Growth of molybdenum on silicon: Structure and interface formation

J. M. Slaughter, Arye Shapiro,* Patrick A. Kearney, and Charles M. Falco
Department of Physics and Optical Sciences Center, University of Arizona, Tucson, Arizona 85721
(Received 8 February 1991)

Investigations of interface formation in the Mo-Si system were carried out by depositing Mo onto Si(100)-(2×1) and Si(111)-(7×7) surfaces in ultrahigh vacuum, followed by characterization with *in situ* reflection high-energy electron diffraction, low-energy electron diffraction, Auger-electron spectroscopy, and x-ray photoelectron spectroscopy (XPS). Continuous growth of multiple Mo coverages on a single Si wafer was accomplished with a technique involving a movable sample shutter. The formation of an amorphous interfacial silicide was observed at all substrate temperatures studied: ~50 °C, 100 °C, and 200 °C. However, the composition quickly becomes Mo rich as the deposition continues. The data are consistent with a composition profile that has an atomically abrupt transition between Si and amorphous $MoSi_x$, where $x=2$ for the first 4 Å and then decays with an error-function form with increasing overlayer thickness. The error-function interface-width parameter was found to be 10.0 Å at ~50 °C and 12.1 Å at 200 °C. Significant differences were seen between Auger intensities calculated by two standard methods: the derivative-amplitude method and the linear-background integrated-intensity method. We attribute these differences to peak-shape changes (due to the varying chemical environment in the interfacial region) that invalidate the use of the derivative method. The XPS measurements revealed shifts in the energies of the Mo $3d_{3/2}$ and Mo $3d_{5/2}$ lines due to the reaction with the Si substrate. The maximum peak shift was −0.4 eV and originated from the Mo nearest the Si substrate.

I. INTRODUCTION

Interest in understanding the physics and chemistry of the Mo-Si interface stems from applications in the semiconductor industry as well as in multilayer x-ray optics. In the semiconductor community, the narrow width of the metal-silicon interface, reproducible Schottky barrier heights, and low resistivities are some of the desirable properties that have stimulated interest in films made of this material combination.[1] For example, such films have been used as gates and contacts in very large scale integrated (VLSI) circuits.[2] For soft-x-ray optics, multilayer mirrors made of Mo as the absorber and Si as the spacer yield high reflectivity values for the wavelength region between 125 and 250 Å. A few examples of applications involving Mo/Si soft-x-ray mirrors are plasma diagnostics and x-ray microscopy at 182 Å,[3] astronomical telescopes for 133, 171, and 186 Å,[4] and x-ray laser cavity mirrors for 206–209 Å.[5] Although excellent Mo/Si mirrors have been produced, it has become apparent that the quality is extremely process dependent.[3,6] Thus, understanding the nucleation and growth mechanisms which affect the microscopic interface structure is very important to understanding and improving this technology.

Interfaces within Mo/Si multilayers have been studied extensively by several groups. High-resolution transmission electron microscope (HTEM) images of these multilayers have revealed amorphous interlayers between the amorphous Si layer and the polycrystalline Mo layer.[7–9] The amorphous interlayers are thicker at the Mo on Si interface than at the Si on Mo interface. This asymmetric interfacial reaction has been observed both in sputtered Mo/Si multilayers by Petford-Long *et al.*[7] and Holloway *et al.*,[8] as well as in UHV vapor deposited multilayers by Slaughter *et al.*[9] The fact that these interfacial layers have been observed in samples made under extremely different conditions implies that it may be an intrinsic property of the Mo-Si interface. However, no previous studies have investigated this possibility in detail.

In addition to these studies of interfaces within multilayer samples, in the past few years the deposition of Mo onto Si in ultrahigh vacuum has been studied by several groups.[10] In these studies, techniques which probe the chemistry of the surface, such as ultraviolet photoelectron spectroscopy (UPS), x-ray photoelectron spectroscopy (XPS), and Auger-electron spectroscopy (AES), were used to characterize the Mo-Si interface for various Mo coverages. However, the various research groups disagree over the sharpness and reactivity of the Mo-Si interface, as well as the growth mode. Nguyen and Cinti[11] and Balaska *et al.*[12] conclude that the interface is abrupt, with no reaction, while Rossi *et al.*[13] and Abbati *et al.*[14] report intermixing and silicide formation. As for the growth mode, Balaska *et al.*[12] claim a layer-by-layer mode, while Abbati *et al.*[14] conjecture that there is incomplete coverage by silicide islands. Meyerheim *et al.*[15] studied growth of Mo on Si(100) at room temperature and concluded from AES and extended x-ray absorption fine-structure (EXAFS) measurements that a chemical reaction takes place between the Mo and Si in a thin region at the interface.

In these previous UHV growth studies, data for different Mo coverages were obtained by successive cycles of sample preparation and Mo deposition, or by inter-

rupted deposition. Unfortunately, the former method has uncertain reproducibility, while the latter is not equivalent to continuous film deposition. It would be considerably better to deposit several Mo coverages during a single experiment onto a single well-characterized substrate. In addition, all of the previous workers who employed AES used the derivative method for determining peak intensities. This method can lead to systematic errors in the measured intensity when the peak shape changes due to changes in the chemical environment.[16] We have found that such peak shape changes do occur in the Mo-Si system and, in fact, lead to large errors in measured intensities when the derivative method is used.

In the present work careful attention was paid to avoiding the above-mentioned problems with sample reproducibility and Auger data analysis. Our technique for sample deposition involved sequentially moving a shutter in front of a large substrate during the deposition, thus exposing different regions of the substrate to the incident flux for various, controlled lengths of time. Our technique for producing such "moving shutter" samples is described in more detail in Sec. II A. Using such moving shutter samples, we then applied a variety of surface analysis techniques to the problem of understanding the formation of the Mo-Si interface on Si(111) and Si(100) at substrate temperatures from $\sim 50\,^{\circ}\mathrm{C}$ to $200\,^{\circ}\mathrm{C}$. Reflection high-energy electron diffraction (RHEED), low-energy electron diffraction (LEED), AES, and XPS were used *in situ* to study the interface for Mo coverages ranging from submonolayer to tens of angstroms.

