Stabilization mechanisms of polar surfaces: ZnO surfaces

Mao-Hua Du,[1] S. B. Zhang,[2] J. E. Northrup,[3] and Steven C. Erwin[4]
[1]*Materials Science and Technology Division and Center for Radiation Detection Materials and Systems, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA*
[2]*Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, New York 12180, USA*
[3]*Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, California 94304, USA*
[4]*Center for Computational Materials Science, Naval Research Laboratory, Washington, DC 20375, USA*
(Received 2 April 2008; revised manuscript received 26 September 2008; published 23 October 2008)

First-principles calculations reveal two mechanisms that compete to determine the structure of ZnO polar surfaces. One is the electron-counting rule, which favors semiconducting surfaces. The other is the large ZnO cohesive energy, which favors unreconstructed metallic surfaces with 1×1 periodicity. Their close competition results in crossovers in the preferred surface structure as the oxygen chemical potential is varied, consistent with a variety of surface morphologies observed under different experimental conditions.

DOI: 10.1103/PhysRevB.78.155424 PACS number(s): 68.35.B−, 68.47.Gh

I. INTRODUCTION

Identifying the mechanisms responsible for the morphology of surfaces is crucial to gaining better control of materials grown both by epitaxial and colloidal methods, for which surface structure can strongly affect crystal quality and even impurity incorporation.[1–3] For semiconductors, surface-energy minimization is often governed by the electron-counting rule (ECR).[4,5] The ECR concerns the minimization of the energy of the surface dangling bonds (DBs), which can be achieved by the surface reconstruction that balances the numbers of cation and anion DBs. The electronegativity of the surface atoms determines that the electrons should transfer from the cation DBs to anion DBs. The ECR is said to be fulfilled if the available DB electrons on the reconstructed surface exactly fill the anion DBs while leaving the cation DBs empty. Such surface is semiconducting (with no high-energy partially occupied DBs) and is likely to be the most stable surface. The reconstructions of many semiconductor surfaces are consistent with the ECR.[4–7]

For ZnO, both of its polar surfaces exhibit 1×1 periodicity in low-energy electron diffraction (LEED).[8] Scanning tunneling microscopy (STM) shows irregularly distributed triangular islands and pits of various sizes on Zn-terminated (0001) surface.[8,9] On O-terminated $(000\bar{1})$ surface, only wide steps have been observed.[8] No published STM images have atomic resolutions. We note that the ECR does not offer an apparent explanation to the observed triangular islands and pits or the 1×1 periodicity.

Many previous studies on the ZnO polar surfaces have focused on classical electrostatic considerations.[8–14] The diverging electrostatic potential associated with the polar surface may be compensated either by the modification of the surface stoichiometry or by the surface metallization. The former involves surface reconstruction without significant changes of the charges of the surface atoms while the latter modifies the charges of the surface atoms from the bulk values without undergoing surface reconstruction. Both approaches, with and without surface reconstructions, can converge the electrostatic potential. Hence, the electrostatic compensation does not determine whether a particular surface reconstruction should or should not occur. This argument is supported by recent studies of several polar surfaces.[15–17] For the same reason, one cannot explain the triangular islands and pits observed on the ZnO (0001)-Zn surface using the electrostatic argument.[18] One can show, within a simple classical electrostatic model considering modified surface atomic charges, that the electrostatic potential is always convergent no matter the ZnO polar surfaces are reconstructed or not (see Appendix).

In this paper, we show that ZnO surface reconstructions can be understood as arising from a competition between two mechanisms. The first is the ECR, which favors semiconducting surfaces. The second, which favors metallic surfaces, arises from the large ZnO cohesive energy (typical for many oxides) and hence opposes the bond breaking necessary to create surfaces that satisfy the ECR. The close competition between the two mechanisms has two consequences. (1) There exists in the surface-energy diagram a stability crossover between semiconducting and metallic surfaces as the oxygen chemical potential, μ_O, is varied. (2) There is a crossover in the size preference for triangular islands and pits on the surface. Near the crossover point, islands and pits of different sizes may form due to their similar formation energies. These two crossovers occur near the same value of μ_O. Various observed ZnO surface structures, including those with triangular islands and pits,[8] can be understood as arising from these two competing surface stabilization mechanisms.

II. METHODS

Our calculations are based on density-functional theory (DFT) within the generalized gradient correction,[19] using the projector augmented-wave method.[20] The wave functions are expanded in a plane-wave basis with cutoff energy of 280 eV. ZnO surfaces were modeled by slabs of five double layers. All atoms were relaxed, except in the bottom two double layers, until their forces were less than 0.05 eV/Å. Surface energies were calculated relative to the relaxed unreconstructed surface.

III. RESULTS

We now show how the ECR and the cohesive energy compete to determine the structure of the ZnO surface. For clarity, in the rest of this paper we discuss mainly the $(000\bar{1})$-O surface and then show that the stabilization mechanisms for the $(000\bar{1})$-O surface also apply to the (0001)-Zn surface.

