Structure of Si(100)-(2×1) Surface Using UHV Transmission Electron Diffraction

Ganesh Jayaram, P. Xu, and L. D. Marks
Department of Materials Science and Engineering, Northwestern University, Evanston, Illinois 60208
(Received 30 July 1993)

Details of the atomic structure for the Si(100)-(2×1) surface using UHV transmission electron diffraction are described. Reliability factor minimizations of the dynamical diffraction intensities establish conclusively the asymmetry in the structure. Fits performed using multilayer subsurface relaxations to match analytical strain solutions demonstrate the existence of long range subsurface strain fields extending up to six layers into the bulk.

PACS numbers: 61.16.Bg, 61.14.Rq, 68.35.Bs

Although the (2×1) dimer structure for the clean Si(100) surface was proposed by Schlier and Farnsworth [1] using low energy electron diffraction (LEED) about thirty years ago, there still remains a great degree of controversy about the details of the surface. Total-energy calculations yield contradictory results: While some empirical tight-binding calculations [2] and total-energy minimizations [3,4] predict symmetric dimers to be unstable with respect to the asymmetric and/or buckled dimers, core-level shift calculations incorporating spin effects [5] and other total-energy and force calculations [6] suggest that symmetric dimers are more stable. The first scanning tunneling microscopy (STM) studies [7,8] indicated that on defect-free (2×1) surfaces only symmetric dimers were present with buckled dimers in local regions with $c(4\times2)$ and $p(2\times2)$ symmetry near surface defects; however, recent low temperature measurements [9] suggest that the dimers are asymmetric and time averaged at room temperature by thermal excitations to give a symmetric appearance. Grazing incidence x-ray diffraction (XRD) [10] and most recently optical spectroscopy [11] studies also show a preference for an asymmetric structure. Apart from a simple buckled structure (tilt out of the surface plane), an asymmetric structure with an in-plane dimer axis twist has also been suggested by LEED [12]. Energy minimizations [13] also indicate that long range subsurface distortions are associated with the Si(100) surface reconstructions; however, there still exists some controversy over the extent.

In the last decade or so, plan view transmission electron diffraction (TED) has rapidly emerged as a powerful technique for solving surface structures at the atomic level [14–16]. This technique is highly sensitive to atomic displacements perpendicular to the beam. Smaller probe sizes (than conventional XRD) and parallel detection (unlike STM) capabilities make it an ideal tool to study local surface structures. For highly reactive surfaces like Si(100)-(2×1) (atomic hydrogen dereconstructs the surface [17]), ultrahigh vacuum (UHV) conditions are therefore paramount. In this Letter, we present details of the structure of the Si(100)-(2×1) surface using UHV-TED, in conjunction with a reliability factor (R factor) and χ^2 minimizations.

Thin samples of n-type Si(100) (B doped at 1 ohm cm) were mechanically polished, dimpled, and ion milled before being transferred into a UHV surface science chamber attached to a Hitachi UHV-H9000 300 keV electron microscope [18]. *In situ* sample preparation involved a cyclic combination of sputtering using 2–4 keV argon ions and electron-beam annealing cycles. Clean surfaces were characterized by the appearance of the (2×1) and (1×2) surface spots. No additional spots corresponding to the $c(4\times2)$ or $p(2\times2)$ structures were ever observed during the entire experiment. Figure 1 shows a typical selected area diffraction pattern of the clean Si(100) surface [19]. A series of such patterns were recorded at different exposures (the photographic film was calibrated to be linear over this range). These were then digitized to 8 bits using an Optronics P1000 microdensitometer (14 electrons per count) and analyzed using SEMPER imaging software on Apollo work stations. The relative intensity value for each spot was obtained by a cross-correlation technique [20]. Two sets of data (from orthogonal domains) were obtained from each pattern; typically, each set contained 50–100 independent beams.