II. EXPERIMENTAL DETAILS

The depositions were performed in a Perkin-Elmer 433-S molecular-beam-epitaxy (MBE) system[17] which was specifically designed for growth of Si and refractory metals. For the work reported here, this MBE machine was configured with two 40-cc electron-beam evaporators with a 99.9999% Si charge in one crucible and $\sim 99.99\%$ Mo pellets in the other. The growth geometry has the substrate facing down and the two evaporators facing up, with their centers offset with respect to the central axis of the wafer. The substrate temperature is measured with a calibrated thermocouple behind the 3-in.-diam Si wafer, and controlled with a feedback stabilized graphite heater. The base pressure is 5×10^{-11} torr, and the pressure during deposition is typically 1×10^{-9} torr, with H_2 making up 70% of that total. Deposition rates are held constant by active feedback from an Inficon Sentinel III deposition controller[18] which uses an electron-impact emission spectroscopy monitor to measure the flux. A Mo deposition rate of 3.0 Å/min was used for all samples discussed here. *Ex situ* Rutherford backscattering spectrometry (RBS) was used for absolute calibration of the deposition controller. In addition, RBS was used to measure the actual amount of Mo deposited with an accuracy of $\pm 3\%$. Since RBS measures the coverage (number of atoms per unit area) rather than the physical thickness, we use the density of bulk Mo (6.4×10^{22} at./cm^3) to calculate the coverage in units of angstroms of bulk material. Thus, the Mo "thicknesses" quoted in the present work should not be thought of as the actual thickness of the overlayer (which is affected by reaction with the Si substrate), but rather as the coverage expressed in terms of the equivalent thickness that an unreacted bulk Mo film would have. In Sec. III D we consider a model which allows for the formation of a graded alloy at the interface and which provides an estimate of the actual thickness of the overlayer.

RHEED and LEED analyses were performed with the samples in the growth chamber. The 10-keV RHEED apparatus is configured such that it can be used during growth. The LEED equipment is a reverse-view system[19] mounted on a flange at the top of the growth chamber of the MBE system. This geometry requires interrupting growth for LEED observations. Following RHEED and LEED analysis, wafers were transferred under UHV conditions to the analysis chamber ($P_{\mathrm{base}} < 3\times 10^{-11}$ torr), where AES and XPS measurements were performed.

The detection system used in this study, for AES and XPS, consists of a double-pass cylindrical-mirror electron-energy analyzer (DPCMA), and microcomputer-based control and data-acquisition electronics.[20] The microcomputer controls the analyzer power supply which in turn controls the voltages applied to the DPCMA. In contrast to most systems in which the derivative (dN/dE) spectrum is acquired for AES and the direct spectrum for XPS, this system acquires the direct spectra for both AES and XPS. A magnesium anode was used as the x-ray source, resulting in a strong Mg $K\alpha$ peak at 1253.6 eV.

Single-crystal Si wafers (p type) with (100) and (111) crystal orientations were used as substrates in these studies. No additional cleaning steps were performed prior to insertion into the MBE machine, where an *in situ* preparation was performed. This *in situ* substrate preparation procedure consists of two steps: first the wafer is heated to $850\,^{\circ}\mathrm{C}$ to desorb the native oxide and volatile contaminants, then a 100-Å-thick buffer layer of homoepitaxial Si is grown at $800\,^{\circ}\mathrm{C}$. The buffer layer buries any residual contaminants (primarily carbon) which survive the $850\,^{\circ}\mathrm{C}$ anneal. This procedure is sufficient for our purposes since we are concerned only with the structure and purity of the final Si surface, and not with the electronic properties of the buffer layer (which are affected by the buried impurities).

We performed LEED and RHEED studies before and after deposition of the epitaxial Si buffer layers in order to check that the Si surface was atomically smooth and crystalline. The LEED and RHEED patterns obtained after deposition of the buffer layers are sharp with a low-intensity background. We found the thickness of the Si buffer layer to be relatively unimportant. In one experiment we found that the RHEED pattern was much improved after deposition of a homoepitaxial layer of only 10 Å in thickness, and remained unchanged for layers 100- and 1000-Å thick. To confirm the chemical purity of the surface of the buffer layer, an experiment was performed in which Auger scans of a Si(100) surface were made immediately after preparation. No signs of C, N, or O were detected, indicating that there was less than 10^{-2} monolayer of each of these contaminants.

A. Moving shutter samples

To avoid ambiguities caused by typical run-to-run variations in deposition conditions, a "moving shutter" sample was devised for our growth studies. Several different film thicknesses can be deposited on the same wafer during one deposition run by sequentially moving the sample shutter to successively shadow parts of the wafer during growth. As many as 15 strips of different Mo thicknesses have been grown on one 3-in.-diam wafer, including calibration strips of completely exposed Si buffer layer and of "thick" Mo (50–150 Å). The advantages of this method include completion of an entire growth study in a single deposition run and internal calibration of AES and XPS intensities.

After being transferred to the analysis chamber, the samples were moved under the electron-energy analyzer, and AES measurements were performed at 0.28-mm increments along the diameter of the wafers. The Si and Mo AES measurements, made across the 5.0 cm of the wafer which were accessible to the DPCMA, took several hours to perform. Periodically, AES measurements were taken on the exposed Si buffer layer and on the thick Mo portion, to be used as time series data. These data were used to normalize the data taken on the thin Mo strips, to correct for any instrumental drifts or contamination occurring during the measurements. Following the AES measurements, XPS data were taken on selected areas of the sample.

In the course of this study, we found that the operation of the nude ionization gauge in the analysis chamber caused a significant increase in carbon and oxygen contamination on our freshly deposited Si and Mo surfaces. Presumably, this is due to generation of CO and CO_2 by the tungsten filaments in the gauges.[21] In order to avoid undue contamination, we turned off all ionization gauges prior to transfer of the samples from the growth chamber to the analysis chamber. The gauges remained off during the AES and XPS measurements.

As mentioned above, the direct spectra were acquired for both AES and XPS. Peak intensities were determined by integration of the data after subtracting a linear background. The normalized intensity used for our analysis is the ratio of the integrated intensity of the peak of interest to that of the pure material. The intensities of the pure materials were measured periodically during the data acquisition process by taking Si and Mo spectra on the bare substrate and on the thick Mo slice, respectively. As mentioned in Sec. I, previous workers[12,15] used the peak-to-peak amplitude of the derivative spectra as a measure of the Auger intensity, rather than the integrated intensity. For comparison with our integrated intensities, we also computed the derivative amplitude for each Auger peak, using a Savitzky-Galay convolution.[22] We determined from these comparisons that the derivative amplitude is not a good measure of the peak intensity, due to changes in the peak shape of the signal from the interface region. To understand this, consider, for example, a Si *LVV* Auger peak measured on a segment of the sample which has had 5 Å of Mo deposited on it. The Si peak from the interface region will be shifted in energy due to reaction with the Mo. However, since the signal from the interface region is added to the unshifted peak from the Si substrate immediately below the interface, the shape of the peak will change. This shape change will lead to a change in the derivative amplitude, and thus a systematic error in the intensity determination relative to the integrated intensity. The same effect occurs when the Mo peak is broadened due to a nonuniform reaction across the thickness of the overlayer. We observed such systematic differences between the normalized intensities determined by the integral and derivative methods. These differences were negligible for very low Mo coverage (<4 Å), but increased with increasing coverage. This behavior is expected, since very thin overlayers cannot have significant composition variations and the Si signal is dominated by the substrate. However, for thicker layers, the addition of signals from various depths alters the peak shape. In the worst case, after depositing ~ 15 Å Mo, the Si derivative intensities were greater than the integrated intensities by factors of almost $2\frac{1}{2}$.