On the ZnO$(000\bar{1})$-O surface each O DB creates 1/2 hole in the valence band. Removing one surface O atom creates three Zn DBs (each is a $\frac{1}{2}$-electron donor) and eliminates one O DB. Therefore, an O vacancy (V_O) is a net double-electron donor. The same is true for a Zn adatom (Zn_{ad}), which has two valence electrons, but an isolated Zn_{ad} is less stable than an isolated V_O by 0.11 eV (as calculated in a 4×4 supercell). Increasing the V_O coverage (ρ_{V_O}) reduces the hole density on the surface. When ρ_{V_O} reaches 1/4, all the holes are eliminated and the ECR is satisfied. However, the removal of surface holes by forming V_O, which is favored by the ECR, also incurs an energy cost for Zn-O bond cleavage, which is strongly penalized because of the large cohesive energy of ZnO. To analyze the competition between the ECR and the large ZnO cohesive energy, we express the formation energy of a V_O as

$$\Delta H_{V_O} = 3E_{\text{Zn-O}} + \Delta E_{V_O}^{\text{ET}} + \mu_O - E_O^{\text{coh}}. \quad (1)$$

The first term, $3E_{\text{Zn-O}}$, is the energy cost for breaking three Zn-O bonds on the surface. The second term, $\Delta E_{V_O}^{\text{ET}}$, is the energy change arising from the transfer of electrons from Zn DBs around the V_O to surface O DBs, including the effects of relaxation on the DB levels and on the local geometry around V_O. This term is always negative for spontaneous charge transfer. The remaining term μ_O-E_O^{coh} is the lowered energy arising from transferring the isolated O atom to the O reservoir. Here μ_O is the relative O chemical potential referenced to half of the energy of an isolated O_2 molecule ($E_{O_2}/2$). When surface is in equilibrium with bulk, we obtain $-\Delta H_f(\text{ZnO}) < \mu_O < 0$ where ΔH_f(ZnO) is the heat of formation of ZnO, which we calculate to be 3.24 eV. $E_O^{\text{coh}} = E_{O,\text{atom}} - E_{O_2}/2$ is the binding (or cohesive) energy per atom in O_2. We note that $\Delta E_{V_O}^{\text{ET}}$ is the energy gain from fulfilling the ECR at the cost of $3E_{\text{Zn-O}}$ (determined by the ZnO cohesive energy). In Eq. (1), only $3E_{\text{Zn-O}}$ is positive and may render the formation of V_O on the surface unstable when $\mu_O > \mu_O^T = E_O^{\text{coh}} - \Delta E_{V_O}^{\text{ET}} - 3E_{\text{Zn-O}}$, where μ_O^T is the transition O chemical potential at which the formation of the single vacancy costs zero energy (i.e., $\Delta H_{V_O}=0$).

Indeed, our DFT calculations show such a transition on the ZnO$(000\bar{1})$-O surface [see Fig. 1(a)]. Under O-poor conditions, the lowest surface energy is obtained for 1/4 coverage of oxygen vacancies, which gives rise to a semiconducting surface. Different arrangements of vacancies with 1/4 coverage, e.g., $3V_O/(3\times4)$, $4V_O/c(4\times8)$, etc., lead to negligible differences in energy.[21] When the V_O coverage deviates from 1/4, e.g., 2/9 $[2V_O/(3\times3)]$, 1/16 $[1V_O/(4\times4)]$, and 1/3 $[1V_O/(\sqrt{3}\times\sqrt{3})]$, the surfaces are metallic and have higher surface energies, as shown in Fig. 1(a). The calculated

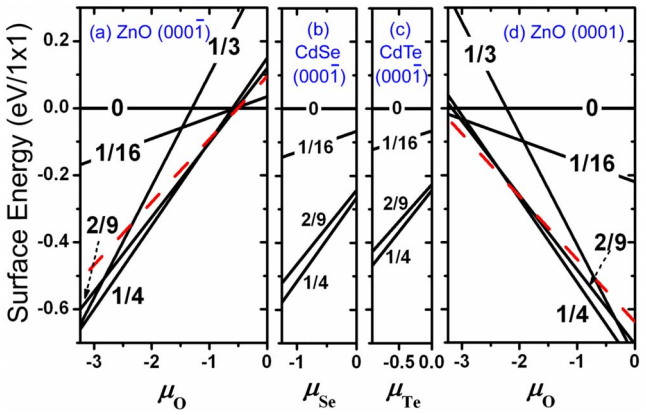

FIG. 1. (Color online) DFT surface energies, in eV per 1×1 cell, of reconstructed (a) ZnO$(000\bar{1})$, (b) CdSe $(000\bar{1})$, (c) CdTe $(000\bar{1})$, and (d) ZnO (0001) surfaces with varying V_O, V_{Se}, V_{Te}, and O_{ad} coverages, respectively. The red dashed lines in (a) and (d) are the surface energies for one pit ($n=3$) per 4×4 area. The lower bound of μ_O is determined by the ZnO heat of formation (Ref. 23). The bounds of the Se (Te) chemical potential were determined by crystalline Se (Te) with the α-Se (γ-Se) structure and the heat of formation for CdSe (CdTe) (Ref. 23).

density of states for the surfaces of several different V_O coverages (Fig. 2) confirms that the surface is semiconducting when $\rho_{V_O}=1/4$, and becomes metallic with hole carriers when $\rho_{V_O}<1/4$ and with electron carriers when $\rho_{V_O}>1/4$.

The results in Figs. 1(a) and 2 demonstrate that, under O-poor conditions, the semiconducting surfaces satisfying the ECR are favored. However, the ECR fails under O-rich conditions, where the metallic 1×1 unreconstructed surface is most stable, in agreement with Meyer.[22] This can be attributed to the large ZnO cohesive energy (which we calculate to be 2.32 eV per Zn-O bond) that favors a surface stoichiometry same as that of bulk. For semiconductors with smaller cohesive energies, e.g., wurtzite CdSe and CdTe with cohesive energies of 1.36 and 1.19 eV per bond, respectively, the energy cost of bond cleavage associated with the formation of surface vacancies are insignificant compared to the energy benefit from fulfilling the ECR. As a result, the CdSe and CdTe surfaces with 1/4 monolayer anion vacancies are calculated to be more stable than the unreconstructed surface over the entire range of allowed chemical potential [Figs. 1(b) and 1(c)], in contrast to the ZnO surface [Fig. 1(a)].

