Native point defects in ZnO

Anderson Janotti and Chris G. Van de Walle
Materials Department, University of California, Santa Barbara, California 93106-5050, USA
(Received 18 April 2007; published 4 October 2007)

We have performed a comprehensive first-principles investigation of native point defects in ZnO based on density functional theory within the local density approximation (LDA) as well as the LDA+U approach for overcoming the band-gap problem. Oxygen deficiency, manifested in the form of oxygen vacancies and zinc interstitials, has long been invoked as the source of the commonly observed unintentional n-type conductivity in ZnO. However, contrary to the conventional wisdom, we find that native point defects are very unlikely to be the cause of unintentional n-type conductivity. Oxygen vacancies, which have most often been cited as the cause of unintentional doping, are deep rather than shallow donors and have high formation energies in n-type ZnO (and are therefore unlikely to form). Zinc interstitials are shallow donors, but they also have high formation energies in n-type ZnO and are fast diffusers with migration barriers as low as 0.57 eV; they are therefore unlikely to be stable. Zinc antisites are also shallow donors but their high formation energies (even in Zn-rich conditions) render them unlikely to be stable under equilibrium conditions. We have, however, identified a different low-energy atomic configuration for zinc antisites that may play a role under nonequilibrium conditions such as irradiation. Zinc vacancies are deep acceptors and probably related to the frequently observed green luminescence; they act as compensating centers in n-type ZnO. Oxygen interstitials have high formation energies; they can occur as electrically neutral split interstitials in semi-insulating and p-type materials or as deep acceptors at octahedral interstitial sites in n-type ZnO. Oxygen antisites have very high formation energies and are unlikely to exist in measurable concentrations under equilibrium conditions. Based on our results for migration energy barriers, we calculate activation energies for self-diffusion and estimate defect-annealing temperatures. Our results provide a guide to more refined experimental studies of point defects in ZnO and their influence on the control of p-type doping.

DOI: 10.1103/PhysRevB.76.165202 PACS number(s): 61.72.−y, 71.55.Gs, 61.72.Bb, 66.30.−h

I. INTRODUCTION

Understanding the behavior of native point defects is essential to the successful application of any semiconductor. These defects often control, directly or indirectly, doping, compensation, minority carrier lifetime, and luminescence efficiency. They also assist the diffusion mechanisms involved in growth, processing, and device degradation.[1–3] Doping forms the basis of much of semiconductor technology and can be drastically affected by native point defects such as vacancies, self-interstitials, and antisites. Such defects may cause *self-compensation*: for instance, in an attempt to dope the material p type, certain native defects which act as donors may spontaneously form and compensate the deliberately introduced acceptors. In ZnO, specific native defects have long been believed to play an even more important role. As-grown ZnO frequently exhibits high levels of unintentional n-type conductivity, and native point defects have often been invoked to explain this behavior. Oxygen vacancies and zinc interstitials have most often been mentioned as sources of n-type conductivity in ZnO.[4–15] Nevertheless, most of the arguments have been based on circumstantial evidence, in the absence of unambiguous experimental observations. The availability of higher-quality bulk crystals and epitaxial layers has contributed to significant progress in the experimental observation of point defects in the last few years.[16–24] However, the impact of individual defects on the electronic properties of ZnO is still a subject of much debate.

Besides knowing their electronic properties, it is also important to know how native point defects migrate in the crystal lattice. Knowledge of migration of point defects greatly contributes to the understanding of their incorporation during growth and processing, and it is essential for modeling self-diffusion and impurity diffusion, which is nearly always mediated by native defects. Information about atomic diffusion or migration of point defects in ZnO is currently limited. Neumann has summarized experimental results for self-diffusion in ZnO up to 1981.[25] Activation energies of zinc self-diffusion were reported to be in a range from 1.9 to 3.3 eV, while activation energies for oxygen self-diffusion were reported to span a much wider range, from 1.5 to 7.5 eV. Interpreting these results or using them in a predictive manner is not straightforward. The activation energy for self-diffusion (Q) is the sum of the formation energy of the defect that mediates the self-diffusion and its migration energy barrier:[26]

$$Q = E^f + E_b. \tag{1}$$

The migration energy barrier E_b is a well defined quantity, given by the energy difference between the equilibrium configuration and the saddle point along the migration path, and can be obtained with good accuracy from first-principles calculations.[27–31] The first term in the activation energy, namely, the formation energy of the defect (E^f), however, strongly depends on the experimental conditions, such as the position of the Fermi level and the zinc or oxygen chemical potentials. These parameters can cause large changes (by several eV) in the formation energy. It is usually not straightforward to assess how the environmental conditions affect

the formation energy and hence the activation energy for self-diffusion. This explains the wide spread in the reported values and makes it difficult to extract values from experiment.

In addition to self-diffusion measurements, investigating the behavior of point defects through annealing can also provide valuable information about their migration.[17,23,32–35] The point defects in such experiments are often deliberately introduced into the material through nonequilibrium processes such as electron irradiation or ion implantation. Once point defects are introduced, they can be identified by their optical, electronic, or magnetic responses (signatures). These signatures are then monitored as a function of annealing temperature. Changes in defect signatures at a given annealing temperature indicate that the relevant defects have become mobile. In principle, one can perform a systematic series of annealing experiments at different temperatures, extract a time constant for the decay of the signal at each temperature, and then perform an Arrhenius analysis. In the absence of such elaborate studies, estimates for activation energies can still be obtained by performing an analysis based on transition state theory.[36] All of these assume, of course, that the observed changes in defect signatures are solely related to defect migration and do not involve any other processes such as formation of complexes, etc.

For all these reasons, systematic first-principles studies of both migration and formation energies for all relevant defects are very useful. As we will show, the calculated values can be used to interpret results from self-diffusion measurements and annealing experiments. A number of first-principles calculations for point defects in ZnO, based on density functional theory (DFT) and the local density approximation (LDA) or the generalized gradient approximation (GGA), have been reported in the last few years.[37–45] However, interpretation of the results of these calculations is not straightforward. Calculations based on DFT-LDA or DFT-GGA suffer from the well known band-gap error. In the case of ZnO, the calculated band gap using LDA is only 0.8 eV, compared to the experimental value of 3.4 eV. This error leads to large uncertainties in the calculated position of the defect-induced states in the band gap and, consequently, in the defect formation energies and transition levels, precluding direct predictions or comparisons with experimental values. Various attempts have been made to correct formation energies and transition levels given by LDA. Some of these corrections are based on empirical reasoning or on a rigid shift of the conduction band;[37–41] others are based on non-self-consistent approaches,[38,42] LDA+U,[43,44] or calculations using a hybrid functional.[45] These various approaches have led different groups to qualitatively different conclusions about the role of individual point defects in ZnO.[37–45]

A comprehensive analysis of the electronic and structural properties of native point defects requires a systematic, quantitative, and self-consistent approach for correcting formation energies and transition levels. The reason LDA severely underestimates the band gap of ZnO is partially related to the underestimation of the binding energy of the zinc semicore 3d states.[46] In this work, we use the LDA+U approach to correct for the position of the zinc d states in ZnO. As discussed in more detail elsewhere,[46,47] this leads to a systematic improvement in the description of the physical properties of the system. The LDA+U approach also produces a partial correction of the band gap. By combining self-consistent calculations based on LDA and LDA+U, we are able to calculate the dependence of transition levels on the theoretical band gap and extrapolate the results to the experimental band gap. The transition levels are thus corrected according to their valence- and conduction-band character in a consistent and quantitative fashion, distinguishing the approach from previous studies based on LDA+U.[43,44] Based on this formalism, we investigate the electronic and structural properties of all native point defects in ZnO. We explore the local atomic relaxations and their effect on the electronic structure of each defect. We also report results for previously unexplored configurations and present a comprehensive study of migration barriers. Finally, we discuss the influence of native defects on the control of n- and p-type doping.