B. Deposition profile

Each evaporation source in our MBE machine produces a particular deposition profile across the wafer. Since the actual coverage at some point (e.g., the thickest Mo slice) can be determined using RBS, knowledge of this deposition profile allows us to calculate the coverage at any point on the wafer from just one RBS measurement. We have measured the deposition profiles of Mo and Si over 3-in.-diam samples, using RBS, and have shown that they correspond to a cosine law flux distribution from our electron beam evaporators.[23,24] Knowledge of the Mo deposition profile, the total time that each slice was exposed to the Mo flux, and one RBS measurement of the Mo coverage on the thickest Mo slice, allowed us to accurately determine the Mo coverage at each point of interest.

III. RESULTS AND DISCUSSION

A. Electron diffraction studies

RHEED and LEED patterns convey structural information about the top two or three atomic layers of the sample under study. After preparing the Si buffer layers as described above, the Si(111) and Si(100) surfaces exhibited diffraction patterns characteristic of their well-known reconstructions; 7×7 and 2×1, respectively. In the case of Si(100), fractional-order spots corresponding to both (2×1) and (1×2) domains were visible. RHEED and LEED were used to determine that the correct growth conditions for the Si buffer layers had been achieved, prior to experiments involving Mo deposition. After preparing the Si buffer layer, Mo was deposited at various substrate temperatures.

Several electron diffraction experiments were performed on Mo which had been deposited onto substrates maintained at 100 °C. Observing these depositions with RHEED, a diffuse background was seen to rapidly rise with increasing Mo coverage, while first the fractional-order Si diffraction spots and then the (1×1) spots disap-

peared. For both Si(100) and Si(111), we found that after deposition of only ~2 Å of Mo, the LEED pattern characteristic of a crystalline surface was completely erased, and only a diffuse background remained. The corresponding RHEED pattern exhibited dim (1×1) spots with a high-intensity diffuse background. After significantly more Mo was deposited (> 10 Å) a ring pattern characteristic of polycrystalline Mo appeared. Similar results were obtained in experiments involving Mo growth on Si(111) at substrate temperatures of 50 °C and 200 °C.

In all cases, we found that the deposition of no more than 2 Å of Mo was sufficient to remove all traces of any Si LEED pattern. This fact indicates that, for these deposition conditions, the surface becomes amorphous upon deposition of Mo. However, since electron diffraction only yields structural information, we cannot determine from these measurements alone if the amorphous layer is pure amorphous Mo or a silicide formed by reaction with the Si substrate. The composition of the amorphous layer was determined with AES and XPS as described in the sections that follow. The fact that the Si pattern disappears after the deposition of only 2 Å of Mo suggests that the amorphous material is growing as a continuous layer. If there were island formation, the Si pattern would remain to higher Mo coverages, due to the exposed Si between the islands.

Samples made at $T=50$ °C, 100 °C and 200 °C on Si(111) and a sample made at $T=100$ °C on Si(100) were studied by HTEM. In all of these samples, the HTEM images revealed a flat and uniform amorphous interlayer, 10–15-Å thick, between the single-crystal Si and the polycrystalline Mo. These results are in agreement with the interpretation of the RHEED and LEED behavior given above.

B. Low-coverage growth studies

In this section we describe detailed studies of the initial stages of Mo growth on Si for $T=100$ °C and two different Si crystal orientations. By moving the shutter during growth, the samples were made with 14 slices, including bare silicon, a thick Mo slice, and 12 slices with Mo coverages below 4 Å (2.6×10^{15} at./cm^2). The deposition temperature of 100 °C was chosen because it is well below the temperature for formation of molybdenum silicides, while at the same time is high enough so that the substrate temperature could be held fixed throughout the entire deposition. One experiment was performed on Si(100)-(2×1) and another experiment on Si(111)-(7×7). Each experiment involved measurements of the Auger intensity of the Si *LVV* [92 eV (Ref. 25)] and the Mo *MNN* [186 eV (Ref. 25)] lines at many locations across the diameter of the sample as described above in Sec. II A.

Figure 1 shows a plot of normalized Si and Mo Auger intensities versus Mo coverage for the Si(100) substrate. For comparison, the calculated curves for alloy-layer growth of $MoSi_2$ on Si are also shown in this figure. The alloy-layer model is described below and in the Appendix. Although not shown in this figure, we find that the data for the experiment on Si(111) are indistinguishable from

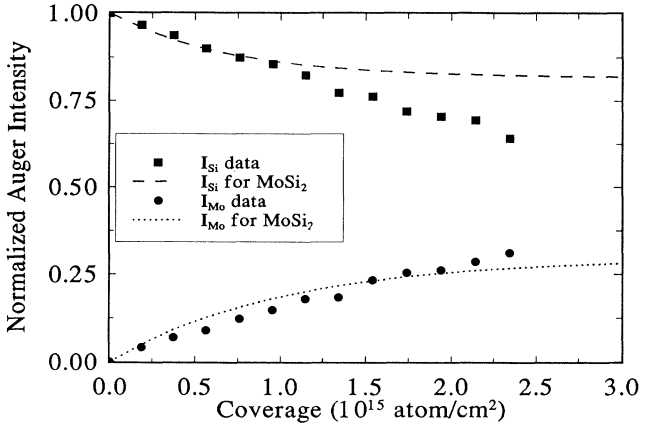

FIG. 1. Auger intensities for the study of Mo on Si(100)-(2×1) at $T=100$ °C. The data for Mo on Si(111)-(7×7) at $T=100$ °C (not shown) are indistinguishable from the data shown. The maximum coverage corresponds to approximately 4 Å of Mo. The dashed curves are calculated for alloy-layer growth of $MoSi_2$ on Si.

the data shown. Note that both the Si and the Mo curves exhibit no breaks, implying a non-layer-by-layer growth mode.[26] If we assume a simultaneous multilayer growth mode[27] or a statistical growth mode[28] of pure Mo on Si, the inelastic mean free path λ of the Si Auger electrons in Mo can be calculated from an exponential fit to the data. Assuming a simultaneous multilayer growth mode we find $\lambda\approx15$ Å, and for statistical growth $\lambda\approx13$ Å. Both values are a factor of 2 to 3 greater than that expected for the Si *LVV* Auger electrons.[29] That is, the Si signal does not decay fast enough with increasing Mo coverage to be consistent with either of these models of the growth mode. Either agglomeration of the Mo or interdiffusion of the Mo and Si could cause the observed behavior of the Auger intensity. However, since the rapidity with which the electron-diffraction patterns become diffuse rules out agglomeration, the large value of λ we determine leads us to conclude that the materials are intermixing at the interface to form a continuous amorphous layer.