For each V_O reconstruction on the ZnO$(000\bar{1})$-O surface with concentration $\rho_{V_O}\leq1/4$, the transition points between favoring and not favoring V_O reconstruction all nearly coincide at μ_O^T. This is because at moderate concentrations the vacancies do not interact appreciably (the V_O repulsion is calculated to be 0.04 eV per V_O for 1/4 ML of V_O.), and hence the surface energy is approximately $\rho_{V_O}\times\Delta H_{V_O}$. For $\rho_{V_O}>1/4$, the number of Zn DBs exceeds the number of O DBs. Thus, electron carriers are created in the conduction band, resulting in much higher ΔH_{V_O} for additional V_O when $\rho_{V_O}>1/4$.

We have demonstrated that the ECR favors the formation of 1/4 ML of V_O, whereas the large ZnO cohesive energy

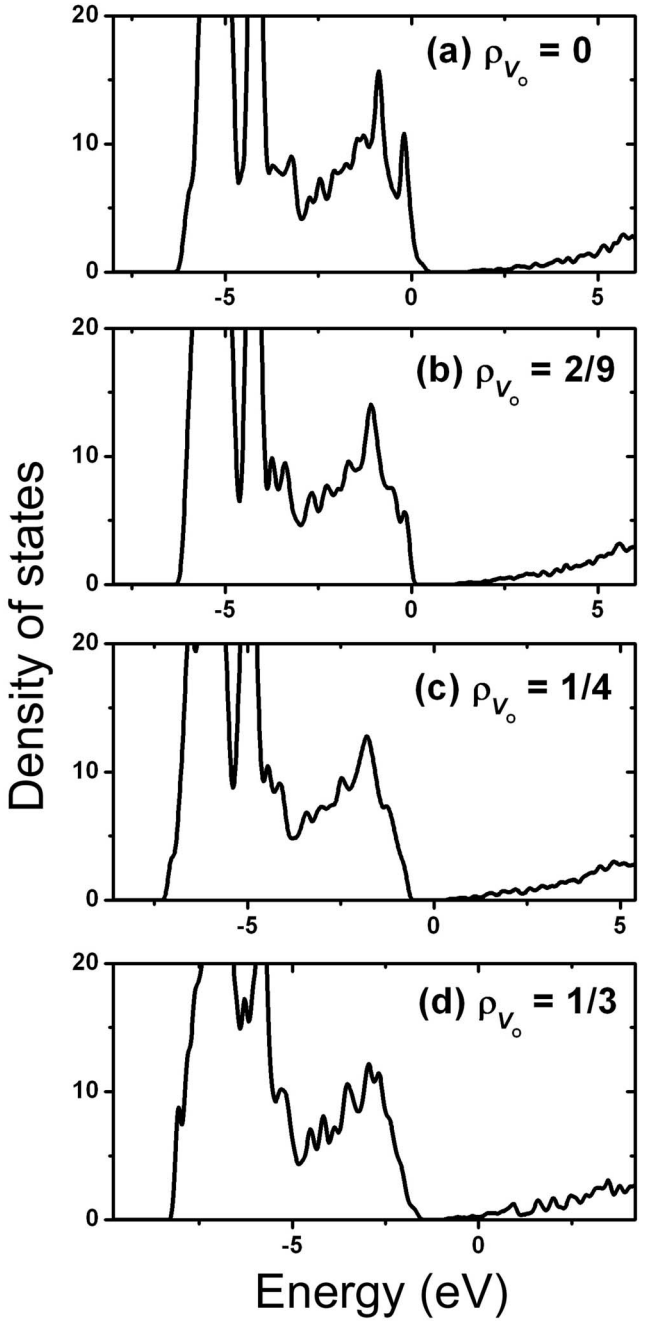

FIG. 2. Calculated density of states for ZnO(000$\bar{1}$) surfaces with ρ_{V_O}=(a) 0, (b) 2/9, (c) 1/4, and (d) 1/3. The density of states correspond to surfaces with five double layers and are normalized by the surface areas. The Fermi energies are set to zero for all figures.

favors the unreconstructed surface. The relative stability among the different surfaces with $\rho_{V_O} < 1/4$ is controlled by μ_O with a crossover at μ_O^T. Surfaces with $\rho_{V_O} > 1/4$ are not favored by either mechanism and thus are unstable. Next, we will show that when $\mu_O \approx \mu_O^T$, more complex surface reconstruction may form, i.e., the formation of triangular islands and pits of various sizes. The random distribution of triangular islands and pits of various sizes, as observed in STM images,[8] cannot be easily studied in a supercell model. Fortunately, the stability trend of islands and pits and the condition for forming the surface with numerous islands and pits of various sizes can be understood from the formation energies of isolated triangular islands and pits. The energetics of an island or a pit is a generalization of that for a vacancy or an adatom. A vacancy and an adatom can be considered as the smallest triangular pit and island, respectively. As will be shown below, the formation energy of an island or pit of arbitrary size should be nearly zero when $\mu_O \approx \mu_O^T$.

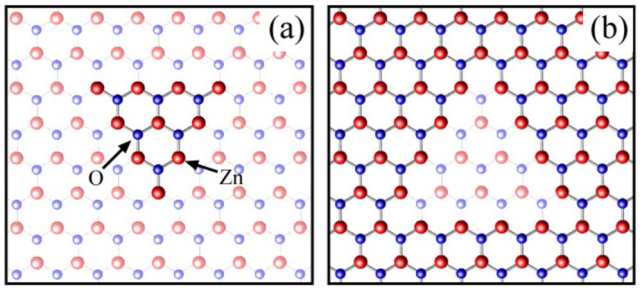

FIG. 3. (Color online) Atomic models of (a) a triangular island (b) a triangular pit with side length n=4 on the (000$\bar{1}$)-O surface.

