Competing stabilization mechanism for the polar ZnO(0001)-Zn surface

Georg Kresse,[1,*] Olga Dulub,[2] and Ulrike Diebold[2]

[1]*Institut für Materialphysik and Center for Computational Materials Science, Universität Wien, A-1090 Wien, Austria*
[2]*Department of Physics, Tulane University, New Orleans, Louisiana 70118, USA*
(Received 27 June 2003; published 11 December 2003)

Density-functional calculations for the (0001)-Zn surface of wurtzite ZnO are reported. Different stabilization mechanisms, such as metallization of the surface layer, adsorption of OH groups or O adatoms, the formation of Zn vacancies, and large scale triangular reconstructions are considered. The calculations indicate that isolated Zn vacancies or O adatoms are unfavorable compared to triangular reconstructions. In the absence of hydrogen, these triangular features are stable under any realistic temperature and pressure. When hydrogen is present, the reconstruction is lifted, and hydroxyl groups stabilize the ideal otherwise unreconstructed surface. The transition between the unreconstructed hydroxyl covered surface and the triangular shaped features occurs abruptly; OH groups lift the reconstruction, but their adsorption is energetically unfavorable on the triangularly reconstructed surface.

I. INTRODUCTION

When the wurtzite ZnO crystal is cleaved parallel to the basal plane, two structurally different surfaces are created. The surfaces are usually termed Zn-terminated (0001) surface and O-terminated (000$\bar{1}$) surface, as Zn and O atoms are located at the vacuum side of the ideal as-cleaved (0001) and (000$\bar{1}$) surfaces, respectively (see Fig. 1). Formally these two surfaces are polar, and it is a matter of simple electrostatic arguments that the cleavage energy is infinite in an *ionic* model.[1,2] But contrary to this formal instability in the ionic model, ZnO crystals exhibit a large fraction of both surfaces, indicating that these two surfaces are rather favorable in nature. Recent theoretical work indicates that the stability of the surfaces might be related to a fractional charge transfer of half an electron per surface area from the O-terminated side to the Zn-terminated surface[3,4] which necessitates partially occupied two-dimensional (2D) metallic bands at both surfaces. The suggested compensation mechanism is in agreement with the electrostatic arguments of Noguera[2] or Tasker.[1] Starting from the ionic model, electrostatics require a fractional charge transfer of half an electron from the O to the Zn side in order to lift the conceived electrostatic divergence. The metallic surface states, however, have not yet been observed using, e.g., photoemission experiments[5,6] or scanning tunneling spectroscopy (STS).[7] Alternatively, the compensation of the ionic excess charge at the surface can occur through two other mechanisms (i) the addition of (1/2)-ML (monolayer) OH$^-$ groups or (1/4)-ML O^{-2} at the surface, or (ii) the formation of (1/4)-ML Zn^{+2} vacancies at the Zn-terminated surface. In fact, surface x-ray diffraction experiments for the ZnO(0001)-Zn surface could best be fitted by allowing a 0.75 occupancy of the topmost Zn layer,[8] in agreement with stabilization mechanism (ii). The experiments gave no indication, how the removal of Zn atoms occurs, but certainly an ordered (2×2) arrangement of vacancies has not been observed in low-energy electron-diffraction or helium scattering experiments.[9] Recent scanning tunneling microscopy (STM) experiments finally solved

the long-standing controversy. They indicate that the removal of Zn atoms happens through the formation of triangular shaped reconstructions.[10] The experiments were accompanied by extensive large scale density-functional calculations, which confirm that these triangular shaped reconstructions are energetically favorable, if hydrogen is not present in the surrounding gas phase. The current work aims to present a detailed discussion of these *ab initio* calculations. The previous results are also refined by performing calculations for thicker slabs than considered before.[10]

Section II briefly describes the computational methods used in the present density-functional study. Results for the ideal bulk terminated surface and a discussion of the compensation mechanism operative at the ideal bulk terminated surface ("metallization") are presented in Sec. III A. Other compensation mechanisms, such as the formation of Zn vacancies or the addition of O adatoms (Sec. III B), hydroxyl adgroups (Sec. III D), and triangular shaped reconstructions (Sec. III E) are investigated as well. All results are combined in the form of an *ab initio* surface phase diagram in Sec. III H. We finish with discussion and conclusion.
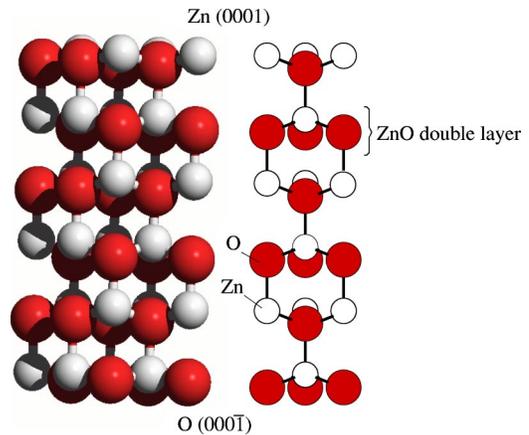

FIG. 1. Ideal as-cleaved Zn-terminated (0001) and O-terminated (000$\bar{1}$) surface of wurtzite ZnO.

II. METHOD

A. *Ab initio* calculations

The present first-principles calculations are based on density-functional theory (see, e.g., Refs. 11 and 12) and employ a plane wave basis set.[13,14] To determine the electronic ground state the Vienna *ab initio* simulation package[15,16] (VASP) is used. In the present calculations, the interaction between the ions and electrons is described by the projector augmented wave[17] (PAW) method in the implementation of Kresse and Joubert.[18] The radial cutoff for the Zn PAW potential is 1.2 Å, and that for the O PAW potential 1.0 Å. For Zn atoms, the $3d$, $4s$, and $4p$ electrons were described as valence, whereas for oxygen the $2s$ and $2p$ electrons were treated as valence. The remaining electrons were kept frozen. A plane-wave cutoff of 280 eV was chosen for the present calculations. At this plane-wave energy cutoff, the relative energies are converged to a few meV (with respect to the energy cutoff). Generalized gradient corrections (GGC) according to Perdew and Wang (PW91) were used throughout this work.[19,20]

With this setting the lattice parameters of ZnO are found to be $a = 3.250$ Å, $c = 5.207$ Å, and $u = 0.375$, in good agreement with experiment, and excellent agreement with previous theoretical studies.[21]

B. The surface model

The surface was modeled by slabs containing eight ZnO double layers (compare Fig. 1). The three topmost ZnO double layers were allowed to relax, except when noted otherwise. As discussed later, the ideal or "as-cleaved" O $(000\bar{1})$ and Zn (0001) surfaces are characterized by metallic surface states. At the O side, the metallic surface states are lying close to the valence-band edge, and, per primitive cell, half an electron is missing to fully occupy them. This electron hole at the edge of the valence band causes a very slow convergence of total-energy differences with respect to the slab thickness, since the Zn side tends to lose a small fraction of its electronic charge to this low-lying O surface state. The effect is an undesirable artifact of the finite-sized slab, since the charge transfer would be negligibly small for an infinite slab (see Sec. III A). To accelerate convergence, the surface O atom at the $(000\bar{1})$ side was replaced by an artificial atom with a valency of 6.5, i.e., intermediate between O and F. The fractional excess electron fills the hole at the valence-band edge, and the charge transfer from the Zn to the O side becomes impossible. In turn the calculations converge more rapidly with respect to the slab thickness. Calculations presented here were performed applying this compensation mechanism on the O side, except when noted. In the previous work,[10] no such correction was used, and only six-layer thick slabs were applied. The present calculations are therefore more accurate than the one reported before.

Various supercells were used to investigate the energetics of defects and surface reconstructions. Most calculations were performed using (4×4) and (6×6) supercells. To model larger reconstructions $(\sqrt{48} \times \sqrt{48})$ supercells were considered as well. Particular care had to be paid to k-point

sampling and convergence. For the (4×4) supercell, relaxations were performed using two special k-points with the coordinates $(0, 0, 0)$ and $(1/3, 1/3, 0)$ with respective weights of 1/3 and 2/3. To test for k point convergence, the energies were calculated using a denser grid of $(3 \times 3 \times 1)$ Monkhorst-Pack k points, as well. Except for very low defect concentrations, the formation energies were within 25 meV of those calculated with two k points (see Sec. III B).

Relaxations for the (6×6) supercell were performed with a single off symmetry special k point with the coordinate $(1/3, 0, 0)$. With our present computational resources it was not possible to perform test calculations using more k points. For the $(\sqrt{48} \times \sqrt{48})$ supercells, the Γ point only was applied. This single k point is identical to the two special points used for the (4×4) supercell, since the k point $(1/3, 1/3, 0)$ folds back to Γ, when the $(\sqrt{48} \times \sqrt{48})$ supercell is built from the (4×4) surface cell.

C. Thermodynamics

To determine the stability of the surface in contact with the gas phase simple thermodynamic arguments are used. The formalism has been applied to a variety of systems before[22–25] and is described in detail in Ref. 26. Here only a brief summary is given, with a particular emphasis to the problems specific to the present system.