It is useful to compare our experimental results to a model in which a uniform alloy-layer forms, with an abrupt interface, on the Si substrate. We refer to this model as "alloy-layer growth." Our calculations are based on the usual exponential attenuation with overlayer thickness, i.e., the fact that the number of electrons which are not inelastically scattered is exponentially attenuated through solids. The inelastic mean free path of an electron is calculated from the formula of Seah and Dench,[29] and the Auger-electron matrix factors are calculated from the backscattering factors of Ichimura and Shimizu.[30] Further details of this model are given in the Appendix. We use the intensity ratio I_{Mo}/I_{Si} as a function of Mo coverage to compare the model calculations to experiment. Figure 2 shows calculated intensity ratio curves for alloy-layer growth of pure Mo and for the

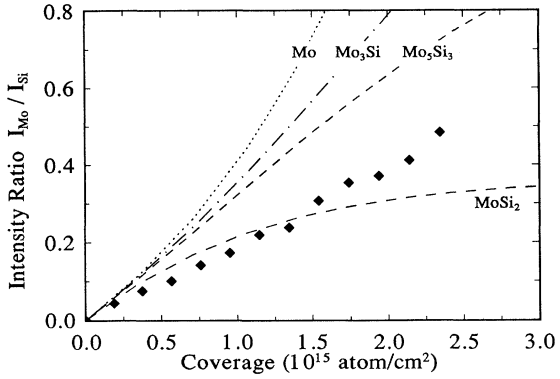

FIG. 2. Auger intensity ratio for Mo on Si(100)-(2×1) at $T=100$ °C. The data for Mo on Si(111)-(7×7) at $T=100$ °C (not shown) are indistinguishable from the data shown. The various dashed curves are the expected ratios for alloy-layer growth of the various silicides and pure Mo as indicated. The maximum coverage corresponds to approximately 4 Å of Mo.

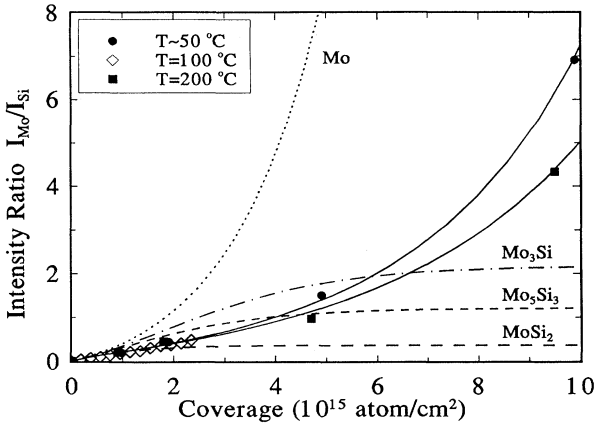

FIG. 3. Auger intensity ratios for growth studies of Mo on Si(111)-(7×7) at $T\approx 50$ °C, $=100$ °C, and $=200$ °C. The various broken curves correspond to alloy-layer growth calculations. The solid curves are gradient model fits corresponding to the composition profiles shown in Fig. 4. The maximum coverage corresponds to approximately 15 Å of Mo which, according to the gradient model calculations, resulted in 22-Å-thick overlayers.

three known molybdenum silicides compared to our data for growth on Si(100). The experimental data are initially quite close to the curve expected for $MoSi_2$, but clearly break away in the direction corresponding to a more Mo-rich alloy for coverages greater than 1.5×10^{15} at./cm^2 (2.3 Å Mo). These data imply that the stoichiometry of the silicide closest to the Si substrate is nearly that of $MoSi_2$, but it quickly becomes more Mo rich with increasing Mo coverage. The observation that the most Si-rich part of the interface has a stoichiometry similar to $MoSi_2$ is consistent with the phase diagram for Mo-Si,[31] which shows $MoSi_2+Si$ for the region with Si concentration greater than 67 at. %.

C. High-coverage growth studies

The results described above, for Mo coverages ≤ 4 Å, show that the early stages of Mo on Si growth involve considerable intermixing. To further determine the nature of the overlayer, two experiments were performed to study the growth over a greater range of Mo coverage. Growth studies were performed on Si(111)-(7×7) for coverages up to 15 Å, and at two different growth temperatures: $T=200$ °C and ≈ 50 °C.

Figure 3 shows the Auger intensity ratios versus Mo coverage for the two high-coverage experiments as well as the data shown in Fig. 2. Comparison of the experimental data to alloy-layer calculations for the various known silicides and pure Mo indicates that, as one might suspect, the surface becomes increasingly Mo rich with increasing Mo coverage. In the early stages of growth (<4 Å Mo), the ratio is independent of substrate temperature within our experimental uncertainty. However, as the deposition continues, the sample grown at the lowest temperature ($T\approx 50$ °C) becomes more Mo rich than the sample grown at the higher temperature ($T=200$ °C). The steep slopes of the curves at the highest coverages (~15 Å Mo) indicate that the surfaces of the overlayers are nearly pure Mo at these coverages. The solid curves are calculations based on the "gradient model" described below.

D. Discussion of Auger results

In order to determine the composition gradient of the overlayer, a model was constructed in which the overlayer is made up of successive slabs of $MoSi_x$, where x may be different for each slab. The composition of each slab is defined by the atomic fraction of Si, $f_{Si}(z)$, where z is the thickness of the overlayer up to and including that slab. We will refer to this as the "gradient model," details of which are given in the Appendix. Calculations using this gradient model yield the intensity ratio as a function of the total amount of Mo deposited (coverage), or as a function of the total thickness of the overlayer up to and including that slab.

Based on the results from our low- and high-coverage experiments, we propose the following model which describes the interfacial composition profile: an atomically abrupt transition between pure Si and a thin layer of amorphous $MoSi_x$, where $x=2$, followed by an error-function decay of the Si concentration with increasing overlayer thickness (and a corresponding increase in the Mo concentration). We define the transition between the pure Si and the silicide overlayer to be located in the $z=0$ plane; then the atomic fractions of Mo and Si, $f_{Mo}(z)$ and $f_{Si}(z)$, are described by the equations

$$f_{Si}(z)=\tfrac{2}{3},\quad 0<z\leq z_0 \tag{1a}$$

$$f_{Si}(z)=\frac{2}{3}\left[1-\mathrm{erf}\left[\frac{z-z_0}{\sqrt{2}\sigma}\right]\right],\quad z>z_0\ , \tag{1b}$$