This paper is organized as follows. In Sec. II, we discuss the theoretical framework for studying native point defects in semiconductors. We describe the calculation of formation energies, transition levels, and the correction scheme based on the LDA and LDA+U calculations. In Sec. III, we focus on the results for native point defects in ZnO. We report the calculated formation energies and discuss the electronic and geometric structure. We devote particular attention to the oxygen vacancy, which has frequently, but incorrectly, been invoked as the source of n-type conductivity in ZnO. We also report the results for diffusion mechanisms and the corresponding migration barriers for interstitials and vacancies. Section IV summarizes our results.

II. THEORETICAL APPROACH

Our first-principles investigations are based on density functional theory within the local density approximation.[48] The DFT-LDA approach allows for calculations of total energies of solids, molecules, and atoms. Lattice parameters, atomic positions, and forces (derivatives of total energy with respect to atomic positions) can be computed quite accurately;[49] the structural parameters are typically predicted to within a few percent of the experimental values. By using appropriate boundary conditions, such as supercells with periodic boundary conditions, it is also possible to investigate the electronic and local lattice structure of defects in solids, search for equilibrium atomic positions of defects and the surrounding host atoms, and calculate migration energy barriers and migration paths. Formation energies of defects and impurities can be derived directly from total energies, allowing calculation of equilibrium concentrations.[30]

A. Technical details

We use the pseudopotential method to separate the chemically inert core electrons from the chemically active valence electrons. Semicore states, such as the zinc 3d states which are less than 10 eV below the valence-band maximum in ZnO, need to be treated explicitly as valence electrons. In the present work, we use the projector-augmented-wave pseudo-

potentials as implemented in the VASP code.[50,51] The defects are simulated by adding (removing) host atoms to (from) a 96-atom supercell with periodic boundary conditions. We use a plane-wave basis set with a 400 eV cutoff and a $2\times2\times2$ mesh of special k points for integrations over the Brillouin zone. Tests as a function of plane-wave cutoff and k-point sampling show that our results for neutral defects are numerically converged to within 0.1 eV; somewhat larger errors may occur for charged defects due to defect-defect interactions in neighboring supercells.

B. Defect concentrations

The concentration of a point defect depends on its formation energy. In thermodynamic equilibrium and in the dilute regime (i.e., neglecting defect-defect interactions), the concentration of a point defect is given by

$$c = N_{\text{sites}} \exp\left(-\frac{E^f}{k_B T}\right), \tag{2}$$

where E^f is the *formation energy*, N_{sites} the number of sites the defect can be incorporated on, k_B the Boltzmann constant, and T the temperature. Equation (2) shows that defects with *high* formation energies will occur in *low* concentrations. Here, we neglect the contributions from the formation volume and the formation entropy. The former is related to the change in the volume when the defect is introduced into the system; it is negligible in the dilute regime and tends to become important only under high pressure. The formation entropy is related mainly to the change in the vibrational entropy. Formation entropies of point defects are typically of the order of a few k_B, and therefore become important only at very high temperatures. Moreover, vibrational entropy contributions largely cancel out when comparing different defects or assessing solubilities.[30]

C. Formation energy of defects

The formation energy of a point defect is not a constant but depends on the growth or annealing conditions.[30] For example, the formation energy of an oxygen vacancy is determined by the relative abundance of Zn and O atoms, as expressed by the chemical potentials μ_{Zn} and μ_{O}, respectively. If the vacancy is charged, the formation energy further depends on the Fermi level (E_F), which is the energy of the electron reservoir, i.e., the electron chemical potential. Forming an oxygen vacancy requires the removal of one oxygen atom; the formation energy is therefore

$$E^f(V_{\text{O}}^q) = E_{tot}(V_{\text{O}}^q) - E_{tot}(\text{ZnO}) + \mu_{\text{O}} + q(E_F + E_v), \tag{3}$$

where $E_{tot}(V_{\text{O}}^q)$ is the total energy of a supercell containing the oxygen vacancy in the charge state q, $E_{tot}(\text{ZnO})$ is the total energy of a ZnO perfect crystal in the same supercell, and μ_{O} is the oxygen chemical potential. Expressions similar to Eq. (3) apply to all native point defects.

The chemical potential μ_{O} depends on the experimental growth conditions, which can be either Zn rich, O rich, or anything in between. It should therefore be explicitly regarded as a variable in the formalism. However, in thermodynamic equilibrium, it is possible to place *bounds* on the chemical potential. The oxygen chemical potential μ_{O} is subject to an upper bound given by the energy of O in an O_2 molecule, $E_{tot}(O_2)$, corresponding to extreme O-rich conditions; spin polarization is included in the energy of the molecule. The calculated bond length of the O_2 molecule of 1.22 Å is in good agreement with the experimental value of 1.21 Å.[52] Similarly, the zinc chemical potential μ_{Zn} is subject to an upper bound given by the energy of Zn in bulk zinc [$\mu_{\text{Zn}}=E_{tot}(\text{Zn})$], corresponding to extreme Zn-rich conditions. It should be kept in mind that μ_{O} and μ_{Zn}, which are free energies, are temperature and pressure dependent.

The upper bounds defined above also lead to lower bounds given by the thermodynamic stability condition for ZnO,

$$\mu_{\text{Zn}} + \mu_{\text{O}} = \Delta H_f(\text{ZnO}), \tag{4}$$

where $\Delta H_f(\text{ZnO})$ is the enthalpy of formation of bulk ZnO (negative for a stable compound). The upper limit on the zinc chemical potential then results in a lower limit on the oxygen chemical potential: $\mu_{\text{O}}^{min}=E_{tot}(O_2)+\Delta H_f(\text{ZnO})$. Conversely, the upper limit on the oxygen chemical potential results in a lower limit on the zinc chemical potential: $\mu_{\text{Zn}}^{min}=E_{tot}(\text{Zn})+\Delta H_f(\text{ZnO})$. The calculated enthalpy of formation of ZnO is $\Delta H_f(\text{ZnO})=-3.5$ eV, compared to the experimental value of −3.6 eV.[53] The host chemical potentials thus vary over a range corresponding to the magnitude of the enthalpy of formation of ZnO.