Figure 3 shows examples of triangular island and pit with side length of n=4 (which is the number of atoms at one edge of the triangle). Forming a pit with an edge length of n requires the transfer of $n(n-1)/2$ Zn-O pairs and n additional O atoms to their respective reservoirs. This is equivalent to the net transfer of n excess surface O atoms to the O reservoir and cleaving $3n$ Zn-O bonds. In addition, $3n \times \frac{1}{2}$ electrons are transferred from $3n$ Zn DBs to $3n$ O DBs at the edges of the pit, and the $3n$ threefold coordinated Zn atoms undergo relaxation. In comparison, the formation of a single O vacancy involves the transfer of one surface O atom to the O reservoir, the transfer of $3 \times \frac{1}{2}$ electrons from three Zn DBs to three O DBs at the edges of the vacancy, and the relaxation of three threefold coordinated Zn atoms. Hence, we arrive at a simple energy relation between a triangular pit of edge length n and a single O vacancy,

$$\Delta H_n = n\Delta H_{V_O} + \delta E_n. \quad (2)$$

The energy difference (δE_n) between a triangular pit of edge length n and n isolated V_O's may be small because it arises from merely a spatial redistribution of $3n$ Zn DBs and the related strain redistribution on the surface. The same analysis also holds for triangular islands. Consequently, the formation energies of islands and pits should all change sign close to μ_O^T because $\Delta H_n = 0$ at $\mu_O = \mu_O^T - \delta E_n/n$. Indeed, DFT calculations (see Fig. 4) show that, except for some small islands ($2 \le n \le 5$), $|\delta E_n/n|$ is small and typically <0.1 eV. The rise of $|\delta E_n/n|$ for small islands is the result of insufficient strain relaxation. The relaxation of the undercoordinated edge atoms of an island causes strain inside the island, which becomes increasingly more difficult to relax as the size of the island decreases. On the other hand, the strain caused by the edge atoms of a pit is applied to the rest of the surface and is thus much easier to relax. As a result, $|\delta E_n/n|$ is generally small except for very small islands.

Since the formation energy of an individual island or pit changes sign close to μ_O^T, the surface energy for a recon-

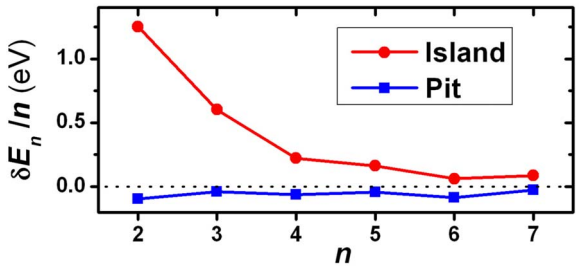

FIG. 4. (Color online) $\delta E_n/n$ for triangular islands and pits with edge length $n=2-7$ on the ZnO(000$\bar{1}$)-O surface. Calculations were performed using a 4×4 surface cell for $n=1,2,3$, and using 5×5, 6×6, 7×7, 8×8 cells for $n=4,5,6,7$, respectively.

structed surface with any number of triangular islands and pits also change sign close to μ_O^T, as demonstrated for example by a surface with one pit ($n=3$) per 4×4 area [see Fig. 1(a)].

From Eq. (2), we can see that the formation energy of a triangular island or pit ($\sim n$) normalized by its occupied surface area ($\sim n^2$) scales as $1/n$. Thus, when ΔH_n is negative, small islands or pits are more favored than large ones and the vacancy reconstruction is most favored; when ΔH_n is positive the reverse is true. Neglecting small δE_n in Eq. (2), the transition of size preference occurs exactly at μ_O^T as schematically shown by Fig. 5(a). When $\mu_O \approx \mu_O^T$, islands and pits of various sizes may coexist because their energy differences are very small. Figure 6(a) shows the calculated formation energies of triangular islands and pits normalized by their areas on the (000$\bar{1}$)-O surfaces. A size preference crossover is clearly seen.

The crossover of size preference of triangular islands and pits shown schematically in Fig. 5(a) can be qualitatively understood by the competition between the ECR and the large ZnO cohesive energy. On the (000$\bar{1}$)-O surface, a triangular island of edge length n terminates $n(n+1)/2$ O DBs on the surface and expose $n(n-1)/n$ O DBs on top of the island and $3n$ Zn DBs at the edges of the island. Since each O or Zn DB has 1/2 hole or 1/2 electron, respectively, overall, a triangular island donates $2n$ electrons to the (000$\bar{1}$)-O surface. Similarly, a triangular pit of edge length of n removes $n(n+1)/2$ O DBs and expose $n(n-1)/n$ O DBs within the pit and $3n$ Zn DBs at the edges of the pit. Overall, a triangular pit also donates $2n$ electrons to the (000$\bar{1}$)-O surface.

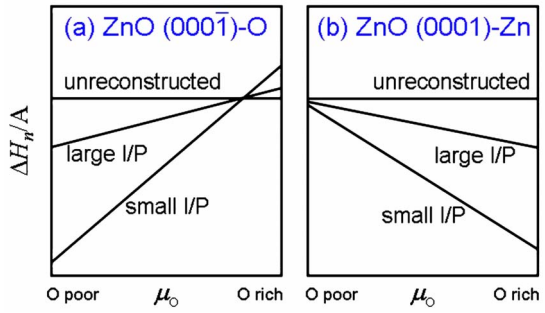

FIG. 5. (Color online) Schematic of formation energies of triangular islands/pits (I/P) normalized by their occupied surface areas ($\Delta H_n/A$) for (a) ZnO(000$\bar{1}$)-O and (b) ZnO (0001)-Zn surfaces, respectively. All the energy lines in (a) or (b) cross exactly at the same μ_O if $\delta E_n=0$ in Eq. (2).