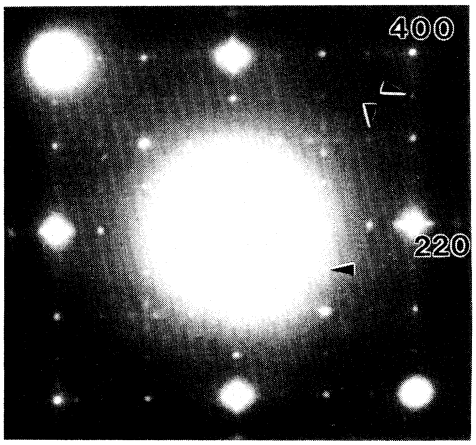

FIG. 1. A typical selected area diffraction pattern of the Si(100)-(2×1) surface. Arrows indicate the surface spots while two bulk spots are also indicated for reference.

0031-9007/93/71(21)/3489(4)$06.00

Experimental intensities at two different sample tilts (54 and 83 mrad) and thicknesses (13.5 and 45.6 nm) were analyzed. Corresponding theoretical intensity values were obtained for a given input structure (of surface and bulk) using both simple kinematical diffraction and rigorous double precision dynamical multislice methods, the input structure being optimized numerically to yield the best fit between the experimental and calculated data. These minimizations [21] included two forms of R factors and reduced χ^2 fittings and were performed using the routine NL2SOL from Netlib [22]. The R factors and χ^2 are defined as follows:

$$R^n = \left[\sum |I_e(g) - I_c(g)|^n\right] \Big/ \sum I_e^n(g) ,$$

$$\chi^n = \sum [|I_e(g) - I_c(g)|^n / \sigma^n(g)] 1/(N-M) ,$$

where $I_e(g)$ and $I_c(g)$ are the experimental and calculated intensities for each diffracted beam $\mathbf{g}$, $\sigma(g)$ is the standard deviation of the error distribution for each $\mathbf{g}$, N is the number of diffraction beams, and M is the number of parameters that are varied. $n=2$ defines a standard form while $N=1$ defines a more robust version.

In the initial stages of the minimization, atoms on the surface and a few layers beneath it were allowed to freely vary from their ideal bulk positions along the "2" direction of the reconstruction (y axis in the calculations). Intensity calculations used a simple kinematical diffraction approach [14,15]; however, the minimizations yielded physically unrealistic atom positions. Although these calculations are simple and save computer time, they can be inadequate [23] and are in any case unnecessary. Minimizations using intensities calculated from a dynamical approach yielded an asymmetric structure; however, χ^2 values indicated an overfit of data.

A strain field with the correct periodic structure for the reconstructed surface was therefore applied to constrain the atom positions to obey the equations of inhomogeneous, isotropic elasticity [21]. This displacement strain field has a periodic character in the x-y plane and is of the form

$$\mathbf{D}(\mathbf{r}) = A\nabla([z+\alpha]\omega) + B\nabla\omega + C\nabla\times(\mathbf{n}\omega) , \qquad (1)$$

where ω is a harmonic function, $\mathbf{n}$ is the unit vector, α [$=2(2-\nu)$, ν is Poisson's ratio] is a constant, and A, B, and C are variables. Using x and y in units of the unit cell, an appropriate form for ω is

$$\omega = \exp(-2\pi q_z z)\exp(-2\pi i[q_x x + q_y y]) ,$$

with $q_z^2 = q_x^2 + q_y^2$, where q_x and q_y are integers representing the Fourier periodicity of the reconstruction. This form has enough generality to model any subsurface distortions and has the correct physical form of decaying into the bulk of the material (while boundary effects may change the strain field for extremely thin materials, e.g., 2–4 nm, they are not a problem for the thicknesses used here). The first two terms in Eq. (1) are associated with a longitudinal wave while the third term represents a shear wave (used only to model twist). Rather than varying the atom positions, A, B, and q_y ($1 \le q_y \le 3$ in most calculations, $q_x = 0$) were varied (i.e., vary the strain field) with an appropriate choice of sine and cosine terms to enforce symmetry requirements if necessary. The symmetric and asymmetric structures proposed by Roberts and Needs [4] were used as the starting point of the minimization. Reconstructions were assumed to exist on both the top and bottom surfaces and a linear incoherent combination of the two was used. Minimizations were performed using both domains of the reconstructions and multislice calculations were performed for each since the orientation with respect to the specimen tilt was different. Also, since intensities are sensitive to the asymmetry direction, minimizations were carried out for both and an average value used. It should be noted that the absolute intensities correlated to approximately a monolayer coverage (of reconstruction) on both surfaces.