The Fermi level E_F in Eq. (3) is not an independent parameter, but is determined by the condition of charge neutrality. In principle, equations such as Eq. (3) can be formulated for every native defect and impurity in the material; the complete problem, including free-carrier concentrations in valence and conduction bands, can then be solved self-consistently, imposing charge neutrality. However, it is instructive to examine formation energies as a function of E_F in order to examine the behavior of defects when the doping level changes. We reference E_F with respect to the valence-band maximum E_v and allow E_F to vary from 0 to E_g, where E_g is the fundamental band gap. Note that the valence-band maximum E_v is taken from a calculation of a perfect-crystal supercell, corrected by the alignment of the electrostatic potential in the perfect-crystal supercell and in a region far from the defect in the supercell containing the defect, as described in Ref. 30. No additional corrections to address interactions between charged defects are included here. It has become clear that the frequently employed Makov-Payne correction[54] often significantly overestimates the correction,[30,55] to the point of producing results that are less accurate than the uncorrected numbers. In the absence of a more rigorous approach, we feel that it is better to refrain from applying poorly understood correction schemes.

D. Defect transition levels

Defects often introduce levels in the band gap of semiconductors;[1,2] these levels involve transitions between different charge states of the same defect and can be derived from the calculated formation energies. The transition levels

are not to be confused with the Kohn-Sham states, which from now on we call *states*, that result from band-structure calculations. The transition level $\varepsilon(q/q')$ is defined as the Fermi-level position for which the formation energies of charge states q and q' are equal. $\varepsilon(q/q')$ can be obtained from

$$\varepsilon(q/q') = \frac{E^f(D^q;E_F=0) - E^f(D^{q'};E_F=0)}{q'-q}, \qquad (5)$$

where $E^f(D^q;E_F=0)$ is the formation energy of the defect D in the charge state q when the Fermi level is at the valence-band maximum ($E_F=0$). The experimental significance of this level is that for Fermi-level positions below $\varepsilon(q/q')$, charge state q is stable, while for Fermi-level positions above $\varepsilon(q/q')$, charge state q' is stable. Transition levels can be observed in experiments where the final charge state can fully relax to its equilibrium configuration after the transition, such as in deep-level transient spectroscopy (DLTS).[1,56] Transition levels correspond to thermal ionization energies. Conventionally, if a defect transition level is positioned such that the defect is likely to be thermally ionized at room temperature (or at device operating temperatures), this transition level is called a shallow level; if it is unlikely to be ionized at room temperature, it is called a deep level. Note that shallow centers may occur in two cases: first, if the transition level in the band gap is close to one of the band edges [valence-band maximum (VBM) for an acceptor, conduction-band minimum (CBM) for a donor]; second, if the transition level is actually a *resonance* in either the conduction or valence band. In that case, the defect necessarily becomes ionized, because an electron (or hole) can find a lower-energy state by transferring to the CBM (VBM). This carrier can still be Coulombically attracted to the ionized defect center, being bound to it in a "hydrogenic effective-mass state." This second case coincides with what is normally considered to be a "shallow center" (and is probably the more common scenario). Note that in this case, the hydrogenic effective-mass levels that are experimentally determined are *not* directly related to the calculated transition level, which is a resonance above (below) the CBM (VBM).

E. Local-density-approximation correction

Unfortunately, LDA seriously underestimates band gaps and therefore corrections are necessary to compare the calculated transition levels $\varepsilon(q/q')$ with experimental results. In the case of ZnO, the error is particularly severe because LDA underestimates the binding energy of the zinc d electrons. The zinc d states form a narrow band that overlaps with the valence band in tetrahedrally bonded II-VI semiconductors. In ZnO, the zinc d states couple with the oxygen p states that form the top of the valence band. This coupling pushes the top of the valence band upward, reducing the band gap.[57] This effect explains why ZnO has a smaller band gap than ZnS, despite being more ionic and having a smaller lattice constant: the lower-lying oxygen p states experience a stronger repulsion from the zinc d states than the sulfur p states.

In the LDA, the underbinding of the d states causes the p-d coupling to be unphysically strong, resulting in too large a reduction of the band gap. One way to correct for the underestimation of the binding energy of the Zn d electrons is by including an on-site Coulomb correlation interaction, as implemented in the LDA+U method.[58] In this approach an orbital-dependent correction that accounts for electronic correlations in the narrow bands derived from Zn d states is added to the LDA potential. This on-site Coulomb correlation corrects the position of the narrow bands derived from the Zn d states and affects both the valence-band maximum and conduction-band minimum.[46] It opens the band gap by shifting the valence band downward and the conduction band upwards. The LDA+U thus provides a partial correction of the band gap, the remaining correction arising from the intrinsic LDA error in predicting the position of the conduction band. An important problem when applying the LDA+U approach is how to choose the value of U. We have proposed an approach[46] to obtain U entirely from first principles by calculating U for the atom and screening this value by using the high-frequency dielectric constant ε^∞ of the solid. Further discussion and justification of this approach is contained in Ref. 46. For ZnO, our procedure gives $U=4.7$ eV for the semicore Zn d states. The difference in enthalpy of formation ΔH_f(ZnO) of ZnO between LDA and LDA+U is only 0.07 eV. Note that in the LDA+U calculation of ΔH_f(ZnO), we take $U=0$ for bulk zinc, consistent with our procedure of obtaining U as the atomic value divided by the high-frequency dielectric constant.

Although the LDA+U approach only corrects part of the band-gap error, it provides us with a basis for obtaining a full band-gap correction through a physically justified extrapolation scheme. Point defects usually induce electronic states in the band gap of semiconductors and insulators. These states may be (fully or partially) occupied with electrons or empty, and the occupation of these states determines the charge state of the defect. The defect states exhibit a mixture of conduction-band and valence-band character. Since the LDA underestimates band gaps, the position of the defect states (with respect to the valence-band edge) also tends to be underestimated. The greater its conduction-band character, the greater the error in the position of the defect state. As a consequence, the LDA also underestimates the formation energy of defects that have occupied states in the band gap. Again, the error in the formation energy is expected to increase with the degree of conduction-band character in the defect state.

Our approach takes advantage of the fact that the extent to which transition levels $\varepsilon(q/q')$ change in going from LDA to LDA+U reflects their relative valence-band and conduction-band character. The procedure is to perform calculations using the LDA, on the one hand, and the LDA+U, on the other hand, and then extrapolate to the experimental gap:

$$\varepsilon(q/q') = \varepsilon(q/q')^{\mathrm{LDA}+U} + \frac{\Delta\varepsilon}{\Delta E_g}(E_g^{expt} - E_g^{\mathrm{LDA}+U}), \qquad (6)$$

with

$$\frac{\Delta\varepsilon}{\Delta E_g} = \frac{\varepsilon(q/q')^{\mathrm{LDA}+U} - \varepsilon(q/q')^{\mathrm{LDA}}}{E_g^{\mathrm{LDA}+U} - E_g^{\mathrm{LDA}}}. \qquad (7)$$

Here, E_g^{LDA} and $E_g^{\mathrm{LDA}+U}$ are the band gaps given by the LDA and LDA+U approximations and E_g^{expt} is the experimental

TABLE I. Calculated lattice parameters a, c/a, and u, band gap E_g, and enthalpy of formation ΔH_f for ZnO using LDA and LDA $+U$. Experimental data were taken from Refs. 59 and 61.