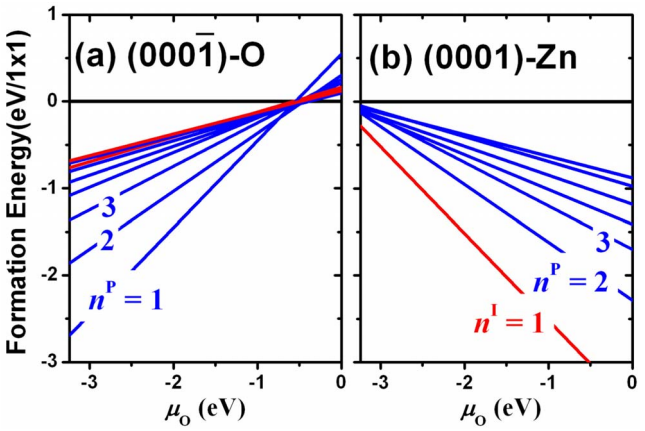

FIG. 6. (Color online) Formation energies for triangular islands and pits normalized by their areas [$n(n+1)/2$] on (a) (000$\bar{1}$)-O surface and (b) (0001)-Zn surface. On the (000$\bar{1}$)-O surface, the order of normalized formation energies for triangular pits at the O-poor limit is from $n^P=1-7$, and for triangular islands is from $n^I=6-7$. The formation energies of triangular islands with $2\leq n^I\leq5$ are not shown because these energies are much higher [large δE_n in Eq. (2) (see also Fig. 4)] and do not cross near μ_O^T (see text). For the (0001)-Zn surface, we show the formation energies of triangular pits, the order of which at the O-rich limit is from $n^P=2-7$, and that of a single O adatom, which is considered as the smallest triangular island ($n^I=1$). Calculations were performed using a 4×4 surface cell for $n=1,2,3$, and using 5×5, 6×6, 7×7, 8×8 cells for $n=4,5,6,7$, respectively.

The unreconstructed (000$\bar{1}$)-O surface has 1/2 hole per 1×1 surface area. Thus, the formation of triangular islands and pits, which are donors, reduces the carrier density on the surface and thus is favored by the ECR. As discussed above, a triangular island or pit of edge length of n donates $2n$ electrons on the (000$\bar{1}$)-O surface. Because the size of the triangular islands and pits scale with n^2, with a given area, a large number of small islands and pits can compensate holes on the (000$\bar{1}$)-O surface more effectively than a small number of large islands and pits. Thus, the ECR favors the small islands and pits. On the other hand, the large ZnO cohesive energy favors large triangular islands and pits because they correspond more closely to the bulk stoichiometry, and consequently their formation requires fewer bond cleavages. The competition between the ECR and the large ZnO cohesive energy results in a transition of the size preference for the triangular islands and pits as μ_O is varied.

Note that although large triangular islands and pits are not favored by ECR, they can host smaller islands and pits within them such that the ECR is also satisfied [e.g., a triangular pit of edge length of $n=3$ hosting an O vacancy within it in a 4×4 supercell on the (000$\bar{1}$)-O surface]. Since each triangular island or pit of edge length of n donates $2n$ electrons to the surface, the number of electrons donated from

such a triangular island or pit complex is given by twice the sum over the edge lengths of all triangular islands and pits, $2\Sigma_i n_i$. However, these structures usually involve relatively larger strain energies because of the close packing of the islands and pits.

We turn now to the Zn-terminated (0001) surface. On this surface, an isolated O_{ad} is more stable than an isolated V_{Zn} by 0.23 eV, as calculated in a 4×4 cell. Figure 1(d) shows that, at the O-rich (i.e., Zn-poor) limit, the most stable surface is the semiconducting surface with 1/4 ML of O_{ad} [$1O_{ad}/c(2\times4)$], which is more stable than the surface structure suggested in Ref. 18 under O-rich conditions. We have considered various possible O_{ad} distributions, all having the same 1/4 ML coverage, and found that they have different surface energies. The simplest oxygen adatom model has one oxygen atom in an H3 site in each 2×2 cell.[24] The present calculations predict that the $c(2\times4)$ arrangement of oxygen adatoms is more stable than the 2×2 pattern by 0.05 eV per 1×1 surface area. The crossover between the O_{ad}-reconstructed surface and the unreconstructed surface occurs near the O-poor limit. In our calculations the formation energy $\Delta H_{O_{ad}}$ of an isolated O_{ad} changes its sign at $\mu_O^T = -\Delta H_f(\mathrm{ZnO}) - 0.28$ eV. Therefore the unreconstructed surface is never favored.

Following the discussions for the $(000\bar{1})$-O surface, we can easily show that a triangular island or pit with edge length n donates $2n$ holes to the (0001)-Zn surface, thus reducing the electron carrier density on the surface. The ECR favors small islands and pits whereas the large cohesive energy favors large ones. The formation energy of a triangular island or pit (ΔH_n) on the (0001)-Zn surface can be related to that of a V_{Zn} ($\Delta H_{V_{Zn}}$), in analogy to Eq. (2). Furthermore, with $\Delta H_{O_{ad}} - \Delta H_{V_{Zn}}$ (=−0.23 eV) being a constant, we obtain $\Delta H_n = n\Delta H_{O_{ad}} + \delta E_n$. DFT calculations show that $|\delta E_n/n| < 0.12$ eV for pits with $n=2-7$. The size preference for triangular islands and pits on the (0001)-Zn surface is also analogous to that on the $(000\bar{1})$-O surface as shown schematically in Fig. 5(b) [neglecting small δE_n in Eq. (2)]. Figure 6(b) shows the calculated formation energies of triangular islands and pits normalized by their areas on the (0001)-Zn surface. The crossover occurs near the O-poor limit.