where

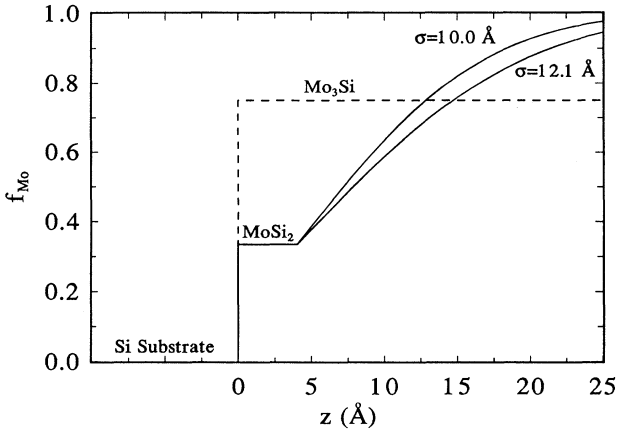

FIG. 4. The solid curves are the composition profiles which resulted in the best gradient model fits to the experimentally determined intensity ratios for Mo on Si(111). The error-function form of these curves is described in the text. The $\sigma=10.0$ and 12.1 Å curves correspond to the $T\approx 50\,^{\circ}\mathrm{C}$ and $=200\,^{\circ}\mathrm{C}$ samples, respectively. The composition profile for alloy-layer growth of Mo_3Si is shown for comparison.

$$\mathrm{erf}(\xi)=\frac{2}{\sqrt{\pi}}\int_0^{\xi} e^{-t^2}dt \tag{1c}$$

and

$$f_{\mathrm{Mo}}(z)=1-f_{\mathrm{Si}}(z)\ . \tag{1d}$$

Physically, this form represents diffusion at an interface in which f_{Si} is fixed at $\frac{2}{3}$ for $z\leq z_0$ and the Si diffuses into the Mo for $z>z_0$, over a characteristic distance σ. The inclusion of a region of constant concentration from $z=0$ to z_0 is based on our conclusion from the low-coverage experiments that the stoichiometry of the thin amorphous silicide layer nearest the Si substrate is approximately that of $MoSi_2$. From fits to our data, we find σ to be 10.0 Å for $T\approx 50\,^{\circ}\mathrm{C}$ and 12.1 Å for $T=200\,^{\circ}\mathrm{C}$ with $z_0=4$ Å in both cases. The calculated ratios are shown as solid curves in Fig. 3. Figure 4 shows these best-fit composition profiles as f_{Mo} versus z. The composition profile for alloy-layer growth of Mo_3Si is also shown in Fig. 4 for comparison. The fact that the calculated ratio curves are in good agreement with the experimentally determined ratios is not proof that the composition profiles are of this error-function form; however, the actual profile is certainly very similar. In addition, the parameters determined from these fits provide a basis for comparison of graded interfaces formed under various conditions, in a form that has a physical basis.

E. XPS studies

Although the intensities of XPS peaks can, in principle, be analyzed in the same way as Auger intensities, we found that only the Auger data had sufficient surface sensitivity for the present growth mode studies. The kinetic energy of the photoelectrons studied by XPS was much higher than that of the Auger electrons: 88 and 182 eV for Auger; 1155 and 1026 eV for XPS. The higher kinetic energy results in a longer inelastic mean free path λ by a factor of 2 to 3, which in turn causes the intensities to be much less sensitive to the composition of the sample surface. However, the binding energy shifts determined by XPS yield useful information about the interface. The Si $2p$, Mo $3d_{3/2}$, and Mo $3d_{5/2}$ lines corresponding to binding energies 99.15, 224.7, and 227.7 eV were used in these studies.

For the photoelectron energies of interest, the Seah and Dench formula[29] yields $\lambda\approx 17$ Å in Mo and its silicides. However, in order to take into account the geometry of our apparatus, we must integrate the cosine of the emission angle over the aperture of our analyzer to find the effective inelastic mean free path λ_{eff}.[32] By positioning a 90° acceptance angle aperture inside our DPCMA, we can vary λ_{eff} between 16 Å (0.944λ) and 5.7 Å (0.338λ). A fully open (360°) aperture has $\lambda_{\mathrm{eff}}=11$ Å (0.641λ), and also results in maximum signal strength. Although the major contribution to the XPS signal is from a depth of λ_{eff} into the sample, there is a significant contribution from material as deep as $\sim 3\lambda_{\mathrm{eff}}$ below the surface. Thus, for the thinner regions on the $T\approx 50\,^{\circ}\mathrm{C}$ and $=200\,^{\circ}\mathrm{C}$ samples, the Mo signal comes from the whole overlayer while the Si signal is mainly from the substrate with only a small contribution from the Si in the overlayer. Although for the thick regions (10- to 20-Å thick) both signals originate mainly from the top of the overlayer, there is a significant contribution from the material at the bottom of the layer, and from the substrate in the case of the Si signal. These effects allow us to use spectra from different thickness slices on the sample to study variations in the binding energies across the width of the interface. In these studies we found the peak position of Si difficult to interpret due to the strong substrate signal, but interpretation of the Mo peak to be straight forward since all of the Mo is in the overlayer. As an internal calibration, Mo and Si peak shifts were measured with respect to the peak position for the thick (150-Å) Mo slice deposits on each sample and the bare Si slice, respectively.

At both $T\approx 50\,^{\circ}\mathrm{C}$ and $=200\,^{\circ}\mathrm{C}$, the shift in the Mo $3d$ peaks is the same (-0.4 ± 0.05 eV) for low coverages, indicating the same type of initial reaction. However, at 50 °C the binding energy moves toward its bulk value more quickly with increasing coverage than at 200 °C. After the deposition of 15 Å of Mo, the shifts were reduced to -0.20 eV at $T\approx 50\,^{\circ}\mathrm{C}$ and -0.30 eV at $T=200\,^{\circ}\mathrm{C}$.

The maximum peak shift -0.4 eV originated from the Mo nearest to the Si substrate. Peak shifts of -0.2 eV (Ref. 33) to -0.32 eV (Ref. 34) have been reported for $MoSi_2$. Thus the binding energy of the Mo nearest to the substrate has a peak shift in the same direction as that for $MoSi_2$ and of similar magnitude. The difference between our value and those reported for $MoSi_2$ could be related to the fact that our interfacial silicide is amorphous and the shifts measured in Refs. 33 and 34 were for crystalline $MoSi_2$.