	LDA	LDA$+U$	Expt.
a (Å)	3.195	3.148	3.249
c/a	1.615	1.612	1.602
u	0.379	0.379	0.381
E_g (eV)	0.80	1.51	3.43

gap.[59] The coefficient $\frac{\Delta\varepsilon}{\Delta E_g}$ is the rate of change in the transition levels with respect to the change in the band gap, and it is obtained by performing calculations using LDA and LDA$+U$ [Eq. (7)]. The correction is in the spirit of schemes discussed by Zhang *et al.*[38] using different methodologies. It is important to note that this extrapolation is valid when the valence- and conduction-band character of the defect (Kohn-Sham) state does not change when the gap is corrected, and only the position of the state is shifted in the gap. The extrapolated values for the transition levels are very close to results obtained using self-interaction and relaxation-corrected pseudopotentials,[60] enhancing confidence in the approach.

The corrections for transition levels enable us to also apply corrections to formation energies. The procedure for doing so is more easily explained in the context of specific examples, after presenting some results for transition levels, and therefore we postpone this discussion to Sec. III B.

III. RESULTS AND DISCUSSION

A. ZnO perfect crystal

In Table I, we list the calculated equilibrium lattice parameters and band gaps using the LDA and LDA$+U$ for ZnO. These calculations were performed for a four-atom primitive cell of the wurtzite structure and a $4\times4\times4$ Monkhorst-Pack special k-point mesh. For a detailed discussion of the effects of the on-site Coulomb correlation energy U on the electronic and structural properties of ZnO, we refer the reader to Ref. 46. The defect calculations discussed in the next sections were performed at the theoretical equilibrium lattice parameters in the respective approximation (LDA or LDA$+U$). Use of the theoretical lattice parameters is necessary in order to avoid spurious effects in the atomic relaxations. Although the absolute value of the atomic displacements may differ between LDA and LDA$+U$, we find good agreement when displacements are expressed relative to the appropriate equilibrium Zn-O bond length.

B. Native point defects in ZnO

In this work, we consider all the possible native point defects in ZnO: oxygen and zinc vacancies (V_O and V_{Zn}), interstitials (O_i and Zn_i), and antisites (O_{Zn} and Zn_O). Before giving detailed results for each point defect we discuss the LDA correction for defect transition levels and formation energies.

We illustrate the correction based on the LDA and LDA $+U$ calculations with the detailed example of the oxygen vacancy. Figure 1 shows the calculated formation energies for the oxygen vacancy using LDA and LDA$+U$. The removal of an oxygen atom from the lattice breaks four bonds. The four "dangling bonds" on the surrounding Zn atoms combine to form a fully symmetric a_1 state in the band gap, and three almost degenerate states located above the CBM. In the neutral charge state of the oxygen vacancy, the a_1 state is occupied with two electrons. The other three states above the CBM are always empty; we do not need to consider them further. In LDA, the a_1 state is close to the VBM. As the band gap is corrected, the a_1 state is shifted upward, as expected for a state that has significant Zn (and hence conduction-band) character. Since the a_1 state is occupied with two electrons, this shift increases the formation energy of the oxygen vacancy in the neutral charge state with respect to the positive charge states, as illustrated by comparing the LDA and LDA$+U$ calculations in Fig. 1. Because the LDA$+U$ affects both the valence band and the conduction band, the shift in the a_1 state (and hence in the formation energy) reflects the correct physics, according to the relative amount of valence- versus conduction-band character in the state. It is noteworthy that these shifts could not simply be guessed based on the qualitative character of the defect state. Indeed, as noted above, the defect state for V_O is made up largely of Zn dangling bands, and hence one might guess that it would exhibit mainly conduction-band character and shift rigidly with the conduction band when the band gap is corrected by 0.7 eV (from E_g=0.8 eV in LDA to 1.5 eV in LDA$+U$). For a state occupied with two electrons, the for-

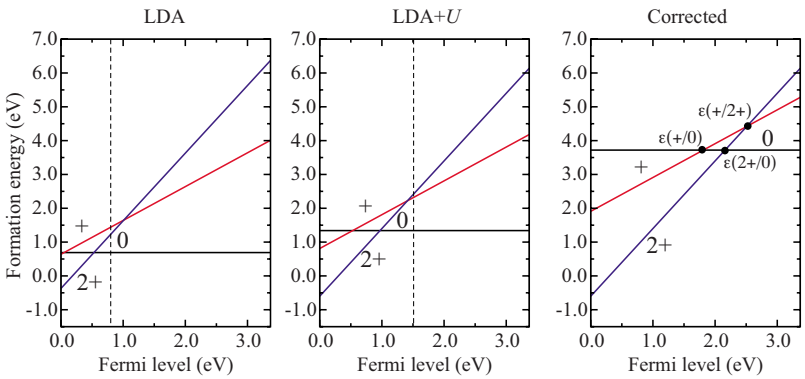

FIG. 1. (Color online) Calculated formation energies as a function of Fermi level for the oxygen vacancy in ZnO: (a) LDA results, (b) LDA $+U$ results, and (c) extrapolated formation energies. The zero of Fermi level corresponds to the valence-band maximum, and the dashed lines in (a) and (b) indicate the band gap in the respective calculations.

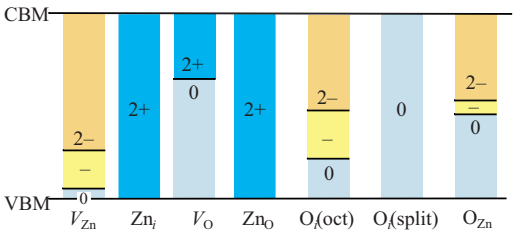

FIG. 2. (Color online) Thermodynamic transition levels for defects in ZnO. These values were corrected based on the LDA and LDA+U calculations according to the procedure described in the text.

mation energy should therefore increase by 1.4 eV. Instead, Fig. 1 shows that the correction is less than half of that. The reason is twofold: first, the LDA+U shifts not only the conduction band but also the valence band; second, evidently, the V_O defect state is by no means purely conduction-band-like in character.

We note in Fig. 1 that the transition levels are shifted upward in the band gap going from LDA to LDA+U. Since the difference between LDA and LDA+U reflects the correct physics of valence-band versus conduction-band character of the Kohn-Sham states, this justifies the use of Eq. (6) to correct defect transitions levels $\varepsilon(q/q')$ (which are referenced with respect to the VBM). The procedure above differs from the assumption made by Lany and Zunger in Refs. 62 and 44 that the transition levels of the oxygen vacancy in ZnO are well described by the LDA+U and do not shift at all with the conduction band. In their paper, the conduction band is rigidly shifted to agree with the experimental band gap, while the transition level is kept fixed at the LDA+U values. We feel that Lany and Zunger's assumption is unjustified and inconsistent with their own results. The assumption that the transition levels associated with the oxygen vacancy do not shift when the conduction band is corrected is equivalent to saying that the a_1 state has purely valence-band character. If this were true, then LDA and LDA+U would give the same values for the transition levels, which is clearly not the case. By increasing the band gap by going from LDA to LDA+U, the transition levels $\varepsilon(2+/+)$, $\varepsilon(+/0)$, and $\varepsilon(2+/0)$ all shift upward in the band gap, in accordance with the relative conduction versus valence-band character of the a_1 state.