Note that Fig. 6 shows the energies for isolated triangular islands and pit, and neither Fig. 5 nor 6 takes into account the strain energy cost for the real surface where islands and pits may be located close to each other. Usually the compressive strain on the surface is easier to relax because the surface atoms can relax toward the open space above the surface. This is evidenced by the small repulsion energy of only 0.04 eV/V_O for 1/4 ML of V_O (which applies compressive strain) on the $(000\bar{1})$-O surface. However, the tensile strain on the surface is more difficult to relax as evidenced by a large repulsion energy of 0.44 eV/O_{ad} for 1/4 ML of O_{ad} (which applies tensile strain) on the (0001)-Zn surface. Thus, for the (0001)-Zn surface, although the area-normalized formation energy for an isolated O_{ad} is lower than that for an isolated larger pit as shown in Fig. 6(b), a surface with less densely packed pits could be more stable than that with 1/4 ML of O_{ad} under the O-poor conditions, as demonstrated for example by a surface with one pit ($n=3$) per 4×4 area in Fig. 1(d). Near the O-poor limit, islands and pits of different sizes have small formation energy differences and may thus coexist.

IV. DISCUSSION

Here we summarize our theoretical predictions of the ZnO polar surface structures and compare with available experimental results. On the $(000\bar{1})$-O surface, we predict that (1) 1/4 ML of oxygen vacancy should form under the O-poor conditions, which results in a semiconducting surface; (2) the unreconstructed metallic surface should be most stable under the O-rich conditions; and (3) triangular islands and pits of different sizes can coexist on the surface if $\mu_O \approx \mu_O^T$ because their formation energies are all close to zero when $\mu_O \approx \mu_O^T$. On the (0001)-Zn surface, we predict that (1) 1/4 ML of oxygen adatom should be most stable under the O-rich conditions, which results in a semiconducting surface, and (2) triangular islands and pits of different sizes may form near the O-poor limit because μ_O^T is close to the O-poor limit.

Experimentally, ZnO surfaces are usually prepared in ultrahigh vacuum (UHV) with only residual O_2 pressure, corresponding to O-poor conditions. Hence the observation of triangular islands or pits with varying sizes on the (0001)-Zn surface[8] is consistent with our results. On the $(000\bar{1})$-O surface our prediction of 1/4 ML V_O must await STM images with atomic resolution. The STM images in Ref. 8 only show the surface morphology but have no atomic resolution. Note that the 1/4 ML V_O may not be fully ordered because our results show that different arrangement of vacancies with 1/4 coverage lead to negligible differences in energy. This may give rise to a pseudo "1×1" surface pattern and may explain the 1×1 periodicity observed by LEED.[8] The experimental observations of chemically reactive (0001)-Zn surface and inert $(000\bar{1})$-O surface[25,26] support our predictions of metallic (0001)-Zn surface and semiconducting $(000\bar{1})$-O surface under O-poor conditions. To reach the O-rich conditions on either surface may remain an experimental challenge because at a typical annealing temperature of ~700 °C, the O-rich conditions with $\mu_O > -0.5$ eV correspond to unrealistically high O_2 pressures ($>10^6$ bar), as estimated from the ideal-gas law.[16] Fortunately, plasma treatment may offer a route to O-rich conditions. Indeed, a recent STM study of the $(000\bar{1})$-O surface after O-plasma treatment at 800 °C showed the first atomically resolved image of unreconstructed surface,[27] consistent with our calculations.

Annealing in UHV may lead to nonequilibrium effects not addressed in the above thermodynamic model. For example, irreversible desorption of atoms from terraces and at step edges can result in the nucleation and growth of pits. Triangular islands may then be created by the merging of three triangular pits. The rates of these kinetic processes affect the size distribution of islands and pits. It is thus possible that surface morphologies observed in the STM images in Ref. 8 are not necessarily an outcome of equilibrium surface thermodynamics. Nevertheless, the fact that islands and pits are observed on the Zn-terminated surface but not on the

O-terminated surface[8] is consistent with our results, which suggests that equilibrium thermodynamics could play a significant role even during UHV annealing.

We should point out that the possible role of hydrogen has not been examined here. Hydrogen can terminate O dangling bonds on the $(000\bar{1})$ surface as discussed by Meyer.[22] On the (0001) surface there is the possibility that OH groups can saturate the Zn dangling bonds.[18] However, hydrogen is unlikely to play a significant role in the ZnO surface reconstructions observed in STM experiments, in which the residual hydrogen concentration is too low.[8]

Finally, we comment on several x-ray diffraction studies of ZnO polar surfaces in literature,[12,28,29] which attempted to obtain the surface structure information from the x-ray data. For the ZnO$(000\bar{1})$-O surface, a large contraction of the interlayer distance for the top bilayer was reported, although the exact contraction distance differs among the available results, ranging from 0.12 Å (Ref. 29) to 0.25 Å (Ref. 28) and 0.44 Å.[12] Our DFT calculations show a contraction distance of 0.26 Å for the top bilayer of the unreconstructed ZnO$(000\bar{1})$-O surface, in agreement with previous DFT calculations.[28] On the ZnO$(000\bar{1})$-O surface with 1/4 ML of O vacancy, the threefold coordinated Zn atoms relax away from the vacancy and have a lower position compared to the fourfold coordinated Zn atoms in the top Zn layer. In general, the contraction distance is reduced (or the interlayer distance increases) when the O vacancy is present. It should be noted that the experimental atomic positions were deduced from the parameter fitting of the x-ray data. The fitting procedure also yielded the occupancy probability of 0.7–0.75 for the Zn sites in the first Zn layer of the ZnO$(000\bar{1})$-O surface.[12,29] However, this result may not necessarily be interpreted as evidence of large number of Zn vacancies under the top O layer. It may be an indication of structure disorder or additional surface features that is absent in the model used to fit the x-ray data. For the ZnO(0001)-Zn surface, our DFT calculations show that the interlayer distance for the top bilayer of the unreconstructed surface decreases by 0.10 Å, in agreement with previous DFT calculations.[28] The presence of O adatoms or Zn vacancy increases or decreases the interlayer distance, respectively. The x-ray result suggests an increase in the interlayer distance by 0.05 Å for the top bilayer.[12] However, as the STM images suggested, the Zn surface possesses more disordered surface features than the O surface,[8] which may complicates the parameter fitting and the interpretation of the x-ray data.