Initially, a four layer relaxation was assumed; the Debye-Waller (DW) terms for each of these layers were used as variables since anisotropic mean square displacements have been reported [24]. However, visual fits for higher order reflections were poor and the minimized DW values were physically unreasonable (i.e., much too small). Since intensities at larger reciprocal lattice vectors are very sensitive to both the strain field and the DW term, an incorrectly defined strain field can be compensated by unrealistic DW values. When eight layers were allowed to relax, good visual fits for higher order reflections were obtained and the DW terms minimized to physically realistic values. For such a relaxation model (i.e., eight layers), the minimizations were only weakly sensitive to variations in DW terms; therefore, fixed anisotropic DW values were used.

The surface structure obtained from our work was clearly asymmetric within both the scatter bars (between different minimizations) and error bars (for each minimization). Table I lists the average atom positions obtained from the different procedures at a 90% confidence level with the associated error bars and maximum scatter values for an eight layer relaxation model [atom displacements in the seventh and eighth layers are not listed since they are very small ($\sim 5\times 10^{-6}$ nm)]. This structure is shown schematically in Fig. 2. R factors (both standard and robust versions) associated with this structure were in the range 0.10–0.12 with χ^2 values of 1.08–1.5 (for the two diffraction cases), signifying excellent reliability of the structure. The lower R factors and χ^2 were obtained for the set with a larger number of beams, also a much better data set. The surface dimer bond is inclined out of the surface at an angle of 5°37′ and the bond length of 0.22 nm is in excellent agreement with the asymmetric structures proposed by Yin and Cohen [3] as well as by Roberts and Needs [4] (i.e., YC and RN structures).

In contrast to all other models in literature, the second layer atoms do not move towards each other (the in-

TABLE I. Atom positions for the asymmetric (2×1) structure (as fractions of cell parameters, $A=0.383$ nm, $B=0.767$ nm, $C=0.543$ nm). Error bars for only the y positions are reported since z positions are measured only indirectly, while the x positions were not varied in the minimizations. Scatter values for the y and z positions are reported to give an idea of the reliability of the structure.

	x	y (scatter, error)	z (scatter)
Layer 1	0.500	0.332 (3×10^{-3}, 5×10^{-3})	−0.024 (0.003)
	0.500	0.616 (2×10^{-3}, 5×10^{-3})	0.024 (0.003)
Layer 2	0.000	0.225 (3×10^{-3}, 2×10^{-3})	0.232 (0.002)
	0.000	0.698 (3×10^{-3}, 2×10^{-3})	0.268 (0.003)
Layer 3	0.000	0.007 (7×10^{-3}, 3×10^{-4})	0.569 (0.06)
	0.000	0.498 (6×10^{-3}, 4×10^{-4})	0.431 (0.05)
Layer 4	0.500	0.001 (5×10^{-4}, 1×10^{-4})	0.770 (0.02)
	0.500	0.499 (5×10^{-4}, 1×10^{-4})	0.730 (0.02)
Layer 5	0.500	0.254 (2×10^{-3}, 1×10^{-5})	0.9995 (1×10^{-4})
	0.500	0.746 (2×10^{-3}, 1×10^{-5})	1.0005 (1×10^{-4})
Layer 6	0.000	0.251 (2×10^{-3}, 1×10^{-5})	1.2499 (1×10^{-4})
	0.000	0.749 (1×10^{-3}, 1×10^{-5})	1.2501 (4×10^{-5})