Figure 2 and Table II show the position of the corrected transition levels $\varepsilon(q/q')$ in the band gap for all native point defects in ZnO. The justification provided above applies to transition levels and therefore also to *relative* formation energies. The correction of the *absolute* formation energies requires further considerations regarding the electronic structure. First, we take note of the fact that the formation energies of charged defects depend on the position of the VBM, as evident from Eq. (3). We already emphasized that the LDA+U approach shifts the position of the VBM as well as the CBM, and this effect will be felt in the formation energies as well. LDA+U of course does not fully correct the band-gap error; indeed, the usual errors intrinsic to the DFT-LDA are still present, and the LDA+U band gap (1.51 eV) is still significantly smaller than the experimental gap (3.43 eV). To extrapolate to the experimental gap, we therefore need to make an assumption about the position of

TABLE II. Calculated transition levels $\varepsilon(q/q')$ (in eV) for native point defects in ZnO. LDA, LDA+U, and corrected values according to Eq. (6) are listed. The differences $\Delta\varepsilon=\varepsilon(q/q')^{\mathrm{LDA}+U}-\varepsilon(q/q')^{\mathrm{LDA}}$ (in eV) are also listed.

Defect	q/q'	$\varepsilon(q/q')^{\mathrm{LDA}}$	$\varepsilon(q/q')^{\mathrm{LDA}+U}$	$\Delta\varepsilon$	$\varepsilon(q/q')$
V_O	2+/+	1.01	1.41	0.40	2.51
	+/0	0.05	0.53	0.47	1.82
V_{Zn}	0/−	0.08	0.11	0.03	0.18
	−/2−	0.29	0.45	0.16	0.87
Zn_i	2+/+	1.41	2.01	0.60	3.65
	+/0	1.44	2.06	0.62	3.75
O_i(oct)	0/−	0.27	0.39	0.12	0.72
	−/2−	0.86	1.06	0.20	1.59
O_i(split)	2+/+	−0.12	−0.09	0.03	−0.01
	+/0	−0.07	−0.05	0.02	0.00
Zn_O	4+/3+	0.33	0.75	0.42	1.88
	3+/2+	−0.25	0.09	0.34	1.01
	2+/+	1.59	2.21	0.62	3.91
	+/0	1.62	2.25	0.63	3.97
O_{Zn}	0/−	0.60	0.85	0.25	1.52
	−/2−	0.55	0.88	0.32	1.77

the valence band. Here, we assume that this position is given by the value in the LDA+U approach and does not undergo any further shifts when the band gap is corrected; i.e., we assume that all of the remaining correction occurs in the conduction band. This approximation is justified by studies in which LDA band structures were compared with quasiparticle calculations based on the GW approximation.[63]

Additional corrections to the formation energy occur, of course, due to the correction in the position of the defect states. In case the defect state is not occupied, this shift will not affect the formation energy. Therefore, for charge states in which the defect (Kohn-Sham) states are empty, the formation energy of the defect should be well approximated by its LDA+U value. In the case of the oxygen vacancy, this is true for the 2+ charge state, for which the a_1 defect state is empty of electrons. The corrected formation energy for V_O^{2+}, corresponding to the experimental band gap, is therefore equal to its LDA+U value (see Fig. 1). Once we have an absolute value for the 2+ charge state, we can then obtain extrapolated values for the formation energies of the other charge states (+ and 0) as well, taking advantage of the fact that our extrapolation scheme for transition levels [Eq. (6)] also provides corrected values for the *differences* between formation energies of different charge states.

Some point defects are not stable in any charge state for which all associated defect states are unoccupied, complicating the question of how to obtain an absolute formation energy. Indeed, for a charge state in which the defect state is partially occupied, the shift in formation energy from LDA to LDA+U contains contributions from both the shift in the VBM *and* the shift in the defect state. Since further correction to the experimental gap should leave the VBM intact, we need to somehow disentangle these two contributions. Fortunately, this is also feasible by closer inspection of the LDA and LDA+U results. Indeed, the part of the formation-energy correction due to the shift in the Kohn-Sham state is proportional to the number of electrons that are occupying the defect state(s). We can calculate this shift by looking at the energy difference between two charge states q and q' and analyzing how this energy difference changes in going from LDA to LDA+U. The energy difference between these charge states is contained in the transition level $\varepsilon(q/q')$ defined in Eq. (5). Therefore, we can identify the energy shift in the Kohn-Sham state(s) as the shift of the calculated transitions levels between the LDA and LDA+U calculations in Eq. (7) [$\Delta\varepsilon=\varepsilon(q/q')^{\mathrm{LDA}+U}-\varepsilon(q/q')^{\mathrm{LDA}}$]. Once we know the shift in the defect state(s), we can extrapolate the absolute formation energies according to

$$E^f = E^{f,\mathrm{LDA}+U} + \frac{(E_g^{expt} - E_g^{\mathrm{LDA}+U})}{(E_g^{\mathrm{LDA}+U} - E_g^{\mathrm{LDA}})} n\Delta\varepsilon, \tag{8}$$

where n is the occupancy of the defect states in the band gap for a given defect in charge state q. For example, in the case of the oxygen vacancy, $n=0$ for $q=2+$, $n=1$ for $q=1+$, and $n=2$ for $q=0$; for the zinc vacancy, $n=4$ for $q=0$, $n=5$ for $q=1-$, and $n=6$ for $q=2-$. Note that Eq. (8) correctly incorporates our notion about the VBM position being fixed by the LDA+U calculation. As a check, our previous discussion of LDA+U giving the correct absolute formation energy for V_O^{2+} is also correctly reflected by Eq. (8) (since $n=0$). For some defects, states with very different character or symmetry can simultaneously be occupied in certain charge states. In the case of the zinc antisite, the different $\Delta\varepsilon$ values can be explicitly evaluated (see Table II) and taken into account in the corrections for formation energies (Table III). In other cases, one of the states is actually a deep resonance in the valence band (e.g., the a_1 state of the zinc vacancy), and hence it is not possible to obtain explicit information about the $\Delta\varepsilon$ value for this state. In such cases, we have assumed that the state has mostly valence-band character, and therefore it follows the valence band and does not contribute to the correction in the formation energy as the band gap is corrected. In the case of the zinc vacancy, we have checked the effect of this assumption on the calculated formation energies by also extrapolating formation energies based on including the electrons in the a_1 state in the electron count n; this corresponds to assuming that the a_1 state shifts up as much as the t_2 states, which would provide an upper limit to the expected shift; even then, the results differed by less than 0.5 eV, assuring us that our results are not very sensitive to this approximation.

In Table III, we list the calculated LDA, LDA+U, and corrected formation energies and the occupancy of the defect states in the gap for each charge state. Our method for extracting $\Delta\varepsilon$ from the difference $\varepsilon(q/q')^{\mathrm{LDA}+U}-\varepsilon(q/q')^{\mathrm{LDA}}$ is not unique, since most defects can be stable in more than two different charge states. Each combination of charge states q and q' therefore, in principle, leads to a different value for $\Delta\varepsilon$. Fortunately, we find that all of these different evaluations provide very similar results, as illustrated by the values in Table III for cases where q and q' differ by 1. The values of $\Delta\varepsilon$ extracted for different combinations of charge states agree with each other to better than 0.1 eV (the case of the Zn antisite reflects different physics and will be discussed below). This provides a confirmation of the validity of our analysis of the various contributions to the formation energy and how they shift with the band-gap correction. For purposes of the extrapolation in Eq. (8) we use an average of the various $\Delta\varepsilon$ values for a given defect state, which we call $\Delta\bar{\varepsilon}$.