As mentioned above, changing the setting of the 90°

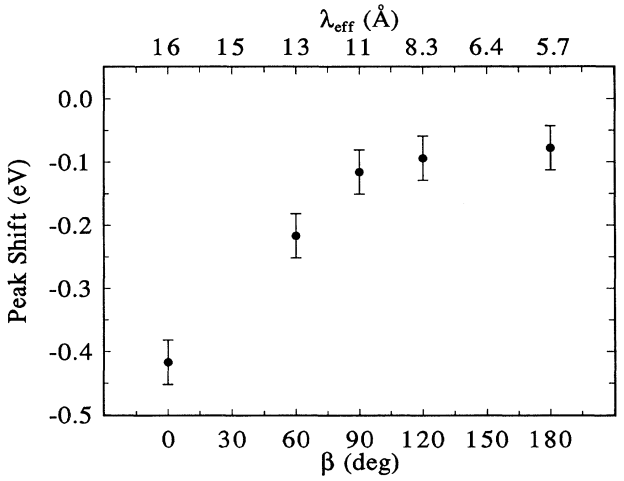

FIG. 5. Mo 3*d* peak shifts from angle-resolved XPS for Mo on Si(111)-(7×7) at $T=50$ °C. Shifts for $3d_{3/2}$ and $3d_{5/2}$ are identical within experimental accuracy. This portion of the sample had 15.5 Å of Mo deposited which, according to the gradient model calculations, resulted in a 22-Åthick overlayer.

aperture in our DPCMA allows us to vary the depth probed by about a factor of 2.8. Figure 5 shows the variation in the position of the Mo peaks for various positions of the 90° aperture for the $T\approx 50$ °C sample after the deposition of 15.5-Å Mo (corresponding to a 22-Å-thick overlayer according to the gradient model calculation). The λ_{eff} corresponding to each angle is shown on the abscissa. As seen in Fig. 5, there is a variation of 0.35 eV in the peak positions as λ_{eff} is varied from 5.7 to 16 Å, indicating a gradient in the Mo bonding across the thickness of the overlayer. In contrast, at $T=200$ °C the variation was less than 0.1 eV, with the peaks remaining shifted by -0.3 eV even for $\lambda_{\text{eff}}=5.7$ Å, indicating that the Mo bonding is more uniform in this case.

Peak shifts of -0.3 eV were observed in the Si spectra after the deposition of 15.5-Å Mo (22-Å-thick overlayer from gradient model calculations). These spectra have contributions from Si at various depths but since $\lambda_{\text{eff}}=11$ Å they are dominated by the signal from Si in the region with decaying Si concentration. Unfortunately, it was impossible to determine the shift for the Si in the thinnest overlayers since in this case the overlayer signal is overwhelmed by the substrate signal. The shifts we observed were, in all cases, toward a lower binding energy. A negative shift has also been observed by Wagner *et al.* for $MoSi_x$ (-0.04 eV).[34] We conclude that the XPS peak shifts observed in the present work, together with the AES analysis, paint a consistent picture of a graded interface beginning with $MoSi_x$, with $x\approx 2$, and becoming increasingly Mo rich over approximately 25 Å.

IV. CONCLUSIONS

We have performed detailed studies of the interface which forms upon deposition of thin layers of Mo on Si. These studies were performed at $T=100$ °C for Si(100), and at $T\approx 50$ °C, 100 °C, and 200 °C for Si(111). In all cases, an amorphous silicide layer is formed which quickly becomes more Mo rich with increasing Mo coverage. The data are consistent with a composition profile which has an atomically abrupt transition between Si and amorphous $MoSi_x$, where $x=2$ for the first 4 Å and then decays with an error-function from with increasing overlayer thickness. The error-function interface-width parameter was found to be 10.0 Å at ~50 °C and 12.1 Å at 200 °C. Our observation that the most Si-rich part of the interface has the stoichiometry of $MoSi_2$ is consistent with the phase diagram for Mo-Si, which shows $MoSi_2+Si$ for compositions with a Si concentration greater than 67 at. %. The $T=100$ °C studies, which focused on Mo coverages below 4 Å (2.6×10^{15} at./cm^2), showed that the growth mode is independent of substrate orientation.

Significant differences were seen between Auger intensities calculated by two standard methods—the derivative-amplitude method and the linear-background integrated-intensity method. We attribute these differences to peak-shape changes (due to the varying chemical environment in the interfacial region) which invalidate the use of the derivative method.

The principal motivation of the present work was to determine the interfacial structure of Mo-Si in order to understand its effect on the performance of multilayer x-ray optical structures. The formation of the interfacial silicide will only slightly degrade the performance of multilayer x-ray mirrors in the wavelength region for which they are most used, 125–250 Å. The problem of silicide formation is not serious in this region because the interface width determined from our studies is much smaller than the wavelengths of interest. However, the problem becomes serious if such mirrors are desired for much shorter wavelengths. Even though the theoretical reflectivity of Mo/Si mirrors is reasonably high below 30 Å,[36] formation of the interfacial silicides observed in the present work, at deposition temperatures as low as 50 °C, would make such mirrors useless. A method of inhibiting the reaction, or use of a different combination of materials, is required for the short-wavelength region.

ACKNOWLEDGMENTS

The authors thank John A. Leavitt and Laurence C. McIntyre, Jr. for the RBS measurements, Dean W. Schulze for the deposition profile calculations, and Charles Hills for HTEM characterization. This work is supported by the Air Force Office of Scientific Research under Contract No. AFOSR-90-0140 and the Joint Services Optics Program under Contract No. F-49620-88-C-0009.

APPENDIX: AES/XPS OF THIN ALLOY LAYERS

In this appendix we first derive expressions for AES and XPS intensities for the alloy-layer growth model. This model assumes that the deposition of element *A* onto a substrate of element *B* results in a homogeneous binary alloy AB_x of uniform thickness *d* on top of *B*. The interface between the alloy overlayer and the *B* sub-

strate is assumed to be perfectly sharp. We then use these results to find AES/XPS intensities for the gradient model. This model assumes that the overlayer is composed of a binary alloy with a composition gradient in the direction normal to the surface. The alloy is again AB_x, but in the gradient model x continuously varies from its value at the interface with the substrate to zero far from the interface. The composition gradient is modeled by taking successive slabs of AB_x, where x may be different for each slab. The gradient model is intended to approximate interfaces which are compositionally graded due to diffusion and/or reaction of the substrate and overlayer. Both models make use of the fact that the number of electrons which are not inelastically scattered is exponentially attenuated through solids. We neglect the effect of surface roughness on the intensities since all measurements in our moving shutter experiments are made on separate segments of the same substrate, and therefore have similar roughness. In addition, by comparing intensity ratios rather than the individual Si or Mo intensities the roughness dependence is canceled out. The electron backscattering factors $R_\mu(E)$, which are functions of material μ and kinetic energy E, are included in the intensity expressions for generality. The backscattering factors of Ichimura and Shimizu[30] were used for our AES calculations. For XPS calculations $R_\mu(E)\equiv 1$.