tralayer bond length is identical to the RN and YC structures). We believe in the plausibility of this structure (low scatter and error bars) and attribute its detection to the high sensitivity of TED to atomic displacements parallel to the surface. The dimer background lengths of 0.24–0.25 nm lie in between the values for the RN (0.23 nm) and YC (0.25–0.28 nm) structures. While atom displacements extending into the fourth layer are asymmetric, scatter in the atom positions for the third layer makes it impossible to resolve the fine details of the structure of this layer. Atoms in all other subsequent layers in the unit cell move towards each other with small displacements noticeable up to the sixth layer. In general, for the first four layers, our structure is in better agreement with the YC structure than the RN structure. Although absolute atom displacements are small beyond the third layer, a strain field needs to be modeled for at least an eight layer relaxation to obtain a good fit. Since TED is relatively insensitive to displacements along the beam direction (z), high error and scatter bars are associated with the Δz values, especially below the second layer. It was possible to constrain the Δz values by increasing the number of q_y variables and still obtain the same Δy displacements. The possibility of an asymmetric structure with twist (suggested by [12]) was explored by including the shear term in Eq. (1); however, structures with large scatter in the x values (0.1–0.2 nm) and error bars were obtained.

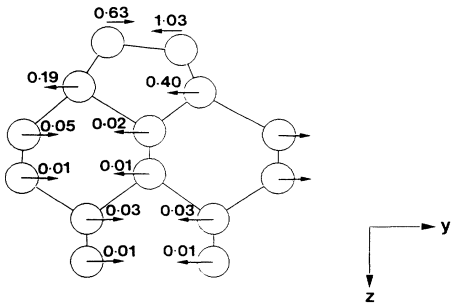

FIG. 2. A schematic of the Si(100)-(2×1) structure for a six layer relaxation model; the dimer bond lies in the plane of the paper. Arrows show the direction of dimerization while the numbers denote the magnitude (in Å) of atom displacements (along the "2" direction of the reconstruction, i.e., the y axis in the calculations) from the bulk positions.

When a constraint was imposed in the calculations to yield a symmetric structure, although the R factors were higher by only 0.02, χ^2 values were 1.6–2.0 (in the two diffraction cases), signifying high errors in the fit. Table II gives a quantitative comparison for the asymmetric (simple tilt) and symmetric structures obtained in this work using the better data set (similar behavior, albeit with higher values for the other set). Large enough step sizes were used in the calculations to avoid a local minimum. It should be mentioned here that with the strain field minimized from rigorous dynamical calculations as the starting point, minimizations run in conjunction with pseudokinematical diffraction calculations yielded an asymmetric structure with identical atom positions. Also, both pseudokinematical and dynamical diffraction calculations run with free atom variability calculations on this structure minimized to exactly the same positions. Larger unit cell calculations carried out using atom positions in the $c(4\times2)$ and $p(2\times2)$ structures also yielded the asymmetric (2×1) model. This structure thus appears to be a minimum for all the methods.

We have presented here results from a detailed study of the Si(100)-(2×1) surface testing the symmetric and asymmetric models using experimental intensities in UHV-TED patterns. Kinematical diffraction calculations prove to be only useful [25], even in the off-zone axis conditions (in the two diffraction cases), to check the relia-

TABLE II. Summary of standard R factors and χ^2 values for models calculated for the better data set shows the conclusive preference for the simple asymmetric structure (i.e., no twist).

	Four layer relaxation	Eight layer relaxation	Forced symmetry
Structure obtained	Asymmetric	Asymmetric	Symmetric
Standard R factor	0.1158	0.113	0.133
χ^2	1.32	1.08	1.61
Visual fit	Poor	Best	N/A

bility of the structure predetermined from a rigorous dynamical analysis. While the structure obtained from fitting a large set (~100–200 beams/pattern) is statistically highly probable, one should avoid the pitfall of overfitting the data by increasing the number of variable parameters (e.g., twist case) in the minimizations. Also, while our structure is asymmetric, due to the inherent insensitivity of the calculations to the DW terms, it is impossible to hypothesize on the validity of the dynamically fluctuating model suggested by STM.

We believe that our work represents the first attempt using rigorous dynamical transmission electron diffraction analyses in combination with R factor and χ^2 minimizations to predict the structure of the (2×1) surface to such a high degree of accuracy. The low values of R factors and χ^2 associated with these calculations reflect on the quality of the data and consequently on the excellent reliability of the structure obtained.