For thermal equilibrium between the gas phase and the surface, the thermodynamic quantity of interest is the surface energy

$$\gamma = \left(G(T,p,\{n_x\}) - \sum_x n_x \mu_x(T,p_x) \right) \bigg/ A, \qquad (1)$$

where T and p are the temperature and the pressure, G is the Gibbs free energy of the solid exposing the surface of interest, n_x are the number of particles x in the solid, and μ_x and p_x are the chemical potentials and the partial pressures of the respective particles in the reservoirs. In the present case, the Gibbs free energies are calculated for a number of reconstructions using finite-sized slabs as described in the preceding section. This implies that the (0001) and $(000\bar{1})$ facets are exposed. For A in Eq. (1), we use the surface area per side (half the total surface area). This implies that the value γ is the sum of the surface energy of the (0001) and $(000\bar{1})$ facets (cleavage energy):

$$\gamma = (\gamma_{(0001)} + \gamma_{(000\bar{1})}).$$

Since the $(000\bar{1})$ surface was identical throughout all calculations, the difference of γ between two calculations is identical to the change of $\gamma_{(0001)}$, which is the desired quantity. Furthermore, zero-point vibrations, vibrational entropy contributions, and enthalpy changes are neglected, and the Gibbs free energies, $G(T,p,\{n_x\})$, are approximated by the energies calculated by density-functional theory (DFT) $E(\{n_x\})$. A recent work indicates that this approximation is sufficiently accurate for oxides and hydroxylated surfaces.[26,27] An estimate for the incurred errors can be obtained by considering the zero-point vibrations of a OH group and a H_2 mol-

ecule. The former has a zero-point vibration energy between 150 and 180 meV, whereas the zero-point vibration energy of H_2 is 260 meV (130 meV per H atom). This suggests that phase boundaries for OH adsorption are wrong by ≈ 50 meV (at zero temperature), but inclusion of zero-point vibrations for hydrogen will cause only a *rigid* shift of the hydrogen potential and of the phase boundaries between hydrogen rich and hydrogen poor terminations. Quite generally, we expect that vibrational entropies will affect the phase boundaries only by the order of 100 meV, which is rather small if one keeps in mind that the present density functionals can be in error by 100 meV/per bond.

In the present case, three different species were considered: Zn, O, and H. To reduce the complexity, it is assumed that the Zn and O particle reservoirs are in thermal equilibrium with bulk ZnO (if this were not the case, the ZnO bulk would either grow or decompose). This requires that the chemical potentials of the Zn and O species satisfy[26]

$$\mu_O + \mu_{Zn} = E_{ZnO}, \qquad (2)$$

where E_{ZnO} is the DFT energy of the ZnO bulk. It is common practice to eliminate μ_{Zn} from the equation for the surface energy, which yields the following expression for the surface energy γ:

$$\gamma = [E(\{n_x\}) - n_H \mu_H - n_{Zn} E_{ZnO} - (n_O - n_{Zn})\mu_O]/A. \quad (3)$$

To determine the most stable surface for a given set of chemical potentials, μ_H and μ_O, γ is calculated for all considered surface reconstructions. The one with the lowest surface energy γ is the thermodynamically stable phase for the corresponding conditions. By repeating these calculations for all sensible sets of chemical potentials μ_H and μ_O, a 2D surface phase diagram can be determined. In this diagram, the two axes correspond to μ_H and μ_O, and each surface reconstruction is assigned one particular gray tone (compare Fig. 2 of Ref. 10, and Fig. 13 of the present work).

In the present work, *all* energies and chemical potentials are referenced to $1/2E_{O_2}$ and $1/2E_{H_2}$, i.e., the energy zero is chosen such that oxygen and hydrogen molecules have zero energy. Using these energy zeros, the chemical potentials can only vary in the following ranges:

$$\mu_O < 0, \quad \mu_H < 0, \quad \text{and} \quad \mu_{Zn} < E_{fcc\,Zn}.$$

If the first or second equation were not fulfilled, oxygen or hydrogen would move from the particle reservoirs to the surface and would start to condensate there. If the third equation were not fulfilled, O particles would flow from the bulk ZnO to the O reservoir leaving back metallic fcc Zn. Eliminating from the third equation μ_{Zn} using Eq. (2) gives

$$E_{ZnO} - E_{fcc\,Zn} < \mu_O.$$

The lower bound for the chemical potential of oxygen is therefore given by the heat of formation of ZnO, which is -3.04 eV in the present case (neglecting zero-point vibrations). In practice, we will restrict the hydrogen potential also to the range -3.0 eV $< \mu_H < 0$ eV in the 2D surface phase diagrams. This range is sensible, when one considers
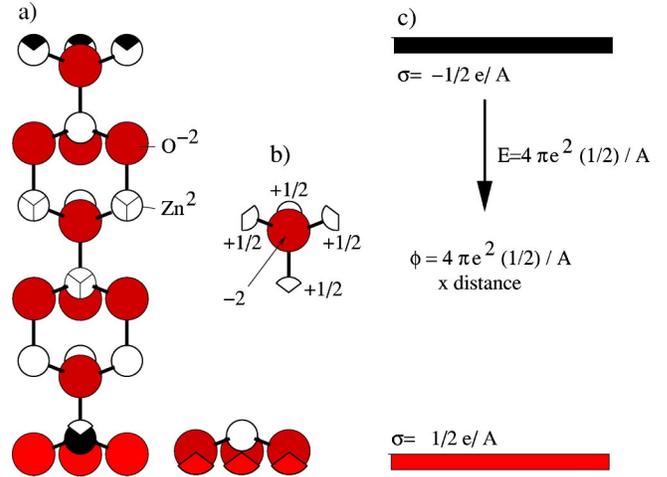

FIG. 2. (a) Sketch of a six-layer slab model of the ZnO (0001) and (000$\bar{1}$) surface, (b) neutral non-polar building unit, and (c) electrostatic field in the ionic model. On the top side, the black triangles cannot be accounted for by concatenation of the nonpolar building unit, whereas at the bottom, the black part and the outermost O atoms cannot be accounted for. This is equivalent to an unaccounted charge of half an electron (bottom middle).

that the chemical potentials can be related to the partial oxygen and hydrogen pressures through

$$\mu_x(T, p_x) = \mu_x(T, p^0) + 1/2 k_B T \ln(p_x/p^0). \qquad (4)$$

Convenient values for the quantities $\mu_x(T, p^0)$ have been tabulated in Refs. 26 and 27, and under typical UHV conditions -3.0 eV $< \mu_H$ is usually fulfilled.

A further restriction is obtained by the restriction that H_2O should not condensate on the surface. This is equivalent to requiring

$$\mu_O + 2\mu_H < E_{H_2O},$$

where the formation energy of water E_{H_2O} is calculated to be -2.64 eV. In the graphs the area where water condensates on the surface will be drawn white.

III. RESULTS

A. The bulk terminated surface

To understand the electronic properties of the ideal as-cleaved ZnO (0001) and (000$\bar{1}$) surfaces it is instructive to consider a simple ionic model.[1] Figure 2 shows a six-layer thick slab exhibiting the two polar surface on the two sides of the slab. Conceptually, one can build the slab by concatenation of a neutral and ''nonpolar'' building unit, which is shown in Fig. 2(b). In fact, the indicated building unit carries a small dipole moment, since the Zn atoms do not have a perfect tetrahedral environment in the wurtzite structure, but for the present purpose we neglect this small deviation from zero polarity. The important point is that repeating this building unit does not account for 1/4 of the Zn^{+2} at the (0001) surface, and 3/4 of the Zn ions and the oxygen ions at the (000$\bar{1}$) surface of the slab. In the ionic

model, these unaccounted parts correspond to an excess surface charge of $(1/2)e$ at the oxygen side of the slab and $(-1/2)e$ at the Zn side of the slab. Solving the Poisson equation for such surface charges (or using the equations of a capacitor) immediately yields a potential difference of $4\pi e/2/A \times d$ between both sides of the slab, where A is the surface area per ZnO surface cell and d is the thickness of the slab. Therefore, in an ionic model the electrostatic potential diverges (Tasker-type three polar surface).[1] In practice, an infinite potential difference between both surfaces is unrealistic, but, as noted by other authors before, the divergence is only an *artifact* of the oversimplistic ionic model and can be removed by transferring half an electron from the O side to the Zn side. Formally, this corresponds to a reduction of the ionicity of the Zn and O ions in the surface layer to $Zn^{+3/2}$ and $O^{-3/2}$. Using such redefined Zn and O surface ionicities, the Madelung energy of the polar ZnO slab is convergent and well defined. It should be emphasized that this "charge transfer" occurs only in the context of the oversimplistic ionic model and it is possible to derive the same formal ionicities in a more natural manner starting from the atomic instead of ionic limit and transferring half an electron along each bond from each Zn atom to neighboring O atoms. Since the Zn atoms and O atoms at the surface have only three O and Zn neighbors, respectively, their ionicity is necessarily only $Zn^{+3/2}$ and $O^{-3/2}$.

An important result of this simple consideration is that *polar surfaces are generally not unstable, their surface energy is not infinite*, but it is true that polar surfaces are often *less favorable* than nonpolar surfaces. For ZnO, the partial ionicity of $\pm 3/2$ implies that electrons at the ideal (0001) Zn side are forced to stay in the conduction band and vice versa, the valence band is not entirely filled at the ideal $(000\bar{1})$ O side, since half an electron is missing. This property is common to most polar surfaces and it is the main reason for their lower stability. Additionally, for ZnO, the fractional occupancy (half an electron and hole, respectively, per Zn and O surface atom) also implies that the Zn and O ideal as-cleaved surfaces are metallic.

