cross term, f_ν^{cross}.

Although we have considered two very different growth modes, the results for alloy decomposition are strikingly similar. There are also some differences, such as the factor of $\beta\pi/4$, which arises from the particular wavelength. Note that in the continuum case the phase of the modulation and decomposition are static on the surface, while for step flow these follow a traveling wave. Thus we expect that the resulting material will exhibit very different patterns of decomposition in the two cases.

In discussing the implications of these results we focus on the step-flow case, since this case is arguably the more relevant for semiconductor growth, and the formulas are slightly simpler. In Eqs. (8)–(11) we see that when ΔG_ν and $\Delta\varepsilon_\nu/\overline{\varepsilon}$ have the same sign, the two decomposition mechanisms add together. When the overall stress is compressive, this occurs if the larger atom is the more mobile. On the other hand, for an overall tensile stress, there will be enhanced decomposition when the larger atom is the less mobile. In the opposite case, when ΔG_ν and $\Delta\varepsilon_\nu/\overline{\varepsilon}$ have opposite signs, the two mechanisms cancel each other at least partially.

In the important case of an AB alloy grown on a substrate of B we can write

$$f_A^{\text{tot}}=\frac{2}{\overline{c}_A}\left(\frac{1}{1+G_B/G_A}\right), \tag{23}$$

where $f_\nu^{\text{tot}}=f_\nu^{\text{mob}}+f_\nu^{\text{size}}+f_\nu^{\text{cross}}$. Thus f_A^{tot} does not depend on the sizes of the atoms, but only on the composition and on the ratio of the adatom mobilities.

If the adatom mobilities of A and B are the same, then there is decomposition due only to the difference in atom sizes, and $f_A^{\text{tot}}=1/\overline{c}_A$. However, if $G_A \gg G_B$ [as for SiGe on Si (Ref. 7)] then

$$f_A^{\text{tot}}\approx\frac{2}{\overline{c}_A}(1-G_B/G_A)\rightarrow\frac{2}{\overline{c}_A}, \tag{24}$$

nearly double the decomposition in the case where $G_A=G_B$. Conversely, if $G_A \ll G_B$ (as for SiGe on Ge)

$$f_A^{\text{tot}}\approx\frac{2}{\overline{c}_A}\frac{G_A}{G_B}\rightarrow 0. \tag{25}$$

This is a remarkable result. In this case the effects of size and mobility cancel almost completely, giving a decomposition about G_A/G_B times smaller than it was in the $G_A \gg G_B$ case. More generally, whenever $\Delta\varepsilon_\nu$ and ΔG_ν have opposite signs the effects of atomic size and mobility will add for compressive strain but partially cancel for tensile strain.

We can incorporate all the effects of composition, strain, and atomic size and mobility difference in a term $g_\nu^{\text{tot}}=\overline{c}_\nu\overline{c}_{-\nu}\overline{\varepsilon}^2f_\nu^{\text{tot}}$, with corresponding definitions for g_ν^{mob}, g_ν^{size}, and g_ν^{cross}. Then the total decomposition is proportional to g_ν^{tot}, the morphological perturbation amplitude, and various other parameters. In Fig. 1 we show g_ν^{mob}, g_ν^{size},

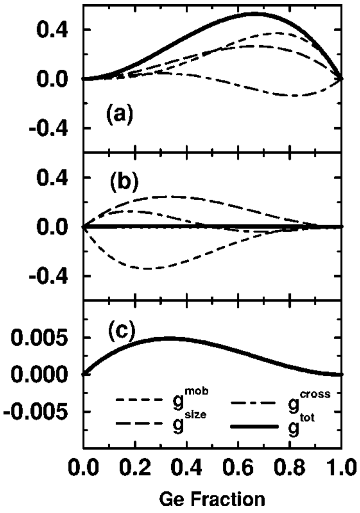

FIG. 1. Alloy decomposition factor g_ν^{tot}, and its three constituent terms g_ν^{mob}, g_ν^{size}, and g_ν^{cross}, versus average composition; from Eqs. (9)–(11). All curves are multiplied by 10^{-4}. Mobility ratios represent (a) SiGe on Si, and (b),(c) SiGe on Ge, as described in the text.

g_ν^{cross}, and g_ν^{tot} for an AB alloy grown on A or on B, as a function of the composition. For concreteness, we refer to $\mathrm{Si}_x\mathrm{Ge}_{1-x}$ grown on Si and Ge, taking a reasonable value of $G_{\mathrm{Ge}}=100G_{\mathrm{Si}}$ for the mobility difference.[7] (The actual value will vary greatly with temperature, because of the different activation energies.)