V. CONCLUSIONS

We have shown that the reconstruction of polar surfaces is not driven by electrostatic compensation as discussed in Appendix. Our theoretical results demonstrate that the reconstruction of the ZnO polar surfaces is governed by the competition between the energy benefit of satisfying the ECR by creating surface vacancies or adatoms, and the energy lost of breaking Zn-O bonds (which is substantial, due to the large ZnO cohesive energy) as required for vacancy or adatom formation. The ECR favors semiconducting surfaces whereas the large ZnO cohesive energy favors metallic surfaces. Our calculations show a crossover between semiconducting and metallic surfaces as the O chemical potential (μ_O) is varied. On the $(000\bar{1})$-O surface, semiconducting surfaces with 1/4 ML oxygen vacancies are stable at low μ_O, and the unreconstructed metallic surface is stable at high μ_O. Analogously, on the (0001)-Zn surface, semiconducting surfaces with 1/4 ML O adatoms are stable at high μ_O, while metallic surfaces are stable at low μ_O. Near the crossover, triangular islands and pits of various sizes should form on both surfaces due to their small energy differences.

ACKNOWLEDGMENTS

We thank X. L. Du, J. F. Jia, and Q. K. Xue for sharing unpublished results. This work was supported by the Office of Naval Research, by the NRL-NRC program, by DOE Office of Nonproliferation Research and Development NA 22, and by DOE/BES and DOE/EERE under Contract No. DE-AC36-99GO10337. Computations were performed at NERSC and the DoD Major Shared Resource Center at ASC.

APPENDIX

Here we demonstrate that the electrostatic compensation requirement is fulfilled no matter the ZnO polar surfaces are reconstructed or not. In a point-charge model, it is well known that alternating cation and anion layers stacked along the [0001] direction of a wurtzite crystal lead to a diverging electrostatic potential.[10,11] To cancel this divergence the charge at the surface must be modified. In general, the electrostatic potential of such a stack will converge if the top m layers have charge densities σ_i per unit area satisfying[11]

$$\sum_{i=1}^{m} \sigma_i = -\frac{\sigma_{m+1}}{2}\left[1 + (-1)^{m-1}\frac{R_2 - R_1}{R_2 + R_1}\right], \tag{A1}$$

where σ_{m+1} is the charge per unit area in the bulklike layers, and R_1 and R_2 are the separation within and between double layers, respectively. Partitioning the valence electrons among the bonds, Zn and O contribute $\frac{1}{2}$ and $1\frac{1}{2}$ electrons to each Zn-O bulk bond, respectively. On the surface, each threefold coordinated Zn or O has $\frac{1}{2}$ electron or hole in its Zn or O DB, respectively. Thus, on the unreconstructed ZnO (0001)-Zn surface, each surface Zn has a formal charge of $+2-\frac{1}{2} = +1\frac{1}{2}$, which is 75% of the bulk Zn formal charge state of +2 in the ideal ionic model for ZnO. Similarly, on the unreconstructed $(000\bar{1})$-O surface, each surface O has a formal charge of $-1\frac{1}{2}$, which is 75% of the bulk O formal charge of −2. Considering $m=1$ in Eq. (A1) for unreconstructed surfaces, the charge configurations on both surfaces, $\sigma_1/\sigma_2 \approx -0.75$, satisfy Eq. (A1) because $R_2/R_1 \approx 3$.

It is easy to demonstrate that Eq. (A1) also holds for surfaces containing triangular islands and pits for which m is an odd number. Consider a ZnO$(000\bar{1})$-O surface with N^I islands and N^P pits, examples of which are shown in Fig. 3 in the main text. Regardless of how the islands and pits are arranged (including islands within pits or pits within islands) the left side of Eq. (A1) (multiplied by total surface area) is given by $-1.5N_{\mathrm{ODB}}-0.5N_{\mathrm{ZnDB}}$, where N_{ODB} and N_{ZnDB} are the total number of O and Zn DBs, respectively. To determine N_{ODB} and N_{ZnDB}, consider first a triangular island, which has $n^I(n^I+1)/2$ Zn atoms and $n^I(n^I-1)/2$ O atoms; here n^I is the number of Zn atoms at one edge of the triangle. Similarly, a triangular pit has $n^P(n^P+1)/2$ O atoms and $n^P(n^P-1)/2$ Zn atoms removed; here n^P is the number of O atoms at one edge of the removed triangle. For an unreconstructed surface, $N_{\mathrm{ODB}}=N$, where N is the total number of atoms in one layer. The addition of one triangular island terminates $n^I(n^I+1)/2$ O DBs underneath the island and creates $n^I(n^I-1)/2$ O DBs on top of the island and $3n^I$ Zn DBs at the three edges of the island. The addition of one triangular pit removes $n^P(n^P+1)/2$ O DBs since $n^P(n^P+1)/2$ O atoms on the top of the surface is removed and exposes $n^P(n^P-1)/2$ O DBs at the bottom of the pit and $3n^P$ Zn DBs at the three edges of the pit. After summing over all the O and Zn DBs associated with all islands and pits on the surface, we have