The band structures shown in Fig. 3 essentially confirm this simple model. At the Zn side, a very steep surface state exists, which is localized mostly at the Zn atoms in the first layer. The surface state is only partially occupied and holds roughly half an electron. The oxygen side on the other hand exhibits two rather flat surface states, with a total electron hole of 0.5 electrons. In practice, the model presented so far lacks some subtleties that become apparent upon closer inspection of Fig. 3. First, the position of the Zn surface band is such that it partly overlaps with the top of the valence band of layers deeper in the bulk. The reason for this behavior is that the $4s$ Zn surface states penetrate the bulk, leading to a charge transfer from the surface layer into deeper layers. To counteract this penetration a counter field is created shifting the Zn $4s$ conduction bands and the O $2p$ valence bands to higher (less favorable) energies in deeper layers. The evolution of the electrostatic potential in the bulk is shown in Fig. 4, together with an illustration indicating the position of the local conduction and valence bands. The penetration of charge into the surface is realistic, and occurs even in infi-
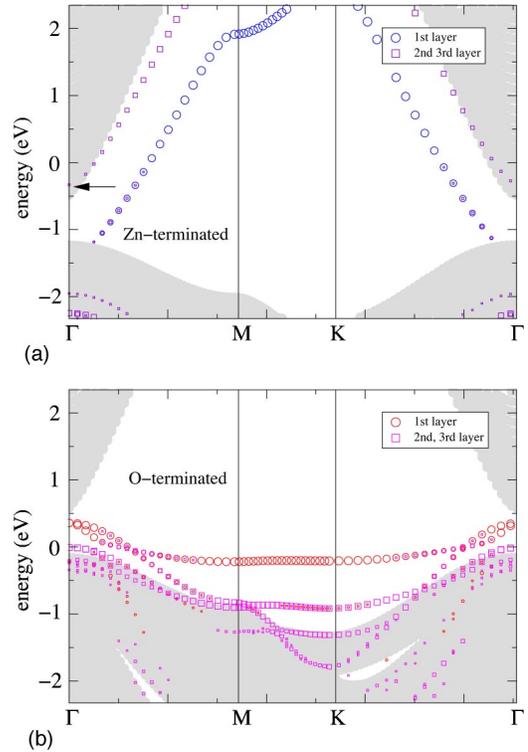

FIG. 3. Band structure of the (0001) Zn-terminated (top) and $(000\bar{1})$ O-terminated (bottom) ZnO surface as calculated for a 26 double-layer slab. The gray shaded areas correspond to the projected band structure of the bulk and have been shifted in such a manner that they represent roughly the positions of the bands at the fourth and fifth ZnO double layer (counted from the surface). Only bands that are strongly localized in the first ZnO double layer or second and third ZnO double layer are shown, and the size of the symbols corresponds to the degree of localization in the respective layers. The Fermi level is located at 0 eV.

nitely thick slabs (the observed behavior is similar to the band bending at semiconductor interfaces). The calculations for thin slabs, however, exhibit an additional artifact. A small amount of charge δQ is also transferred from the Zn side to the O side, in order to align the Fermi levels of both sides. One can estimate the amount of transferred charge assuming that the central ZnO layers behave like a capacitor with a dielectric constant ϵ. The surface charge is then given by (atomic units, i.e., charges are counted as multiples of the electronic charge)

$$\sigma(d) = \frac{\phi\epsilon}{4\pi e^2 d},$$

where d is the slab thickness and ϕ the "initial" potential difference between the local Fermi levels on both sides of the slab (as indicated in Fig. 4). "Initial" here means that long-range charge transfer has not yet occurred, and that the compensation mechanism discussed before is not operative. Using the same model one can also calculate the surface energy gain $\Delta\gamma$ associated with the charge transfer from the electron rich side (Zn-terminated side) to the holes at the electron
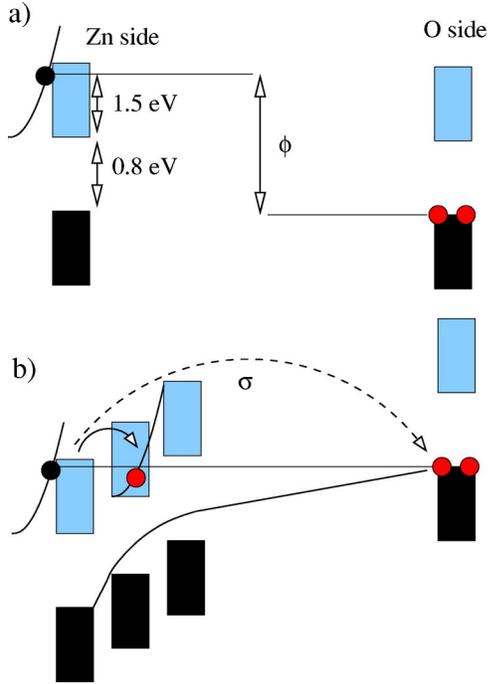
FIG. 5. Cleavage energy as a function of the number of ZnO double layers n in the slab. The present results are shown by triangles and results of Ref. 21 are indicated by squares. Diamonds correspond to the surface energy of a slab, which is terminated with pseudoatoms with a valency 6.5 at the O side (the energy zero is shifted arbitrarily in this case).

FIG. 4. (a) Initial positions of the conduction and valence bands at the ZnO (0001) and (000$\bar{1}$) surfaces, and (b) electrostatic potential in a finite-sized ZnO slab after the charge transfer has occurred. The gray and black bars indicate the local positions of the conduction and valence bands, respectively. The surface band at the Zn side has a free-electron-like behavior and is filled up to 1.5 eV above the bottom of the conduction band. The DFT band gap is roughly 0.8 eV. Electrons are transferred from the Zn-surface state localized in the first layer to deeper double layers. A strong local field (band bending) develops in the vicinity of the Zn surface to counteract the electron transfer. Alternatively electrons can be transferred from the Zn side to the electron holes at the O side (broken arrow), leading to a finite field through the slab.

poor side (O-terminated side). Initially the local Fermi levels have an offset of ϕ. After transfer of a charge σ' the offset has decreased to

$$\phi'(\sigma') = \phi - \frac{1}{\epsilon} 4\pi e^2 \sigma' d.$$

Integration of the energy gain $-\phi'(\sigma')d\sigma'$ over the transferred charge

$$-\int_0^\sigma \phi'(\sigma')d\sigma',$$

yields a surface energy change $\Delta\gamma$ of

$$\Delta\gamma(d) = -\frac{\phi^2\epsilon}{8\pi e^2 d}. \quad (5)$$

It is again emphasized that this energy gain is an artifact of the finite slab, as evidenced by the $1/d$ dependency. In the limit of very thick slabs, the charge transfer must become zero.
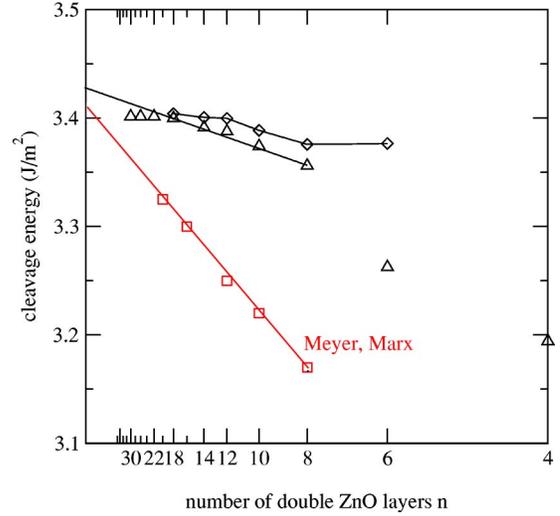
With the previous insight the strong dependency of the cleavage energy on the thickness of the considered slab becomes clear. Results of our present calculations are shown in Fig. 5, together with a comparison to a previous report.[21] For the calculations presented in this section, the oxygen atoms at the (000$\bar{1}$) surface, were not replaced by the pseudoatom with a valency of 6.5. Additionally only two double layers were allowed to relax on the Zn and O side of the slab. The first noticeable observation is that the previous calculations of Meyer and Marx exhibit a much steeper slope than the present calculations. The reason for this, at first sight, unexpected behavior is that only the two outermost double layers on both sides of the slab were allowed to relax in the present calculations, whereas Meyer and Marx allowed for relaxation of all layers.[21] To obtain estimates for the slope of the line, the ion-clamped static dielectric function ϵ^∞ must be applied in Eq. (5) in the present case, whereas the ionic contribution to the dielectric function has to be included, when the central bulklike part is allowed to respond. For ZnO the static dielectric function is $\epsilon_{33}^\infty = 4.4$, whereas the ionic contribution is roughly two times larger,[28] explicating the difference in slope.

Despite this difference, both calculations clearly converge to similar values for the surface energy. We find a value of 3.4 J/m^2 in excellent agreement with the previous GGC calculations.[21] Second, our curve deviates from the expected linear behavior for very thick slabs ($n = 22,26,30$). We have carefully checked that this is not an artifact of some calculational parameters. The reason for the deviation is explained by the observation that for *thin* slabs, only a small penetration of the Zn surface states into the bulklike region is observed, and that the alignment of the local-Fermi levels of the Zn and O sides is achieved mainly by a charge transfer from one side of the slab to the other (dashed arrow in Fig.

4). Therefore, Eq. (5) correctly models the dependency of the cleavage energy on the slab thickness. For thick slabs, however, the surface states at the Zn side penetrate the bulk efficiently, with the band marked by an arrow in Fig. 3 being mostly responsible for this. For up to 18-layer thick slabs, this band is entirely above the Fermi level, but it becomes progressively populated when the slab thickness is further increased. Hence for thick slabs, the alignment of the local Fermi levels of both surfaces is mostly achieved through a *local* band bending at the Zn-terminated side of the slab, and the correction formula, Eq. (5), does not apply. For thin slabs however, it yields quantitatively sensible corrections to the cleavage energy: the local-Fermi levels of the Zn and O side differ by roughly 2.3 eV (compare Fig. 4). Using a dielectric constant of 4.4, an energy correction of 0.052 J/m^2 is obtained for the eight-layer thick slab. This would give a corrected cleavage energy of 3.41 J/m^2 which is within 0.01 J/m^2 of the converged value.

Slightly faster convergence with respect to the slab thickness is achieved by filling the electron holes at the O-terminated side. In the present work, we achieved this by replacing the outermost O atoms at the $(000\bar{1})$ side by a special pseudoatom with a valency of 6.5. The surface energies are now indeed converging reasonably rapidly with the thickness of the slab, as shown in Fig. 5. The slab thickness was fixed to eight ZnO double layers for the remainder of this work, as already stated in Sec. II B. Clearly the surface energy reaches only a plateau for six to eight double layers, and it would have been desirable to perform calculations for 12 double layers. But with the present computational resources it was impossible to model large scale reconstructions with 12 double layers. A few additional convergence tests, illustrating that energy differences are reasonably well converged, will be presented below.