We would like to acknowledge the support of the National Science Foundation in funding this work.

[1] R. E. Schlier and H. E. Farnsworth, J. Chem. Phys. **30**, 917 (1959).
[2] D. J. Chadi, Phys. Rev. Lett. **43**, 43 (1979).
[3] M. T. Yin and M. L. Cohen, Phys. Rev. B **24**, 2303 (1981).
[4] N. Roberts and R. J. Needs, Surf. Sci. **236**, 112 (1990).
[5] E. Artacho and F. Yndurain, Phys. Rev. Lett. **62**, 2491 (1989).
[6] I. P. Batra, Phys. Rev. B **41**, 5048 (1990).
[7] R. M. Tromp, R. J. Hamers, and J. E. Demuth, Phys. Rev. Lett. **55**, 1303 (1985).
[8] R. J. Hamers, R. M. Tromp, and J. E. Demuth, Phys. Rev. B **34**, 5343 (1986).
[9] R. W. Wolkow, Phys. Rev. Lett. **68**, 2636 (1992).
[10] N. Jedrecy *et al.*, Surf. Sci. **230**, 197 (1990).
[11] A. I. Shkrebtii and R. Del Sole, Phys. Rev. Lett. **70**, 2645 (1993).
[12] J. A. Appelbaum and D. R. Hamann, Surf. Sci. **74**, 21 (1978).
[13] W. S. Yang, F. Jona, and P. M. Marcus, Phys. Rev. B **28**, 2049 (1983).
[14] K. Takayanagi *et al.*, Surf. Sci. **164**, 367 (1985).
[15] J. M. Gibson, Surf. Sci. Lett. **239**, L531 (1990).
[16] L. D. Marks, P. Xu, and D. N. Dunn, Surf. Sci. **294**, 324 (1993).
[17] J. Boland, Phys. Rev. Lett. **65**, 3325 (1990).
[18] L. D. Marks *et al.*, Ultramicroscopy **37**, 90 (1990).
[19] Although the TED experiments were conducted using 300 keV electrons, the surface structure did not undergo radiation damage; in fact, the lifetime was controlled solely by the partial pressure of water vapor in the system and longer survival times were observed when the sample was left in the UHV surface science chamber (a relatively cleaner system) before observation. However, recent investigations have shown that radiation damage is indeed a problem in other systems, e.g., Au-Si(111)-(5×2). Annealing effects due to the 300 keV beam are very minor (50–80 °C) and do not affect the surface structure.
[20] P. Xu, G. Jayaram, and L. D. Marks (to be published); standard deviation for each intensity value $I_g(g)$ obtained using this technique was calibrated at $\sigma=0.025\sqrt{I_e}(g)$.
[21] L. D. Marks (to be published).
[22] J. E. Dennis, D. M. Gay, and R. E. Welsch, ACM Trans. Math. Software **7**, 348 (1981).
[23] L. D. Marks, Ultramicroscopy **45**, 145 (1992).
[24] A. Mazur and J. Pollmann, Surf. Sci. **225**, 72 (1990).
[25] The raw data did not have the P_{2mm} symmetry which it would have if the diffraction were kinematical. In addition there is no way of estimating the effect of systematic errors in kinematical theory on the atomic positions.

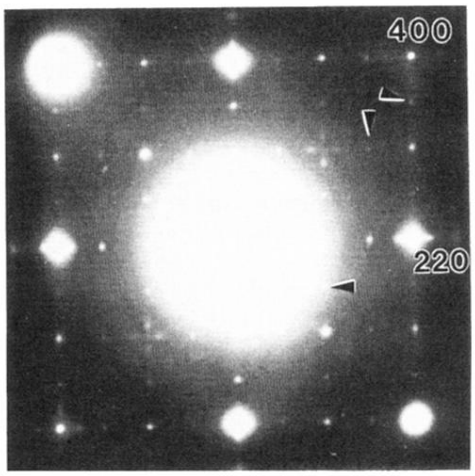

FIG. 1. A typical selected area diffraction pattern of the Si(100)-(2×1) surface. Arrows indicate the surface spots while two bulk spots are also indicated for reference.