The results for SiGe on Si are displayed in Fig. 1(a). The contributions g_ν^{size} and g_ν^{mob} have the same sign, given enhanced decomposition over the entire range. In the limit of high Ge fraction the total decomposition has about the same magnitude as the decomposition due to the mobilities differences, because the size-driven decomposition is almost exactly canceled by the cross term. (g^{tot} differs from g^{mob} only by terms of order c_{Si} and of order $G_{\mathrm{Si}}/G_{\mathrm{Ge}}$.) But in the limit of high Si fraction, the effect of the mobility difference is simply to double the decomposition (via the cross-term), again up to corrections of order c_{Ge} and of order $G_{\mathrm{Si}}/G_{\mathrm{Ge}}$.

Figure 1(b) displays the results for SiGe on Ge. We see that g_ν^{size} and g_ν^{mob} have opposite signs for all compositions. Because of the large mobility difference, the cancellation is nearly complete. In Fig. 1(c) we display g_ν^{tot} for SiGe on Ge again, but with a different y scale to show clearly the form of the curve. According to Figs. 1(a) and 1(c), for any alloy AB grown in a B substrate, the decomposition will be maximum when the fraction of element B is 1/3.

For the continuum case, the same trends apply, but there is not in general such perfect cancellation of terms in the various limiting regimes. This is because the surface energy causes a more complex dependence on wavevector. The short-wavelength behavior is driven by surface energy, and decomposition results mainly from the mobility difference. The long-wavelength limit strongly resembles the step-flow case. (The step-flow regime would presumably share some of these complexities, if the steps were so close that we needed to include their short-range repulsion.)

When the size of the two atoms is nearly the same, as for $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{As}$ on a misfitting substrate, $|\Delta\varepsilon_\nu/\overline{\varepsilon}|\ll 1$ and $f_\nu^{\mathrm{tot}}\approx f_\nu^{\mathrm{mob}}$. In this case there may still be substantial decomposition, but it results almost entirely from the mobility difference.

Alloy decomposition due to different atom mobilities does not require strain at all, in contrast to the effect of atomic size difference. If the surface initially has some profile $\Delta H(x,t)$, during subsequent growth the profile will decay exponentially if there is no strain. From Eq. (17), the alloy decomposition is then

$$c_\nu(x,t)-\overline{c}_\nu=\frac{-k^4\gamma V_0}{F}\Delta H(x,t)\Delta G_\nu .$$

Thus the decomposition can be substantial, depending upon the mobility difference, growth rate, and wavelength. Such decomposition may be particularly important during growth on patterned substrates.

Other cases where decomposition would occur without strain include step bunching due to the diffusion barriers at steps,[5] and surface faceting. The former has recently been reported for unstrained $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{As}$.[8] However, these problems lie beyond the scope of our analysis.

Discussions with B. J. Spencer and P. Voorhees are gratefully acknowledged. P. V. was supported by the Brazilian agency FAPESP.

---

[1] T. Walther, C. J. Humphreys, and A. G. Cullis, Appl. Phys. Lett. **71**, 809 (1997); T. Okada, G. C. Weatherly, and D. W. McComb, J. Appl. Phys. **81**, 2185 (1997); J. Mirecki Millunchick, R. D. Twesten, S. R. Lee, D. M. Follstaedt, E. D. Jones, S. P. Ahrenkiel, Y. Zhang, H. M. Cheong, and A. Mascarenhas, MRS Bull. **22**(7), 38 (1997).

[2] J. E. Guyer and P. W. Voorhees, Phys. Rev. Lett. **74**, 4031 (1995); Phys. Rev. B **54**, 11 710 (1996).

[3] J. Tersoff, Phys. Rev. Lett. **77**, 2017 (1996).

[4] B. J. Spencer, P. W. Voorhees, and J. Tersoff (unpublished).

[5] G. Ehrlich and F. G. Hudda, J. Chem. Phys. **44**, 1039 (1966); R. L. Schwoebel, J. Appl. Phys. **40**, 614 (1969).

[6] J. Tersoff, Y. H. Phang, Z. Zhang, and M. G. Lagally, Phys. Rev. Lett. **75**, 2730 (1995); J. Tersoff, *ibid.* **74**, 4962 (1995).

[7] S. Jay Chey and David G. Cahill, in *Dynamics of Crystal Surfaces and Interfaces*, edited by P. M. Duxbury and T. J. Pence (Plenum Press, New York, 1997), p. 59. The activation energy of G for Si(001) is about 1.8 eV, and for Ge(001) it is about 20% smaller. Assuming comparable attempt frequencies (consistent with what little has been measured), the diffusion parameter G would be roughly one to three orders of magnitude larger for Ge than for Si over the typical range of growth temperatures.

[8] Th. Schäpers, A. Hartmann, A. Schwarz, H. Hardtdegen, M. Bongartz, Ch. Dieker, and H. Lüth, Appl. Surf. Sci. **123/124**, 687 (1998).