From the values of $\Delta\bar{\varepsilon}$, we can also extract a quantitative measure of the relative amount of valence-band versus conduction-band character. Defect states that have exclusively conduction-band character should shift rigidly with the conduction band, i.e., $\Delta\bar{\varepsilon}=\Delta E_g$. Conversely, defect states with exclusively valence-band character should not shift at all when the band gap is corrected, i.e., $\Delta\bar{\varepsilon}=0$. This implies that the parameter $\Delta\bar{\varepsilon}/\Delta E_g$ can be used as a measure of the amount of conduction-band character in the state (and $1-\Delta\bar{\varepsilon}/\Delta E_g$ as a measure of the amount of valence-band character). Inspection of the results in Table III indicates that these measures indeed agree with our intuitive expectations about the nature of the defect states. For instance, for zinc interstitials, the shift $\Delta\bar{\varepsilon}$ is almost equal to the band-gap correction given by LDA+U: $\Delta\bar{\varepsilon}=0.61$ eV compared to $\Delta E_g=E_g^{\mathrm{LDA}+U}-E_g^{\mathrm{LDA}}=0.71$ eV. The resulting degree of conduction- versus valence-band character CB/VB, expressed as percentages, is 87/13. Similar results apply to the high-lying transition states of another donor, the zinc antisite.

TABLE III. Calculated formation energies at $E_F=0$ for native point defects in ZnO under zinc-rich conditions. LDA, LDA+U, and corrected values are listed. The occupancy n of the defect states in the gap, defined in the text, is also given. In the last column, we list the relative degree of conduction-band versus valence-band character (CB/VB, expressed as percentages) for each defect state. The conduction-band character of the defect state is defined as CB=$\Delta\bar{\varepsilon}/\Delta E_g$, where $\Delta E_g=E_g^{\mathrm{LDA}+U}-E_g^{\mathrm{LDA}}=0.71$ eV (see text). All energies are given in eV.

Defect	q	n	$E^{f,\mathrm{LDA}}$	$E^{f,\mathrm{LDA}+U}$	E^f	(CB/VB) (%)
V_{O}	2+	0	−0.37	−0.60	−0.60	
	+	1	0.64	0.81	1.91	62/38
	0	2	0.69	1.34	3.72	
V_{Zn}	0	4	5.94	6.39	7.38	
	−	5	6.02	6.49	7.55	13/87
	2−	6	6.31	6.94	8.43	
Zn_i	2+	0	−0.10	−0.45	−0.45	
	+	1	1.32	1.56	3.20	87/13
	0	2	2.76	3.62	6.95	
O_i(oct)	0	4	6.36	6.83	8.54	
	−	5	6.63	7.22	9.26	23/77
	2−	6	7.49	8.28	10.86	
O_i(split)	2+	2	5.13	5.12	5.25	
	+	3	5.01	5.02	5.24	03/97
	0	4	4.93	4.97	5.24	
Zn_{O}	4+	0	0.14	−0.31	−0.31	
	3+	1	0.48	0.44	1.57	54/46
	2+	2	0.22	0.53	2.59	
	+	3	1.81	2.74	6.49	89/11
	0	4	3.43	4.98	10.47	
O_{Zn}	0	4	9.94	10.04	13.15	
	−	5	10.53	10.88	14.68	41/59
	2−	6	11.08	11.76	16.45	

The oxygen vacancy, however, clearly exhibits a more complex behavior, with only 62% of conduction-band character in its defect state. Defect states induced by acceptors, on the other hand, are expected to exhibit mainly valence-band character. For instance, for zinc vacancies $\Delta\bar{\varepsilon}=0.09$ eV, resulting in 87% valence-band character. This is consistent with the fact that the defect states of the zinc vacancy are composed of oxygen dangling bonds and are expected to follow the valence band.

Figure 3 shows the *corrected* (i.e., extrapolated to the experimental gap) formation energies for the relevant native point defects in ZnO as a function of Fermi-level position. The kinks in the curves for a given defect indicate transitions between different charge states; the corresponding (corrected) transition levels $\varepsilon(q/q')$ were shown in Fig. 2. Note that the results presented in this work differ slightly from some values that were published in a preliminary account of our work.[64] The differences are related to the correction of absolute formation energies, where we now take into account the occupancy of the defect-induced states in the case of defects with partially occupied states in the band gap.

We now discuss the electronic structure and local atomic geometry of each point defect in more detail. For migration barriers, we found that the values calculated with LDA and LDA+U differed by less than 0.1 eV; the results quoted below are all obtained within LDA.

C. Oxygen vacancies

1. Formation energy and transition levels

Among the defects that behave as donors, oxygen vacancies have the lowest formation energy (see Fig. 3). Oxygen vacancies have frequently been invoked as the source of unintentional n-type conductivity in ZnO. Our first-principles results indicate that this assignment cannot be correct. First, we note from Fig. 3 that the formation energy of V_{O} is quite high in n-type material, even under extreme Zn-rich conditions (3.72 eV). This indicates that the V_{O} concentration will always be low under equilibrium conditions. Even more important, however, is that the oxygen vacancy is a *deep* rather

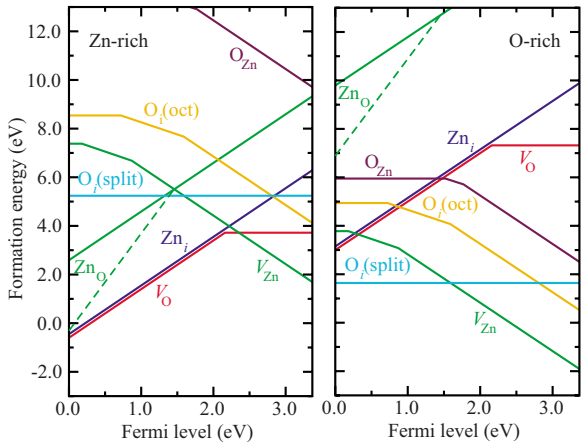

FIG. 3. (Color online) Formation energies as a function of Fermi-level position for native point defects in ZnO. Results for Zn-rich and O-rich conditions are shown. The zero of Fermi level corresponds to the valence-band maximum. Only segments corresponding to the lowest-energy charge states are shown. The slope of these segments indicates the charge state. Kinks in the curves indicate transitions between different charge states.

than a *shallow* donor. Figure 1(c) shows that the $\varepsilon(2+/0)$ transition level occurs at ~1 eV below the conduction band. Therefore, it is clear that the oxygen vacancy cannot provide electrons to the conduction band by thermal excitation in steady state, and therefore it cannot be a source of the often-observed unintentional *n*-type conductivity. It should be noted that, whereas the formation energy of V_O is very high in *n*-type ZnO, it is much lower in *p*-type ZnO, where V_O assumes the 2+ charge state. Thus, oxygen vacancies are a potential source of compensation in *p*-type ZnO.