$$N_{\mathrm{ODB}} = N + \sum_i^{N^I} \frac{n_i^I(n_i^I-1)}{2} - \sum_i^{N^I} \frac{n_i^I(n_i^I+1)}{2} + \sum_i^{N^P} \frac{n_i^P(n_i^P-1)}{2} - \sum_i^{N^P} \frac{n_i^P(n_i^P+1)}{2} = N - \sum_i^{N^I} n_i^I - \sum_i^{N^P} n_i^P,$$

and

$$N_{\mathrm{ZnDB}} = 3\sum_i^{N^I} n_i^I + 3\sum_i^{N^P} n_i^P$$

Hence $-1.5N_{\mathrm{ODB}}-0.5N_{\mathrm{ZnDB}}$ reduces to $-1.5N$. This is equal to the right side of Eq. (A1) (multiplied by total surface area), taking into account that m in Eq. (A1) is always an odd number for a surface with triangular islands and pits. The same analysis also holds for the (0001)-Zn surface. Thus, the electrostatic potential is always converged no matter if the surface is unreconstructed or reconstructed with any number of triangular islands and pits of arbitrary size. We conclude that electrostatic compensation does not provide a driving force for the formation of islands and pits on Zn- and O-terminated polar surfaces.

[1] Z. X. Mei, X. L. Du, Y. Wang, M. J. Ying, Z. Q. Zeng, H. Zheng, J. F. Jia, Q. K. Xue, and Z. Zhang, Appl. Phys. Lett. **86**, 112111 (2005); Z. X. Mei, Y. Wang, X. L. Du, M. J. Ying, Z. Q. Zeng, H. Zheng, J. F. Jia, Q. K. Xue, and Z. Zhang, J. Appl. Phys. **96**, 7108 (2004).

[2] X. Luo, S. B. Zhang, and S.-H. Wei, Phys. Rev. Lett. **90**, 026103 (2003).

[3] S. C. Erwin, L. Zu, M. I. Haftel, A. L. Efros, T. A. Kennedy, and D. J. Norris, Nature (London) **436**, 91 (2005).

[4] D. J. Chadi, J. Vac. Sci. Technol. A **5**, 834 (1987).

[5] M. D. Pashley, Phys. Rev. B **40**, 10481 (1989).

[6] J. E. Northrup and S. Froyen, Phys. Rev. B **50**, 2015 (1994); A. Garcia and J. E. Northrup, J. Vac. Sci. Technol. B **12**, 2678 (1994).

[7] L. Zhang, E. G. Wang, Q. K. Xue, S. B. Zhang, and Z. Zhang, Phys. Rev. Lett. **97**, 126103 (2006).

[8] O. Dulub, L. A. Boatner, and U. Diebold, Surf. Sci. **519**, 201 (2002).

[9] O. Dulub, U. Diebold, and G. Kresse, Phys. Rev. Lett. **90**, 016102 (2003).

[10] P. W. Tasker, J. Phys. C **12**, 4977 (1979).

[11] C. Noguera, J. Phys.: Condens. Matter **12**, R367 (2000).

[12] N. Jedrecy, M. Sauvage-Simkin, and R. Pinchaux, Appl. Surf. Sci. **162-163**, 69 (2000).

[13] U. Diebold, L. V. Koplitz, and O. Dulub, Appl. Surf. Sci. **237**, 336 (2004).

[14] Y. Ding and Z. L. Wang, Surf. Sci. **601**, 425 (2007).

[15] X.-G. Wang, W. Weiss, Sh. K. Shaikhutdinov, M. Ritter, M. Petersen, F. Wagner, R. Schlogl, and M. Scheffler, Phys. Rev. Lett. **81**, 1038 (1998).

[16] K. Reuter and M. Scheffler, Phys. Rev. B **65**, 035406 (2001).

[17] R. Pentcheva, F. Wendler, H. L. Meyerheim, W. Moritz, N. Jedrecy, and M. Scheffler, Phys. Rev. Lett. **94**, 126101 (2005).

[18] G. Kresse, O. Dulub, and U. Diebold, Phys. Rev. B **68**, 245409 (2003).

[19] G. Kresse and J. Furthmuller, Phys. Rev. B **54**, 11169 (1996).

[20] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).

[21] $1\mathrm{Zn_{ad}}/(2\times2)$ is more stable than $1V_{\mathrm{O}}/(2\times2)$, but is less stable than many other reconstructions with 1/4 ML V_{O} such as $3V_{\mathrm{O}}/(3\times4)$ and $4V_{\mathrm{O}}/c(4\times8)$.

[22] B. Meyer, Phys. Rev. B **69**, 045416 (2004).

[23] S. B. Zhang, J. Phys.: Condens. Matter **14**, R881 (2002).

[24] J. E. Northrup and J. Neugebauer, Appl. Phys. Lett. **87**, 141914 (2005).

[25] V. Staemmler, K. Fink, B. Meyer, D. Marx, M. Kunat, S. G. Girol, U. Burghaus, and Ch. Wöll, Phys. Rev. Lett. **90**, 106102 (2003).

[26] Z. L. Wang, X. Y. Kong, and J. M. Zuo, Phys. Rev. Lett. **91**, 185502 (2003).

[27] X.-L. Du, J.-F. Jia, and Q.-K. Xue (private communication).

[28] A. Wander, F. Schedin, P. Steadman, A. Norris, R. McGrath, T. S. Turner, G. Thornton, and N. M. Harrison, Phys. Rev. Lett. **86**, 3811 (2001).

[29] N. Jedrecy, S. Gallini, M. Sauvage-Simkin, and R. Pinchaux, Phys. Rev. B **64**, 085424 (2001).