The intensity of Auger electrons or photoelectrons from an infinitely thick slab of a homogeneous binary alloy AB_x can be written[37]

$$I_A = I_A^\infty \frac{R_M(E_A)\lambda_M(E_A)N_A}{R_A(E_A)\lambda_A(E_A)N_A^\infty} , \tag{A1}$$

where $R_\mu(E)$, $\lambda_\mu(E)$, and N_μ represent the backscattering factor, inelastic mean free path, and atomic number density for material μ at energy E. The subscripts A and M denote material A and the alloy matrix AB_x, respectively, while the superscript ∞ denotes a parameter for bulk (infinitely thick) material. Equation (A1) can be written in terms of the mass density of the materials ρ_μ, the mean atomic weight A_μ, and the molar fraction of material A in the alloy, $1/(1+x)$, to yield

$$\frac{I_A}{I_A^\infty} = \frac{R_M(E_A)\lambda_M(E_A)\rho_M A_A}{R_A(E_A)\lambda_A(E_A)\rho_A A_M}\left[\frac{1}{1+x}\right] . \tag{A2}$$

The intensity $I_A(d)$ from material A in a slab of alloy with thickness d is simply the intensity of the infinitely thick slab I_A, less the contribution from the material below a depth d. The contribution from below a depth d is the intensity for an infinite slab attenuated by the alloy above, i.e., $I_A\exp[-d/\lambda_M(E_A)\cos\theta]$, where θ is the emission angle. We must also include a backscattering correction factor $R_B(E_A)/R_M(E_A)$ to account for the fact that we are considering a thin film of the alloy on a thick substrate of B.[38] The resulting expression for the normalized intensity from material A for the alloy-layer growth model is given by

$$\frac{I_A(d)}{I_A^\infty} = \left[\frac{I_A}{I_A^\infty}\right](1-e^{-d/\lambda_M(E_A)\cos\theta})\left[\frac{R_B(E_A)}{R_M(E_A)}\right] . \tag{A3}$$

The intensity from material B is the sum of contributions from the overlayer and from the substrate. The contribution from the overlayer is given by Eq. (A3) with B substituting for A. The intensity from the substrate is attenuated by the overlayer, yielding a contribution $I_B\exp[-d/\lambda_M(E_B)\cos\theta]$. Thus, the normalized intensity of B for the alloy-layer growth model is

$$\frac{I_B^{\text{total}}(d)}{I_B^\infty} = \frac{I_B(d)}{I_B^\infty} + e^{-d/\lambda_M(E_B)\cos\theta} , \tag{A4}$$

where the first term is given by (A3) with B substituting for A.

For the energy range of interest, the inelastic mean free path is given approximately by

$$\lambda_\mu(E) = 0.41 a_\mu^{1.5} E_\mu^{0.5} , \tag{A5}$$

where $\lambda_\mu(E)$ and a_μ are in nanometers and E_μ is in eV.[29,37] The average atom size a_μ is derived from the relation $\rho_\mu N a_\mu^3 = A_u$, where N is Avogadro's number. Equations (A1) through (A5) define the alloy-layer growth model.

We now consider the gradient model. This model assumes that the overlayer is composed of a binary alloy with a composition gradient in the direction normal to the surface. The alloy is AB_x with x continuously varying from its value at the interface with the substrate, to zero far from the interface. The composition gradient is modeled by taking successive slabs of AB_x, where x may be different for each slab. We wish to find expressions which describe the intensities after the deposition of each slab in order to calculate intensities as a function of coverage of material A. The nth slab is described by its thickness, mass density, and mean atomic mass: d_n, ρ_n, and A_n, respectively. We make the assumption that these parameters do not change as the deposition continues. The total thickness of the overlayer after the deposition of the nth layer is D_n. If we begin with the slab in contact with the substrate ($n=1$), the intensities are given by the alloy-layer growth expressions above. For the other slabs we can write a recursion formula by using the same arguments used to construct Eq. (A4). The intensity from material A after the deposition of n layers $I_A^G(D_n)$ is the sum of contribution from the material below layer n is $I_A^G(D_{n-1})$. However, it is attenuated by layer n, yielding a contribution

$$I_A^G(D_{n-1})\exp[-d_n/\lambda_n(E_A)\cos\theta] .$$

Thus, the normalized intensity of A for the gradient model is determined by the recursion formula

$$I_A^G(D_n) = I_A(d_n) + I_A^G(D_{n-1})e^{-d_n/\lambda_n(E_A)\cos\theta} . \tag{A6}$$

The intensity of material B for $n=1$ is given by (A4). For $n>1$ we construct a recursion formula by the same arguments used to derive (7). The result is

$$I_B^G(D_n) = I_B^G(D_{n-1})e^{-d_n/\lambda_n(E_B)\cos\theta} + I_B(d_n) . \tag{A7}$$

In order to apply Eqs. (A6) and (A7), the mass density of the alloy as a function of composition is needed. We

determined an empirical formula which approximates the densities of the known molybdenum silicides and pure molybdenum to within 1%. We take the general alloy to be $M=\mathrm{Mo}_y\mathrm{Si}_x$ and write the expression for density as the product of the weighted average of the densities of the components and a correction factor which is linear in the atomic fraction of Si, $x/(x+y)$ as shown in Eq. (A8).

$$\rho_M=\left[\frac{y\rho_{\mathrm{Mo}}+x\rho_{\mathrm{Si}}}{x+y}\right]\left[1+0.317\left[\frac{x}{x+y}\right]\right]. \quad \text{(A8)}$$

During the experiment it is necessary to measure the intensities from the bulk materials I_μ^∞ in order to normalize the data for comparison to the equations above. A region of bare Si was left on our samples for the determination of I_{Si}^∞, and a region of thick Mo was deposited for I_{Mo}^∞. Since there is some contribution to the Auger signal from the material in the top $3\lambda_{\mathrm{eff}}$ of the sample, clearly the thick Mo layer must have a thickness greater than $3\lambda_{\mathrm{eff}}$ in order to be thick enough for determination of I_{Mo}^∞. However, one must also consider the effect of backscattered electrons on the Auger signal from the overlayer. In general, the Auger signal has two components: that generated by the interaction of the primary beam with the sample surface, and that generated by electrons backscattered within the sample which return to the surface. For the case of an interface between an overlayer of A on a substrate B, the corrected intensity I_A is approximately given by

$$\frac{I_A}{I_A^\infty}=\frac{I_A(d)}{I_A^\infty}\left[1+\left[\frac{R_B(E_A)}{R_A(E_A)}-1\right]e^{-d/l_b}\right], \quad \text{(A9)}$$

where $I_A(d)$ is the uncorrected intensity from an overlayer of thickness d and l_b is a characteristic length for backscattering.[39] For a 3-keV primary beam incident on Mo, l_b has been found to be approximately 20 monolayers, or ~ 50 Å.[40] In our apparatus the electron beam impinges upon the sample at an angle of 30° from the normal, resulting in an effective l_b of ~ 43 Å. By applying Eq. (A9) we find that the intensities measured on the 50-Å-thick slices require correction by nearly 6%, since in this case the correction factor is 0.944. However, for $d_{\mathrm{Mo}}=150$ Å the correction factor is 0.995, so that the intensities measured on the 150-Å-thick Mo slices provide accurate measurements of I_{Mo}^∞.