2. Atomic geometry and negative-U character

We find that the oxygen vacancy is a "negative-*U*" center, implying that $\varepsilon(2+/+)$ lies above $\varepsilon(+/0)$, with $U=\varepsilon(+/0)-\varepsilon(2+/+)=-0.7$ eV, as seen in Fig. 1(c). As the Fermi level moves upward, the thermodynamic charge-state transition is thus directly from the 2+ to the 0 charge state; the + charge state is unstable for any position of the Fermi level.

Negative-*U* behavior is typically related to unusually large local lattice relaxations that stabilize particular charge states. For the neutral charge state, the four Zn nearest neighbors are displaced *inward* by 12% of the equilibrium Zn–O bond length, whereas for the + and 2+ charge states, the displacements are *outward* by 2% and 23%, as shown in Fig. 4. The origin of these large lattice relaxations lies in the electronic structure of V_O, which was discussed in Sec. III B. In the neutral charge state, the a_1 state is occupied by two electrons, and its energy is lowered as the four Zn atoms approach each other. In this case, the gain in electronic energy exceeds the cost to stretch the Zn–O bonds surrounding the vacancy and the resulting a_1 state lies near the top of the valence band. In the V_O^+ configuration, the a_1 state is occupied by one electron, and the electronic energy gain is too small to overcome the strain energy; the four Zn atoms are displaced slightly outward, moving the a_1 state near the middle of the band gap. In the V_O^{2+} configuration, the a_1 state is empty and the four Zn atoms strongly relax outward, strengthening the Zn–O bonds; the empty a_1 state lies near the conduction band. These large relaxations significantly reduce the formation energies of V_O^{2+} and V_O^0 relative to V_O^+, making the oxygen vacancy a negative-*U* center.

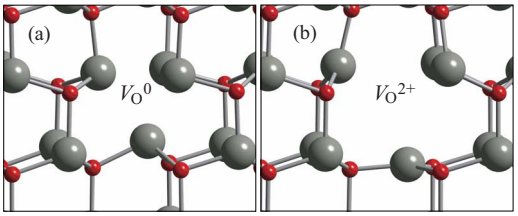

FIG. 4. (Color online) Local atomic relaxations around the oxygen vacancy in the (a) neutral and (b) 2+ charge states. In the neutral charge state, the four Zn nearest neighbors are displaced *inward* by 12% of the equilibrium Zn–O bond length, whereas for the 2+ charge states, the displacements are *outward* by 23%.

3. Comparison with experiment

Most of the experimental investigations of oxygen vacancies to date have relied on electron paramagnetic resonance (EPR) studies, which will be discussed in the next section. One group has studied oxygen vacancies with positron annihilation spectroscopy.[35,65] The samples were electron irradiated and had a Fermi level 0.2 eV below the CBM after irradiation. The dominant compensating defect was found to be the zinc vacancy, consistent with the results presented in Sec. III D; however, the measurements also produced evidence for the presence of a neutral defect, which was proposed to be the neutral oxygen vacancy. These observations are fully consistent with our computational results, both regarding the absence (below the detection limit) of oxygen vacancies in as-grown material and V_O being present in the neutral charge state when $E_F=E_c-0.2$ eV.

4. Electron paramagnetic resonance studies of oxygen vacancies

Because of the negative-*U* character, the positive charge state of the oxygen vacancy, V_O^+, is thermodynamically unstable, i.e., it is always higher in energy than either V_O^{2+} or V_O^0 for any position of the Fermi level in the band gap. This has important implications for the characterization of oxygen vacancies in ZnO. Only the positive charge state, with its unpaired electron, is detectable by magnetic resonance techniques. An EPR signal associated with V_O should thus not be observed under thermodynamically stable conditions. It is, of course, possible to create oxygen vacancies in the + charge state in a metastable manner, for instance, by excitation with light. Once generated, V_O^+ does not immediately decay into the 2+ or 0 charge state due to the existence of energy barriers associated with the large lattice relaxations that occur around the oxygen vacancy. We thus expect that at low enough temperatures, EPR signals due to V_O^+ may be observed upon excitation. When the excitation is removed and the temperature is raised, these signals will decay. In Ref. 66, we reported that the thermal barrier to escape out of the 1+

TABLE IV. Overview of electron paramagnetic resonance observations of donors in ZnO.

g value	Sample	Treatment	Reported assignment	Reference
$g_{\parallel}=1.956$, $g_{\perp}=1.955$	Single crystal	1150 °C, 7 h	Electrons bound at donors	74
$g_{\parallel}=1.957$, $g_{\perp}=1.956$	Powder	900 °C, 2 h	V_O	75
$g_{\parallel}=1.956$, $g_{\perp}=1.955$	Single crystal		Electrons in CB or donor band	71
$g\sim 1.96$	Powder	975 °C, 1–20 h		76
$g=1.9539$	Powder		V_O	77
$g=1.9564, 1.9600$	Powder	575 K, vac+O_2		72
$g_{\parallel}=1.9576$, $g_{\perp}=1.9557$	Single crystal		V_O	8
$g=1.9557$	Powder		V_O	78
$g=1.956$	Powder		V_O	79
$g=1.955$	Powder	930 °C, H_2 then O_2	V_O	10
$g_{\parallel}=1.9573$, $g_{\perp}=1.9557$	Powder		V_O	80 and 81
$g=1.9564$, $g=1.9596$	Powder	700–900 °C, N_2, H_2	V_O	96
$g_{\parallel}=1.9570$, $g_{\perp}=1.9551$	Single crystal		Delocalized electrons	21
$g_{\parallel}=1.957$, $g_{\perp}=1.956$	Single crystal		Shallow donors	22
$g_{\parallel}=1.9948$, $g_{\perp}=1.9963$	Single crystal	Irradiation, illumination	V_O	69
$g_{\parallel}=1.9948$, $g_{\perp}=1.9961$	Single crystal	Irradiation	V_O	86
$g_{\parallel}=1.9945$	Single crystal	Irradiation, illumination	V_O	68
$g_{\parallel}=1.9945$	Single crystal	Irradiation, illumination	V_O	70
$g_{\parallel}=1.9948$, $g_{\perp}=1.9963$	Single crystal	Irradiation, illumination	V_O	109 and 110
$g_{\parallel}=1.9951$, $g_{\perp}=1.9956$	Single crystal	Irradiation, illumination	V_O-Li_{Zn}	109

charge state is 0.3 eV, sufficient to maintain an observable concentration of V_O^+ during excitation and cause persistent photoconductivity at low temperature, but clearly too low to allow for persistent photoconductivity at room temperature. We therefore disagree with Lany and Zunger's[44] proposal that persistent conductivity related to oxygen vacancies is responsible for the unintentional n-type conductivity observed in many ZnO samples.