B. Zn vacancies

Since removal of Zn atoms from the surface decreases the number of electrons in the surface state, it should enhance the stability of the surface. A removal of every fourth Zn atom, for instance, in a regular (2×2) pattern, would entirely deplete the surface state. Figure 6 summarizes the results for the average Zn vacancy formation energy. To this end, the vacancy formation energy is defined as the heat of formation of the following reaction:

$$\text{bulk terminated slab} + n\text{O}_2/2 \Rightarrow \text{slab}/n \text{ vac} + n\text{ZnO},$$
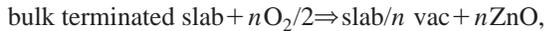
which is equivalent to calculating the formation energy using

$$E_{\text{form}} = (E_{\text{slab}/n\text{vac}} + nE_{\text{ZnO}} - E_{\text{slab}} - n/2E_{\text{O}_2})/n.$$

This corresponds to the oxygen rich limit ($\mu_O = 0$), since it implies that oxygen is abundantly available in the gas phase. Figure 6 shows that the formation of vacancies leads to a substantial energy gain of the order of 3.2 eV. But the repulsion between the vacancies is substantial; already at 0.25 ML, the average vacancy formation energy has decreased to roughly 2.5 eV. This is more clearly expressed through the energy gain for adding vacancies one by one in the (4×4) supercell. The first vacancy has a formation energy of
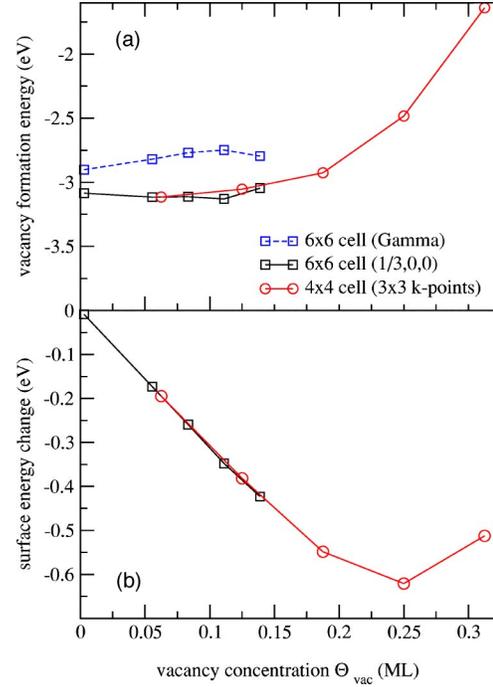

FIG. 6. (a) Average Zn vacancy formation energy and (b) change of the surface energy for the (0001) surface of ZnO.

-3.2 eV, for the second and third the energy changes are -2.9 and -2.65 eV. For the fourth vacancy the energy gain is only -1.15 eV, and addition of a fifth vacancy costs already 1.7 eV. The same result can be expressed through the surface energy γ in the oxygen rich limit ($\mu_O = 0$). The surface energy specifies the energy change per surface area and is related to the average vacancy formation energy (which specifies the energy gain *per vacancy*) through the simple relation

$$nE_{\text{form}} = \Delta \gamma \times \text{surface area},$$

or equivalently

$$\Delta \gamma = E_{\text{form}} \frac{n}{\text{surface area}} = E_{\text{form}} \theta_{\text{vac}},$$

where θ_{vac} is the vacancy concentration per surface area. Again it is clearly visible that the surface energy drops until the vacancy concentration reaches 0.25 ML. Then the surface energy increases again. The calculations are clearly consistent with the picture that Zn vacancies progressively deplete the surface state, which is also confirmed by inspection of the local density of states at the surface (not shown here).

Another point that deserves attention is the large dependency of the vacancy formation energy on the Brillouin-zone sampling in the low coverage limit. For the (6×6) supercell, Γ point only calculations show a significant error of the order of 0.3 eV. But even using a single off symmetry k point $(1/3, 0, 0)$, the calculations are probably not yet fully converged; at low coverage, the initial increase of the vacancy formation energy is most likely only an artifact of the yet still insufficient k-point sampling. The slow convergence

TABLE I. Average vacancy formation energy and average energy gain for addition of an oxygen adatom in eV for the (4×4) supercell for two special k points $(0,0,0)$ and $(1/3,1/3,0)$, and for the Monkhorst-Pack grid 3×3.

Θ	Zn vacancy two k points	Zn vacancy 3×3	O adatom two k points	O adatom 3×3
1/16	-3.375	-3.115	-3.618	-3.422
2/16	-3.072	-3.054	-3.280	-3.270
3/16	-2.941	-2.926	-3.039	-3.026
1/4	-2.508	-2.484	-2.551	-2.527

with k points is a result of the steepness of the Zn surface state. In the low coverage limit, electrons are removed from the steepest section of this band (compare Fig. 3), and converged energies are therefore only obtained using a very accurate k point sampling corresponding to roughly 12×12 points in the Brillouin zone of the primitive surface cell. Once the steep part of the band has been depleted, k point sampling becomes a less critical issue. Typically, meshes corresponding to 8×8 k points in the Brillouin zone of the primitive surface cell are sufficient, when the surface state is half depleted. As a further illustration, the results for the vacancy formation energy and the O adsorption energy (cf. following section) at various coverages are summarized in Table I. At 1/16 ML concentration, the errors in the adsorption energies are clearly large with two k points (200 meV), but they fall off rapidly at higher coverage (20 meV). Additionally, it is noted that the energy differences between Zn vacancies and O adatoms at a particular coverage are roughly four times smaller than the absolute changes of these energies.

Additional convergence tests with respect to the slab thickness were performed as well. For a vacancy concentration of $\theta=1/4$, the vacancy formation energy was -2.528 eV, and the energy gain for addition of an O adatom to the surface was -2.567 eV (for 3×3 k points). These values agree within 40 meV with the eight-layer slab calculations. It is important to emphasize that errors are largest for the ideal nonstoichiometric surface, since this surface exhibits a strong band bending at the surface (see Sec. III A). For nonpolar locally stoichiometric surfaces, convergence with slab thickness is in fact much more rapidly attained, and *energy differences* between surface reconstructions with *similar stoichiometry* converge fast with slab thickness towards their final value (e.g., the energy difference between a Zn vacancy and an O adatom changes only by 4 meV from 8 to 12 layers). Overall, we therefore expect that the energy differences between different reconstructions in the relevant regime ($\theta_{vac}\approx0.25$) are converged to 20 meV for the present setup (k points and slab thickness), which is significantly better than the expected DFT precision.

C. O adatoms

An alternative process to the removal of Zn atoms is the addition of neutral oxygen atoms to the Zn surface. The oxygen atoms will take up the electrons from the surface state in
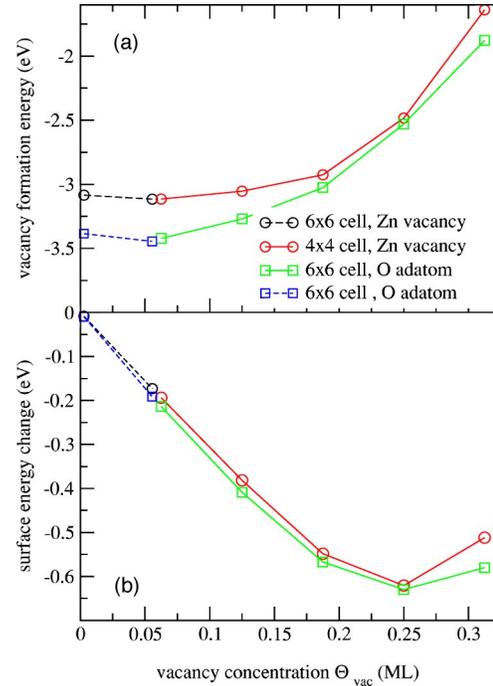

FIG. 7. (a) Average oxygen adsorption energy (compared to the average vacancy formation energy) and b) corresponding change of the surface energy for the (0001) surface of ZnO.

order to form O^{-2} ions. An insulating surface is again obtained when 1/4 ML of oxygen is adsorbed at the ideal (0001) surface. To allow for a direct comparison with the vacancy formation energy, we define the energy as the heat of formation of the following reaction:

$$\text{slab}+n\text{O}_2/2\Rightarrow\text{slab}/n\text{O},$$
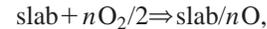
which is equivalent to calculating the adsorption energy using

$$E_{ads}=(E_{slab/n\text{O}}-E_{slab}-n/2E_{\text{O}_2})/n.$$

For the adsorption of oxygen two different sites were considered. The first one is atop Zn atoms, i.e., the conventional bulk continuation of the oxide, which will be termed T4 (tetrahedral fourfold coordinated, the coordination the site would have if the bulk were continued). In addition, the threefold hollow (H3) site with no Zn atoms beneath was considered (see Fig. 8). This site is found to be 2 eV more stable than the T4 site at 1/4 ML. Results for the H3 site, and a comparison to the results of the preceding section are shown in Fig. 7. It is discernible that oxygen adatoms are more stable at any relevant coverage. Particularly at low coverage, they exceed the stability of isolated vacancies by almost 350 meV. At higher coverage the energy difference is smaller indicating a stronger repulsion between the O adatoms, but O adatoms become again significantly more favorable at a coverage of 5/16 ML. This has geometrical reasons, since for two neighboring vacancies, four O atoms at the surface have two broken O-Zn bonds [highlighted by white circles in Fig. 8(a)]; obviously a rather unfavorable situation.
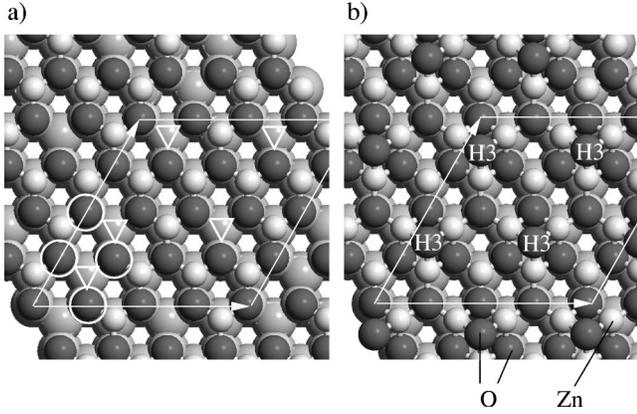
FIG. 8. (4×4) surface cell with (a) five vacancies (positions highlighted by triangles) and (b) four O adatoms adsorbed at the H3 site. For each vacancy, one O-Zn bond is broken at each corner of the small triangles. The four encircled O atoms have two broken bonds each.