*Present address: IBM East Fishkill Facility, Building 330-74A, Route 52, Hopewell Junction, NY 12533.

[1]S. P. Murarka, *Silicides for VLSI Applications* (Academic, New York, 1983).

[2]Raj N. Singh on pp. 77–92 and Anant G. Sabnis on pp. 228 and 229 of *VLSI Electronics: Microstructure Science,* edited by Norman G. Einspruch, Simon S. Cohen, and Gennady Sh. Gildenblat (Academic, Orlando, 1987), Vol. 15.

[3]J. M. Slaughter, M. K. Burkland, P. A. Kearney, A. R. Lampis, Z. Milanovic, D. W. Schulze, J. R. Roberts, J. Kerner, E. B. Saloman, and C. M. Falco, Proc. SPIE **1160**, 235 (1989).

[4]Barham W. Smith, Jeffrey J. Bloch, and Diane Roussel-Dupré, Opt. Eng. **29**, 592 (1990).

[5]N. M. Ceglio, D. G. Stearns, D. P. Gaines, A. M. Hawryluk, and J. E. Trebes, Opt. Lett. **13**, 108 (1988).

[6]S. Ogura, M. Niibe, Y. Watanabe, M. Hayashida and T. Iizuka, Proc. SPIE **984**, 140 (1988).

[7]A. Petford-Long, M. B. Stearns, C. H. Chang, D. G. Stearns, N. M. Ceglio, and A. M. Hawryluk, J. Appl. Phys. **61**, 1422 (1987).

[8]Karen Holloway, Khiem Ba Do, and Robert Sinclair, J. Appl. Phys. **65**, 474 (1989).

[9]J. M. Slaughter, Patrick A. Kearney, Dean W. Schulze, and Charles M. Falco, Proc. SPIE **1343**, 73 (1991).

[10]For a detailed review of pervious work on the Mo-Si system see Arye Shapiro, Ph.D. dissertation, University of Arizona, 1989.

[11]T. T. A. Nguyen and R. C. Cinti, J. Phys. (Paris) Colloq. **45**, C5-435 (1984).

[12]H. Balaska, R. C. Cinti, T. T. A. Nguyen, and J. Derrien, Surf. Sci. **168**, 225 (1986).

[13]G. Rossi, I. Abbati, L. Braicovich, I. Lindau, and W. E. Spicer, J. Vac. Sci. Technol. **21**, 617 (1982); G. Rossi, I. Abbati, L. Braicovich, I. Lindau, W. E. Spicer, U. Del Pennino, and S. Nannarone, Physica B+C **117&118B**, 795 (1983).

[14]I. Abbati, L. Braicovich, B. De Michelis, A. Fasana, E. Puppin, and A. Rizzi, Solid State Commun. **52**, 731 (1984); I. Abbati, L. Braicovich, B. De Michelis, A. Fasana, and A. Rizzi, Surf. Sci. **177**, L901 (1986).

[15]H. L. Meyerheim, U. Döbler, A. Puschmann, and K. Baberschke, Phys. Rev. B **41**, 5871 (1990).

[16]M. P. Seah in *Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy,* edited by D. Briggs and M. P. Seah (Wiley, New York, 1983), p. 193.

[17]Perkin Elmer, Physical Electronics Division, Eden Prairie, MN.

[18]Inficon Leybold-Heraeus, Inc., East Syracuse, NY.

[19]Model RVL 8-120, Princeton Research Instruments, Princeton, NJ.

[20]Physical Electronics 3027 PC DPCMA-based ESCA/AES System, Perkin Elmer Physical Electronics, Eden Prairie, MN.

[21]John F. O'Hanlon, *A User's Guide to Vacuum Technology,* 2nd ed. (Wiley, New York, 1989), pp. 97 and 140.

[22]A. Savitzky and M. J. E. Golay, Anal. Chem. **36**, 1627 (1964) with corrections by J. Steinier, Y. Termonia, and J. Deltour, Anal. Chem. **44**, 1906 (1972).

[23]L. I. Maissel and R. Glang, *Handbook of Thin Film Technology* (McGraw-Hill, New York, 1970), pp. 1–34.

[24]D. W. Schulze, J. M. Slaughter, and C. M. Falco, Proc. SPIE **984**, 75 (1988).

[25]This energy is the nominal position of the negative extreme in the derivative Auger spectrum, typically used for elemental identification. See, for example, Lawrence E. Davis, Noel C. MacDonald, Paul W. Palmberg, Gerald E. Riach, and Roland E. Weber, *Handbook of Auger Electron Spectroscopy,* 2nd ed. (Physical Electronics Division, Perkin-Elmer Corp., Eden

Prairie, MN, 1978).
[26]T. E. Gallon Surf. Sci. **17**, 486 (1969).
[27]G. E. Rhead, M.-G. Barthés, and C. Argile, Thin Solid Films **82**, 201 (1981).
[28]L. C. Feldman and J. W. Mayer, *Fundamentals of Thin Film and Surface Analysis* (Elsevier, New York, 1986), Chap. 6.
[29]M. P. Seah and W. A. Dench, Surf. Interface Anal. **1**, 2 (1979).
[30]S. Ichimura and R. Shimizu, Surf. Sci. **112**, 386 (1981).
[31]*The Handbook of Binary Phase Diagrams,* edited by Joseph O. Accrocco (Genium Publishing, Schenectady, NY, 1984), Vol. 4.
[32]Ours is the same CMA geometry as described by S. Hofmann in *Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy* (Ref. 16), pp. 165 and 166.
[33]W. A. Brainard and D. R. Wheeler, J. Vac. Sci. Technol. **15**, 1801 (1978).
[34]C. D. Wagner, D. E. Passoja, H. F. Hillery, T. G. Kinisky, H. A. Six, W. T. Jansen, and J. A. Taylor, J. Vac. Sci. Technol. **21**, 933 (1982).
[35]C. D. Wagner, H. A. Six, W. T. Jansen, and J. A. Taylor, Appl. Surf. Sci. **9**, 203 (1981).
[36]Patrick A. Kearney, J. M. Slaughter, and Charles M. Falco, Proc. SPIE **1343**, 25 (1991).
[37]M. P. Seah, in *Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy* (Ref. 16), pp. 186 and 187.
[38]M. P. Seah, in *Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy* (Ref. 16), p. 211.
[39]S. Hofmann, Microchim. Acta Suppl. **8**, 71 (1979).
[40]Manfred Schrader, Mikrochim. Acta Suppl. **8**, 377 (1979) and analysis of Schrader data in S. Hofmann, *ibid.* **8**, 71 (1979).