D. Adsorption of OH groups and hydrogen

The excess electrons in the conduction band can be removed by the adsorption of neutral OH groups as well. When they combine with one electron from the surface state, OH^- groups are formed on the surface. At a coverage of 0.5-ML OH, a rather stable situation is expected, since then the surface state has been depleted entirely. To evaluate the adsorption energy as a function of the coverage, the average adsorption energy is defined as the heat of formation of the following reaction:

$$\text{slab} + n O_2/2 + n H_2/2 \Rightarrow \text{slab}/n OH,$$

which is equivalent to calculating the adsorption energy using

$$E_{\text{ads}} = (E_{\text{slab}/n OH} - E_{\text{slab}} - n/2 E_{O_2} - n/2 E_{H_2})/n.$$

We also considered hydrogen adsorption at the surface, defining the adsorption energy with respect to molecular hydrogen:

$$E_{\text{ads}} = (E_{\text{slab}/n H} - E_{\text{slab}} - n/2 E_{H_2})/n.$$

Various adsorption sites for OH groups and H were considered. As for oxygen adatoms, the most stable site for OH groups is the H3 site. Hydrogen atoms, on the other hand, prefer to adsorb atop the Zn atoms. The average adsorption energies are displayed in Fig. 9. For the present definition of the OH adsorption energy, OH adsorption is clearly more favorable than H adsorption. The energy difference amounts to more than 3 eV at most coverages implying that hydrogen adsorbed at the Zn atoms is thermodynamically unstable, since the reaction

$$n \text{ZnO} + \text{slab}/n H \Rightarrow \text{slab}/n OH + n \text{fcc Zn}$$
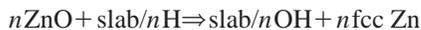
is exothermic. Expressed in another manner: on the Zn side, adsorbed hydrogen is so reactive that it is capable to reduce bulk ZnO units to metallic Zn, forming OH groups. The kinetic barriers for the formation of OH group might be rather
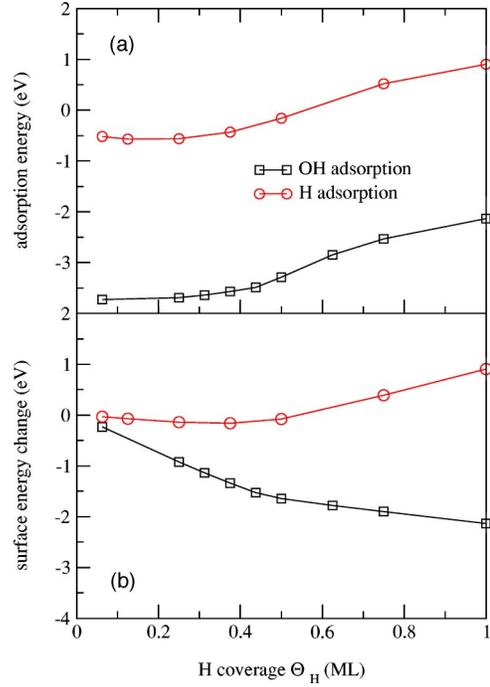

FIG. 9. (a) Average OH adsorption energy and H adsorption energy (with respect to molecular oxygen and hydrogen) and (b) corresponding change of the surface energy for the (0001) surface of ZnO.

large, however, inhibiting the formation of OH groups under certain experimental conditions (low temperature). The calculations are nevertheless, consistent with the experimental observations that hydrogen adsorption is exothermic on the Zn site, and that prolonged exposure to hydrogen destroys the lateral order of the ZnO crystal:[9] the formation of the thermodynamically stable OH groups requires oxygen mass transport on or close to the surface which possibly abolishes order. Disagreement with experiment exists, however, as to how much hydrogen can be adsorbed at the surface. According to the present calculation, the saturation coverage is 0.5 ML for molecular H whereas in experiment indications for an ordered (1×1) overlayer were observed.[9]

Returning to Fig. 9, it is noted that the surface energy shows a rather clear change of the slope between 0.4 and 0.5 ML OH and H coverage. This corresponds to the point where the surface state becomes depleted, and it is an indication that the binding mechanism changes from a binding with an ionic contribution (OH^-) to a less favorable purely covalent bonding (OH) at higher coverage.

E. Triangular reconstructions

Triangular reconstructions were modeled by removing triangular shaped areas from the eight-layer thick slab. In the (4×4) supercell, triangles with an edge length of $n=2$ and $n=3$ were considered, in the (6×6) supercell triangles with an edge length of $n=3-5$ were calculated, and for the $(\sqrt{48} \times \sqrt{48})$ supercell, triangles with an edge of $n=6-7$ were constructed. Since a large fraction of the first double surface layer is removed for the triangular shaped reconstruc-
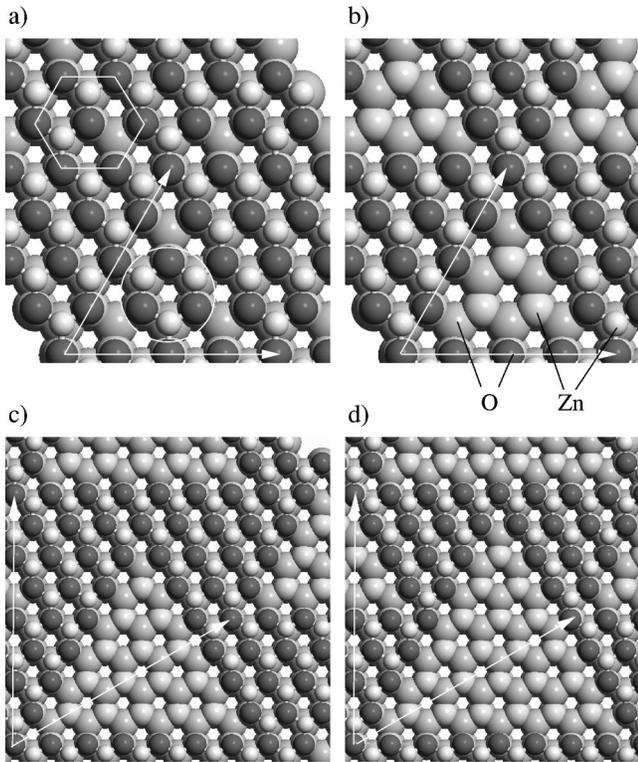
FIG. 10. (4×4) surface cell with (a) three isolated vacancies and (b) a triangular pit with an edge length $n = 3$. $(\sqrt{48} \times \sqrt{48})$ supercell with a triangular pit of (c) edge length $n = 6$ and (d) edge length $n = 7$.

tions, it is essential that the surface energies are identical for seven- and eight-layer thick slabs, since otherwise the formation energies of triangles might be overestimated or underestimated. Figure 5 shows that slabs with six, seven, and eight double layers have almost identical surface energies (plateau), when the charge transfer to the $(000\bar{1})$ surface is inhibited by replacing the O atoms by pseudoatoms with a valency of 6.5 at the $(000\bar{1})$ surface. Nevertheless, convergence tests for the (4×4) cell and a triangle with $n = 3$ were performed. When the number of layers was increased from 8 to 12 double layers, the energy difference between the clean and triangularly reconstructed surface changed by 67 meV for the entire slab. This corresponds to a relative change of 67 meV/16 = 4 meV per primitive surface cell. The energy difference between the clean surface and three isolated vacancies changed by roughly the same number (58 meV), indicating that energy differences between surfaces with similar stoichiometry (see below) are essentially converged for the present setup.

Figures 10(a) and 10(b) show these two reconstructions, the one with three isolated Zn vacancies and the triangle with $n = 3$, respectively. For this particular triangle, six Zn atoms and three O atoms have been removed, with the removed Zn (O) atoms located above the second layer O (Zn) atoms. In general, $n \times (n+1)/2$ Zn atoms and $n \times (n-1)/2$ O atoms are removed for a triangle with an edge length n, which corresponds to
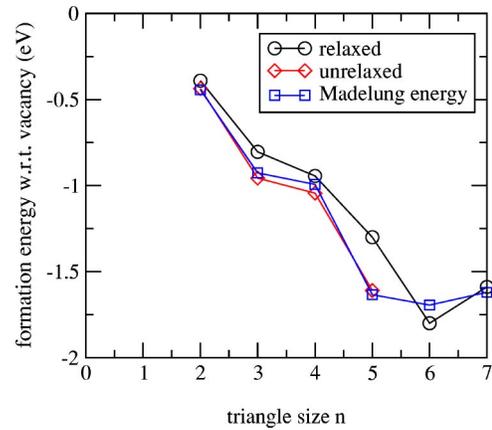
$$n \times (n+1)/2 - n \times (n-1)/2 = n$$

FIG. 11. Formation energy of triangular pits with respect to isolated vacancies, calculated using the fully relaxed DFT structures (circles), using DFT and unrelaxed structures (diamonds), and using a Madelung-type calculation for the unrelaxed geometries (squares).

Zn vacancies from a stoichiometric point of view.[10] For the triangle with an edge length $n = 3$, this is particularly obvious: removal of the encircled three Zn and three O atoms in Fig. 10(a) leaves the local stoichiometry of the surface unaffected and allows us to go from isolated vacancies to the triangular pit.

Furthermore, it is possible to prove that the number of cleaved bonds is identical for three isolated vacancies, and a triangular shaped pit with an edge length $n = 3$. We can show this by first realizing that the number of outermost Zn atoms is exactly identical in both cases. For the three isolated Zn vacancies in the (4×4) cell, $16 - 3 = 13$ inequivalent Zn atoms per surface area are visible. For the triangular pit [Fig. 10(b)], $16 - 6 = 10$ Zn atoms are exposed in the outermost double layer, and three more are visible in the second double layer, which yields again a total of 13 Zn atoms. For each visible Zn atom, one bond has been cleaved. Furthermore, for the nine oxygen atoms which are located next to the three vacancies, one bond has been cleaved. For the triangular reconstruction, the nine oxygen atoms at the edge of the pit possess one cleaved bond each, as well. This concludes the proof for $n = 3$, and the argument can be extended readily to triangles with an arbitrary side length. In general, a triangular pit with side length n possess the *same surface stoichiometry* as n isolated vacancies, and additionally both exhibit exactly *the same number of cleaved bonds*. Since the number of broken bonds is identical, the covalent contribution to the surface energies should be similar for n isolated Zn vacancies and for a triangular indentation with side length n.

The calculated formation energies of triangular pits with respect to isolated vacancies are shown in Fig. 11. The formation energy is defined as

$$E_{\text{form}} = E_{\text{slab/triangle } n} - E_{\text{slab}/n\text{vac}} + n(n-1)/2 E_{\text{ZnO}}.$$

It is a measure of how much energy is gained, when n vacancies combine to form a triangular indentation. Of course, in this process ZnO units must be removed from the slab as well, which is accounted for by adding $n(n-1)/2 E_{\text{ZnO}}$ to the formation energy. At the real ZnO surface, this involves mass

transport, and the removed units might attach to steps, or they might form triangles at the surface.

The calculated formation energies are roughly proportional to the side length of the triangle n. The calculated constant of proportionality implies that 300 meV are gained, whenever an additional vacancy is incorporated into a triangle in turn increasing its edge length by one. The linear dependency also means that six isolated vacancies can either combine to form two triangles with a side length 3 or a single triangle with a side length 6. In the former case, the energy gain is 2×0.805 eV≈ 1.6 eV, whereas in the latter, it is only slightly larger (1.8 eV). If the formation energies were lying exactly on a straight line, many small triangles would have exactly the same energy as few large triangles. As we will see later, the present calculations seem to favor the second scenario, i.e., rather large triangles with additional triangular features inside. Nevertheless, the energetic competition is very tight, which explicates why the surface becomes macroscopically rough in the experiment. In practice, the largest formation energy per vacancy (i.e., the value one obtains after division by the number of vacancies) is obtained for the triangle with side length $n=6$. The drop for the triangle with side length $n=7$ is probably related to the proximity of triangles for the ($\sqrt{48} \times \sqrt{48}$) supercell used for this case [compare Fig. 10(d)].

The intriguing question is the reason for the stability of triangles versus isolated vacancies. Since differences in the covalent bonding can be ruled out on the basis of the introductory arguments of this section, only two possible answers remain. Either the triangles allow for a more efficient surface relaxation or electrostatic arguments explicate the observed behavior. To test the first hypotheses the energies of isolated vacancies and triangular pits were calculated without surface relaxation and are summarized in Fig. 11 using diamonds. Clearly without relaxation the formation energies of triangles are close to those obtained with relaxation, although it should be noted that the individual relaxation energies are large and of the order of several eV, but they are almost identical for n isolated vacancies and a triangle with edge length n. To test the hypothesis that the Madelung energy is responsible for the formation of triangles, the Madelung energies of the unrelaxed triangular pits with edge length n and n isolated vacancies were evaluated and compared. At first sight, this is not a trivial undertaking, since one encounters again the problem that the Madelung energy is not convergent in the ionic model, i.e., the surface energies do depend on the thickness of the considered slab. For the ideal bulk terminated surface, this problem can be circumvented by assigning an ionicity of $+3/2$ and $-3/2$ to the outermost Zn and O ions, respectively (compare Sec. III A). To model the defective surfaces, the ionicity of the outermost Zn atoms must be set to $1.5/(1+\theta_{Zn})$, which ensures that the Madelung energy remains well defined. Here θ_{Zn} is the effective Zn deficiency, as already defined in Ref. 10. For the two models shown in Fig. 10, the Zn deficiency is, for instance, $\theta_{Zn}=-3/16$, which implies that the ionicity of the 12 Zn atoms visible in Figs. 10(a) and 10(b) is 1.84 (the reduced ionicity is also assigned to the visible Zn atoms in the second double layer). Furthermore, a dielectric constant of 5.5 was used in the calculation of the Madelung energies, where we regard this constant as a simple fitting parameter. The energy differences between triangles and vacancies obtained in this manner are shown in Fig. 11 using squares. The calculated points lie almost exactly on top of the energy differences obtained in the DFT calculations (using the same unrelaxed configurations). This is certainly to some degree fortuitous, since the applied model is necessarily an oversimplification. But certainly it is the Madelung energy that stabilizes the triangles. A much simplified explanation for the stability of triangles can be given by inspecting the first coordination shell of undercoordinated oxygen atoms [see Fig. 10(a)]. The number of Zn neighbors is identical around vacancies and triangles. But on the unreconstructed surface, oxygen atoms in the proximity of vacancies have six repulsive oxygen neighbors in the second coordination shell, whereas this number is reduced to four or five when triangles form on the surface.

F. Electronic properties

Figure 12(a) shows the layer resolved electronic density of states (DOS) of the ideal (0001) surface. It is emphasized that the DOS presented in this section was obtained with each O atom at the ($000\bar{1}$) surface replaced by an atom with a valency of 6.5, in order to fill the surface state at the oxygen side. Charge transfer from the Zn to the O side is therefore inhibited, and only local band bending is operative. As a reference the density of states in the S-2 layer is used, which exhibits little modifications compared to the bulk. As a result of the band bending, the onset of the valence band and the positions of the core levels are at 1 eV larger binding energies in the S layer than in the S-2 layer (more negative). In the surface layer, the general structure and shape of the DOS at the oxygen atoms changes little compared to deeper layers, whereas the Zn d band has narrowed somewhat as a result of the reduced coordination. The most obvious feature, however, is the free-electron parabola stemming from the free-electron 2D-surface state. The Fermi-level cuts through this parabola and even in the S-2 double layer a sizable contribution from this surface state is still visible. It is certain that the metallic surface state should be visible in STS and would allow for a clear identification of the ideal surface. Also the local band bending would have a finger print in core-level spectroscopy, but neither the metallic surface state nor the local band bending has been observed,[9,10,29] which again rules out the existence of the unreconstructed surface.

For the OH covered surface [Fig. 12(b)], the surface state has been depleted entirely, and the surface is essentially insulating. This confirms that adsorption of OH creates a low-lying electron acceptor level allowing the depletion of the Zn 4s band at the surface. At the surface, no band bending is found and the Zn and O derived states are similar to those in the bulk. Fingerprints of the OH group are the OH σ peak in valence band (in DFT located -7 eV below the top of the valence-band edge), and the larger O 1s core-level binding energy of the oxygen atoms in the OH group. Therefore in experiments, the identification of a OH covered surface should be straightforward.
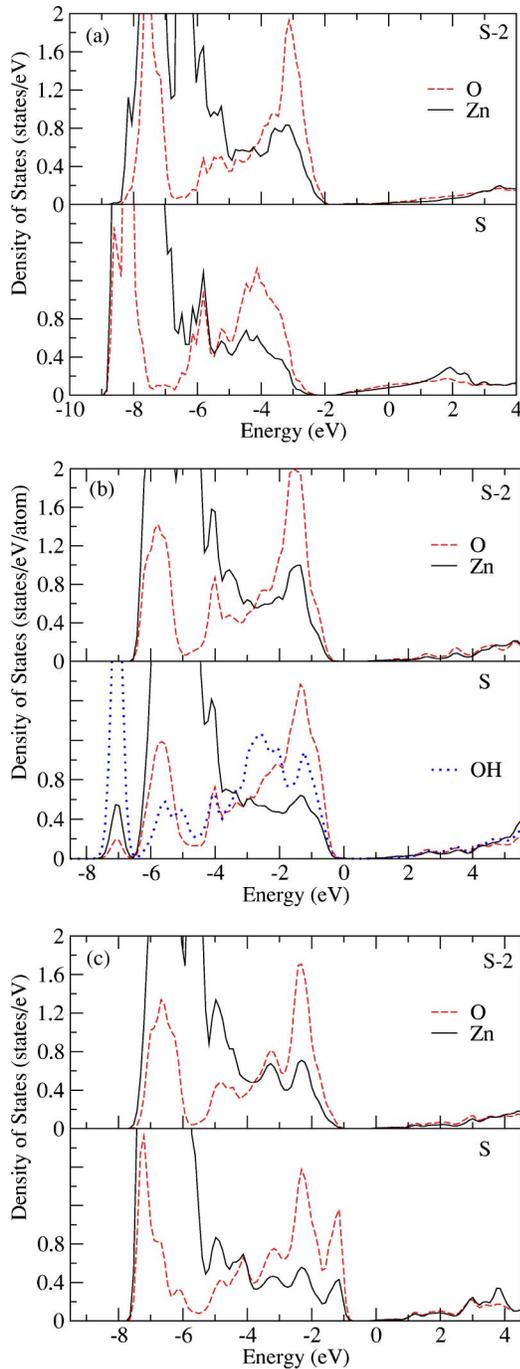
FIG. 12. (a) Electronic density of states (DOS) of the ideal as-cleaved (0001) surface of ZnO. The electronic states are shown for the first double layer (S) and the third double layer (S-2). (b) DOS of the OH covered surface (0.5 ML). (c) DOS of a triangularly reconstructed surface for a triangle with a side length $n=2$ in a (3×3) supercell. The Fermi levels are located at 0 eV. The plots have been shifted such that the edges of the valence bands are aligned for S-2.

For the triangular reconstruction, a (3×3) supercell with a single small triangle with a side length $n=2$ was considered (the DOS does not depend significantly on the size of the triangles). This corresponds to a rather high Zn deficiency of $\theta_{Zn}=-0.222$ ML and the size of the cell still al-

lows for a very accurate k-point sampling. For this triangular reconstruction, the surface state is almost entirely depleted, and the Fermi level is located at the edge of the conduction band. The local band bending is fairly weak; the onset of the valence band and the core levels are shifted by roughly 0.3 eV to higher binding energies in the surface layer.

G. Surface relaxation

Table II summarizes the changes of the interlayer distances with respect to the bulk interlayer spacing. For the ideal bulk terminated surface, a contraction of the Zn-O distance by almost 33% is found inside the first double layer. A similar but smaller contraction is obtained in deeper double layers as well, whereas the distance between the double layers increases slightly. This oscillatory behavior agrees well with previous calculations.[21] The OH and O covered surfaces, however, exhibit only a small relaxation of the interlayer distances. Adsorbed species therefore lift the surface relaxation very efficiently.

On the other hand, with vacancies, the surface shows an even stronger inward relaxation of the first Zn layer than the ideal defect-free surface. The Zn atoms sink into the first oxygen layer. For deeper layers, again only little relaxation is found, which indicates that the surface becomes bulklike rapidly. A similar pattern is observed for the triangularly reconstructed surfaces, if all Zn atoms visible at the surface are counted as first layer Zn atoms. The topmost Zn atoms again sink into the oxygen layer by 0.3 Å, whereas little relaxation is found in deeper layers.

These results clearly show that the metallic 2D surface state—if occupied—has a strong influence on the surface relaxation. For the metallic surface, the oscillatory relaxation is a consequence of the penetration of the surface state into deeper layers. When electrons are removed from the surface state, the surface becomes bulklike after the first ZnO double layer.

The available experimental data are somewhat difficult to interpret using the theoretical results. Most experiments indicate a slight outward relaxation.[8,30,31] This seems to be consistent only with a surface where O or OH groups are adsorbed, since the triangularly reconstructed surfaces show a strong inward relaxation. In the work of Maki et al.,[31] the surface was prepared in flowing oxygen gas, and triangular features were absent in atomic force microscopy. One might speculate that their surface was covered by OH groups or oxygen adatoms (oxygen adatoms might be kinetically stable). The x-ray-diffraction data of Ref. 8 might require a reanalysis considering triangular reconstructions, but it cannot be ruled out that the surface was OH or O covered, as well.

H. The phase diagram

The final phase diagram combining all results of the present work is shown in Fig. 13. It is important to note that this diagram differs quantitatively somewhat from the one reported previously, since a more precise computational setup was used in the present case. The eight-layer calculations generally favor higher vacancy concentrations, but ad-

TABLE II. Average interlayer distance $d_{O_n Zn_m}$ in Å and change relative to the bulk in % of the bulk interlayer spacing for the ideal, OH covered, and O covered surface, for Zn vacancies and for the reconstructed surfaces. The entry $3@(4\times4)$ corresponds to a triangle with side length 3 in the (4×4) supercell, and $3,7@(\sqrt{48}\times\sqrt{48})$ is the structure shown in the inset of Fig. 13 at $\theta_{Zn}=-0.21$. The bulk interlayer distances are $d_{Zn_n O_n}=0.644$ Å in the double layer and $d_{O_n Zn_{n+1}}=2.010$ Å between double layers.

	Ideal	(1/2)-ML OH	(1/4)-ML O
$Zn_1 O(H)$		1.262	1.223
$Zn_1 O_1$	0.430 (-33.3)	0.663 ($+3.0$)	0.606 (-5.9)
$O_1 Zn_2$	2.146 ($+6.7$)	2.003 (-0.4)	2.027 ($+0.8$)
$Zn_2 O_2$	0.528 (-18.1)	0.648 ($+0.6$)	0.624 (-3.1)
$O_2 Zn_3$	2.089 ($+3.9$)	2.000 (-0.5)	2.010 (-0.0)
$Zn_3 O_3$	0.568 (-11.9)	0.639 (-0.8)	0.626 (-2.8)
$O_3 Zn_4$	2.064 ($+2.7$)	2.004 (-0.3)	2.009 (-0.1)
	(1/4)-ML vacancy	$3@(4\times4)$	$3,7@(\sqrt{48}\times\sqrt{48})$
$Zn_1 O_1$	0.272 (-57.7)	0.284 (-55.8)	0.318 (-51.1)
$O_1 Zn_2$	2.042 ($+1.6$)	2.044 ($+1.7$)	2.020 ($+0.5$)
$Zn_2 O_2$	0.652 ($+1.1$)	0.637 (-1.1)	0.673 ($+4.4$)
$O_2 Zn_3$	2.024 ($+0.7$)	2.057 ($+2.3$)	2.036 ($+1.3$)
$Zn_3 O_3$	0.644 (-0.1)	0.606 (-5.9)	0.627 (-2.7)
$O_3 Zn_4$	2.027 ($+0.8$)	2.045 ($+1.7$)	2.031 ($+1.0$)

ditionally more structural models were considered in the present study. In the previous calculations, for instance, only combinations of triangular indentations and Zn vacancies were explored, whereas in the present case triangular pits and O adatoms at the terraces were taken into account as well. The general features are however identical to those reported before. At H-rich conditions, OH groups, adsorbed at the otherwise unreconstructed substrate, are *thermodynamically* most stable. This however does not preclude that the formation of OH groups is possibly kinetically hindered. To form OH groups, water or hydrogen molecules must dissociate at the surface, and additional oxygen atoms must stem from either oxygen diffusing to the surface or oxygen molecules dissociating at the surface. Such processes might involve significant barriers. Nevertheless, the energetics are clear, and they favor the formation of OH groups over a wide regime. Typical room-temperature UHV conditions, for instance, correspond to 300 K and 10^{-10} mbar. At these conditions the oxygen and hydrogen chemical potentials are roughly $\mu_O\approx-0.65$ eV and $\mu_H\approx-0.54$ eV.[26] This implies that the surface should become hydroxylated even in UHV, if it has sufficient time to take up residual water or hydrogen from the gas phase.

Under hydrogen-poor conditions, however, the surface exhibits a completely different morphology,[10] with triangular reconstructions being stable. Preparation conditions of $T>800$ K and $p<10^{-10}$ mbar correspond to $\mu_O<-1.9$ eV and $\mu_H<-1.55$ eV, which is well within the regime where triangles form according to our calculations.

The precise structural morphology depends on the oxygen potential. For oxygen-poor conditions ($\theta_{Zn}=-7/48\approx-0.15$), the stable structure is a triangular pit with side length $n=6$ and an additional O adatom adsorbed in the center of the triangular indentation at an H3 site. At $\theta_{Zn}=-8/48\approx-0.17$, the stable structure is again a triangle

with a side length $n=6$, and two oxygen adatoms adsorbed at H3 sites [see Fig. 14(a)]. Both structures have not been considered in the previous work (and it is admittedly possible that other structures have still been overlooked). At $\theta_{Zn}=-10/48\approx-0.21$, the stable structure is a triangle with an edge length of $n=7$, with another triangular indentation with $n=3$ inside (again a previously unconsidered model). In this case, even the second ZnO double layer has been removed partially. The second triangle is inverted with respect to the first one, which is a consequence of the ZnO stacking sequence: to remove more Zn than O atoms in the second double layer, the triangles must be rotated by 180° with respect to the triangles in the topmost layer. This is fully consistent with experimental results, where small pits were always to form within lower pits.[10] A further increase of the Zn concentration is possible by adding one O atom in the center of the small triangle with side length $n=3$ [see Fig. 14(b)] and corresponds to the final structure at $\theta_{Zn}=-0.23$. Several other plausible configurations were also considered, for instance adsorption of O adatoms at various other H3 sites, but they were energetically slightly less favorable. The complexity and the size of the supercell precludes an exhaustive search of the configuration space with the present computational resources (the eight-layer $\sqrt{48}\times\sqrt{48}$ supercell contains more than 700 atoms). Nevertheless, the present phase diagram is expected to describe the experimental situation reasonably well.

Two important issues should be highlighted at this point. The calculations give a rather clear indication that larger triangles are preferred. But large triangles are necessarily linked to a low Zn deficiency. The smallest supercell for which a triangle with an edge length $n=7$ can be realized is the $\sqrt{48}\times\sqrt{48}$ supercell. Nevertheless, this corresponds only to a rather small Zn deficiency of $\theta_{Zn}=-7/48\approx-0.15$. To
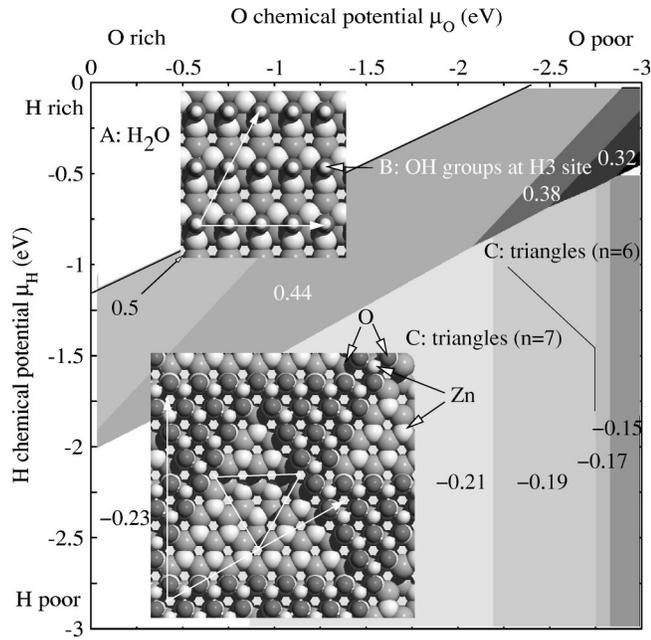
FIG. 13. Phase diagram of the ZnO(0001)-Zn surface in equilibrium with H and O particle reservoirs controlling the chemical potentials μ_H and μ_O. Preparation conditions of $T > 800$ K and $p < 10^{-10}$ mbar correspond to $\mu_O < -1.9$ eV and $\mu_H < -1.55$ eV. The upper left area A indicates the conditions under which H_2O condensates on the surface. The dark and bright areas correspond to a OH covered surface (B) and triangular reconstructions (C), respectively. Positive numbers indicate the OH coverage θ_{OH} and negative numbers the effective Zn deficiency θ_{Zn} per surface area. The two insets show optimized structures corresponding to the area. The basis vectors of the supercells are indicated by arrows.

FIG. 14. Stable ($\sqrt{48} \times \sqrt{48}$) structures at (a) $\theta_{Zn} = -0.17$ and (b) $\theta_{Zn} = -0.23$ (see text).

atoms, the triangles are, however, no longer thermodynamically stable, and the reconstruction is lifted.

IV. DISCUSSION

The combination of simple thermodynamic arguments with energies determined by *ab initio* density-functional theory can contribute significantly to our understanding of surface reconstructions. To date, such calculations have been limited to one-dimensional cuts through the phase diagram; for instance, usually only variations in the oxygen potential were considered. The present study (and two previous publications, Refs. 10 and 27) extends this approach to two independent extensive thermodynamic variables: the chemical potential of oxygen and hydrogen. For the surfaces of oxides, which are usually in contact with water, it is clearly necessary to include the influence of hydrogen, since oxide surfaces are easily passivated by the presence of hydroxyl groups. This has been realized before, for instance, in the pioneering work of Wang, Chaka, and Scheffler, in which only the adsorption of hydrogen could explicate the observed relaxation of the Al_2O_3 (0001) surface.

The situation for ZnO is rather similar. When hydrogen is present, the surface tends to be unreconstructed, and OH groups adsorb at the H3 site, where the oxygen atoms are coordinated to three surface Zn atoms. The typical OH coverage is 1/2 ML, which is exactly the amount of OH required to deplete the Zn 4s surface state. For the ideal surface, this surface state is occupied by half an electron, since electrostatics require that the Zn surface atoms have a formal valency of $+3/2$ at the ideal (0001) surface. When (1/2)-ML OH adsorbs, the surface energy is significantly lowered, because the adsorption of OH leads to the formation of a low-lying acceptor level, which takes up the "excess" electrons from the Zn 4s conduction band. Formally a simple ionic picture captures the calculated behavior: upon adsorption, OH becomes negatively charged (OH^-), and the Zn surface atoms reach their full ionicity of $+2$.

An alternative process to the adsorption of OH groups is the addition of O atoms to the surface and/or the removal of Zn atoms from the surface. The present calculations indicate that the addition of O adatoms is energetically slightly more favorable than the removal of Zn atoms. In the ionic picture,

achieve higher Zn deficiencies either the density of triangles has to increase—which can be achieved only, when the triangles simultaneously shrink—or inverted triangles have to form in the next ZnO layer. For instance, a triangle with an edge length $n = 3$ can be realized in a (4×4) supercell corresponding to a Zn deficiency of $\theta_{Zn} = -3/16 \approx -0.187$, and an even smaller triangle with $n = 2$ can be realized in a 3×3 supercell. The latter one corresponds already to a Zn deficiency $\theta_{Zn} = -2/9 \approx -0.222$. The *ab initio* calculations give a relatively clear indication that smaller triangles (in particular, those with a side length 2) are unfavorable. Hence larger triangles with inverted triangles in the next ZnO layer are preferred over a dense packing of smaller triangles.

The second important point is that the present calculations suggest that OH groups *cannot* adsorb at the reconstructed surface. In Ref. 10, the area C of Fig. 2 corresponded to conditions, at which hydroxyl groups adsorbed on the terraces and in the triangular pits. In the present case, these structures were thermodynamically not stable. The main reason for this discrepancy is that we considered more structures in the present study, in particular, combinations of triangles and oxygen atoms adsorbed at H3 sites. These structures compete with OH groups adsorbed at the surface, and are slightly more stable as long as the hydrogen potential is not sufficiently high. Once the chemical potential of hydrogen becomes sufficiently large to hydrogenate the O ad-
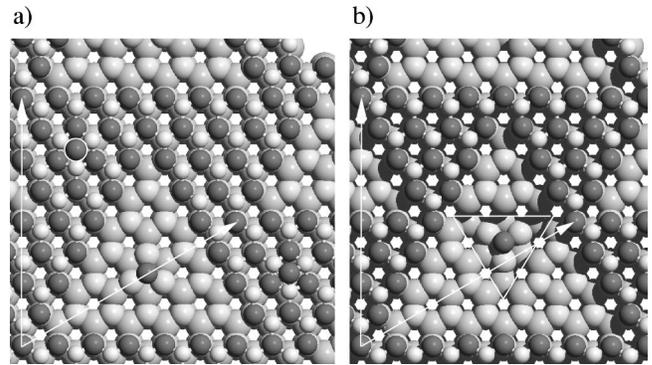
1/4 ML of O should correspond to the energetically most stable situation, since in this case the Zn surface atoms reach again their formal ionicity of $+2$ (and the adsorbed O atom a formal ionicity of -2). Indeed, the calculations agree with these simple arguments: the adsorption of (1/4)-ML oxygen leads to the lowest surface energy at oxygen rich conditions.

In practice, however, triangular reconstructions are even more stable than isolated oxygen adatoms or Zn vacancies, in agreement with recent STM measurements.[10] The calculations indicate that larger triangles are generally preferred over small triangles. To reach the required Zn deficiency of 0.25 ML, large triangles must however possess additional features inside, which were modeled by removing triangles rotated by 180° from the second ZnO double layer. It is found that such an arrangement has a lower surface energy than many small densely packed triangles. The energetic competition between triangles of different size is found to be very tight, explicating the observed macroscopic roughening of the surface in the experiment. The reason for the stability of triangular reconstructions was shown to arise from simple electrostatic arguments, with the Madelung energy of regularly arranged vacancies being significantly higher than that of triangles.

Finally, we have discussed the dependency of the cleavage energy on the thickness of the slab, which is an important issue for polar oxide surfaces lacking inversion symmetry or a mirror plane. It is found that the cleavage energy depends strongly on the thickness of the slab, since a small amount of charge tends to be transferred from one side of the slab to the other. The effect is a genuine finite-size effect, and it is highly undesirable in the present case. A simple analytical formula to model the finite-size effect was suggested, showing that the surface energy is lowered by

$$\gamma = -\frac{\phi^2 \epsilon}{8\pi e^2 d},$$

for a finite slab with thickness d. The magnitude of the finite-size effect depends quadratically on the misalignment of the local position of the Fermi levels ϕ at the two sides of the slab and is proportional to the dielectric constant of the bulk oxide ϵ. Although not discussed in this work, the magnitude of the effect might be strong enough to change the stability of surface reconstructions of thin films compared to bulk oxides.

V. CONCLUSION

Depending on the hydrogen (and to a lesser extent) on the oxygen pressure, the ZnO surface exhibits either triangular shaped reconstructions or remains unreconstructed but covered by hydroxyl groups. The effective Zn vacancy concentration of the surface is found to be between 0.23 and 0.20 under realistic conditions. The concentration of hydroxyl groups on the unreconstructed surface is roughly twice as large as the effective Zn vacancy concentration, i.e., under typical conditions 0.38–0.5 ML OH. From a catalytic point of view the regime where OH groups and triangular reconstructions compete in stability seems to be particularly interesting. The present work forms the basis for further studies involving catalytic properties insofar that realistic microscopic models for the (0001) Zn surface are now available.

ACKNOWLEDGMENTS

*Electronic address: georg.kresse@univie.ac.at

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

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

[3] A. Wander and N.M. Harrison, J. Chem. Phys. **115**, 2312 (2001).

[4] 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).

[5] W. Göpel, J. Pollmann, I. Ivanov, and B. Reihl, Phys. Rev. B **26**, 3144 (1982).

[6] R.T. Girard, O. Tjernberg, G. Chiaia, S. Soederholm, U.O. Karlsson, C. Wigren, H. Nylen, and I. Lindau, Surf. Sci. **373**, 409 (1997).

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

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

[9] Th. Becker, St. Hövel, M. Kunat, Ch. Boas, U. Burghaus, and Ch. Wöll, Surf. Sci. **486**, L502 (2001).

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

[11] W. Kohn and L. Sham, Phys. Rev. **140**, A1133 (1965).

[12] R.O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).

[13] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[14] M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, and J.D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).

[15] G. Kresse and J. Hafner, Phys. Rev. B **48**, 13 115 (1993).

[16] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996); G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11 169 (1996).

[17] P.E. Blöchl, Phys. Rev. B **50**, 17 953 (1994).

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

[19] Y. Wang and J.P. Perdew, Phys. Rev. B **44**, 13 298 (1991).

[20] J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).

[21] B. Meyer and D. Marx, Phys. Rev. B **67**, 035403 (2003).

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

[23] P. Raybaud, J. Hafner, G. Kresse, S. Kasztelan, and H. Toulhoat, J. Catal. **189**, 129 (2000); **190**, 128 (2000).

[24] X.-G. Wang, A. Chaka, and M. Scheffler, Phys. Rev. Lett. **84**, 3650 (2000).

[25] G. Kresse, S. Surnev, M.G. Ramsey, and F.P. Netzer, Surf. Sci. **492**, 329 (2001).

[26] K. Reuter and M. Scheffler, Phys. Rev. B **65**, 035406 (2002).

[27] Q. Sun, K. Reuter, and M. Scheffler, Phys. Rev. B **67**, 205424 (2003).

[28] N.A. Hill and U. Waghmare, Phys. Rev. B **62**, 8802 (2000).

[29] M. Kunat, S.G. Girol, T. Becker, U. Burghaus, and Ch. Wöll, Phys. Rev. B **66**, 081402 (2002).

[30] M. Sambi, G. Granozzi, G.A. Rizzi, M. Casari, and E. Tondello, Surf. Sci. **319**, 149 (1994).

[31] H. Maki, N. Ichinose, N. Ohashi, H. Haneda, and J. Tanaka, Surf. Sci. **457**, 377 (2000).