In previous work,[66] we mapped out a complete configuration-coordinate (CC) diagram that allowed a direct comparison with experimental results and provided a detailed interpretation of recent optically detected EPR experiments.[24] The key quantity for this interpretation is the position of the $\varepsilon(+/0)$ transition level in the band gap. Our results yield $\varepsilon(+/0)=1.82$ eV, in contrast to the value of 1.2 eV obtained by Zhang *et al.*[38] based on empirical corrections, the value of ~ 1 eV obtained by Lany and Zunger[44,62] using LDA+U, or the results of Kohan *et al.*[37] that find the $\varepsilon(+/0)$ in the valence band. From our calculated CC diagram in Ref. 66, the peak for the optical transition $V_O^0 \rightarrow V_O^+$ is 2.0 eV. This value is in good agreement with the experimental value of 2.3 eV.[67,68]

Experimental reports of EPR measurements relating to oxygen vacancies in ZnO are summarized in Table IV. They fall into two categories, depending on the value of the g factor. One set of reports associates oxygen vacancies with a g value of ~ 1.96, the other with $g\sim 1.99$ (see Table IV). We feel that there is overwhelming evidence that it is actually the $g\sim 1.99$ line that can be consistently assigned to oxygen vacancies. This signal has only been observed after irradiation of the samples, clearly indicating it is related to a point defect (and also consistent with our calculated result that V_O has a high formation energy and is thus unlikely to occur in as-grown n-type material). Also, it has been found that illumination is necessary to observe the center,[68–70] consistent with our results that excitation is required in order to generate the paramagnetic 1+ charge state. Further evidence for the identification of the $g\sim 1.99$ line with the oxygen vacancy came from the observations of hyperfine interactions with the ^{67}Zn neighbors of the vacancy.[68,69]

On the other hand, no hyperfine interactions were observed for the $g\sim 1.96$ line. It is likely that the $g\sim 1.96$ signal is associated with electrons in the conduction band or in a donor band, as originally proposed by Müller and Schneider[71] and most recently confirmed by Garces *et al.*[22] The historical tendency for authors to assign the $g\sim 1.96$ line to V_O was probably largely based on the prevailing hypothesis that oxygen vacancies were the donors responsible for the unintentional n-type conductivity. In a collection of experimental results up to 1970, Sancier[72] also favored assigning the $g\sim 1.96$ line to electrons in the conduction band. In a critical review of results up to 1981, Neumann[73] observed that doping with Al, Ga, or In increases the intensity of the $g\sim 1.96$ signal. This behavior is consistent with the $g\sim 1.96$ signal being due to delocalized electrons, but would be hard to reconcile with oxygen vacancies as the source. We

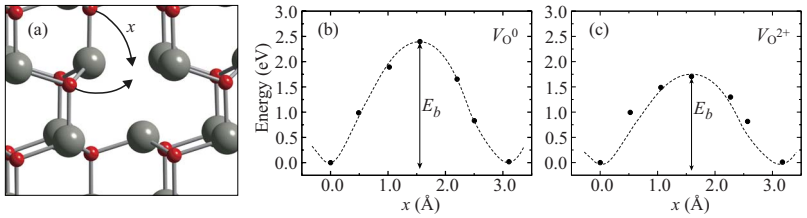

FIG. 5. (Color online) (a) Migration path for oxygen vacancies. Paths where an oxygen atom from an adjacent basal plane or from the same basal plane moves into the vacancy are shown. Both migration paths give quantitatively the same result. (b) Calculated energy along the migration path for the neutral charge state of V_O. (c) Calculated energy along the migration path for the 2+ charge state of V_O.

note that the $g \sim 1.96$ line has also been reported to be photosensitive; in particular, the signal is enhanced after UV illumination.[71,74,76–79,81] This observation is of course consistent with the $g \sim 1.96$ line corresponding to electrons in conduction-band states, since UV light can promote electrons into these states.

Leiter *et al.*[82,83] have performed photoluminescence and optically detected magnetic resonance experiments on a broad emission band around 2.45 in as-grown (i.e., unirradiated) single-crystal ZnO. They detect intense resonances which they attribute to a spin-triplet system ($S=1$) with $g_{\parallel}=1.984$ and $g_{\perp}=2.025$. Based on analogies with anion vacancies in other oxides, they attribute this signal to oxygen vacancies. This assignment clearly disagrees with the experiments cited above, which relate oxygen vacancies with g values of 1.99. In addition, the signal observed by Leiter *et al.* is present already in as-grown material, unlike the observations of the $g \sim 1.99$ signal, which all required irradiation. Moreover, the samples in Refs. 82 and 83 are n type with residual donor concentration of $\sim 10^{17}$ cm^{-3}. Under these conditions, the formation energy of oxygen vacancies is at least 3.7 eV, and therefore, the concentration of oxygen vacancies is expected to be too low to be detectable. In Sec. III D 4, we will suggest that the triplet $S=1$ EPR signal observed by Leiter *et al.* may be related to spin-dependent recombination involving an $S=1/2$ V_{Zn}^- defect exchange coupled to a $S=1/2$ effective-mass donor.

5. Migration

Migration of oxygen vacancies occurs when a nearest-neighbor oxygen atom in the oxygen lattice jumps into the vacant site, leaving a vacancy behind. We obtained the migration barrier of the oxygen vacancy by calculating the total energy at various intermediate configurations when moving an oxygen atom from its nominal lattice site adjacent to the vacancy along a path toward the vacancy, as schematically shown in Fig. 5(a). The coordinate along the path is the distance between the intermediate position of the jumping atom and its original lattice site position. The oxygen vacancy has 12 next-nearest-neighbor oxygen atoms: six are located in the same basal plane as the vacancy and account for vacancy migration perpendicular to the c axis; the other six neighbors are located in basal planes above and below the basal plane of the oxygen vacancy and can account for vacancy migration both parallel and perpendicular to the c axis. We find no anisotropy in the calculated migration barrier for oxygen vacancies. That is, migration barriers involving oxygen atoms from the basal plane of the vacancy and from planes above or below the basal plane of the vacancy have the same value.

However, the migration barrier does depend on the charge state of the oxygen vacancy. Here, we focus on the stable charge states, namely, neutral and 2+ (V_O^0 and V_O^{2+}, respectively). The former is relevant for migration of the oxygen vacancy in n-type samples and the latter for vacancy migration in semi-insulating (E_F below 2.2 eV) or p-type samples. We find migration barriers of 2.4 eV for V_O^0 and 1.7 eV for V_O^{2+}, as shown in Figs. 5(b) and 5(c). A possible qualitative explanation for this difference is that the saddle-point configuration may be regarded as a complex formed by an oxygen interstitial adjacent to two oxygen vacancies. For Fermi energies where V_O^0 is stable, this complex is $2V_O^0+O_i^0$, and for Fermi energies where V_O^{2+} is stable, this complex is $2V_O^{2+}+O_i^{2-}$. Because the distances between the vacancies and the interstitial are smaller than the distance between the two vacancies, the complex $2V_O^{2+}+O_i^{2-}$ can exhibit an attractive Coulomb energy which is absent in $2V_O^0+O_i^0$. Thus, the Coulomb interaction between the point defects that constitute the saddle-point configuration lowers the migration barrier for V_O^{2+} compared to V_O^0. We note that a similar difference between migration barriers for different charge states of the anion vacancy was found in the case of GaN.[29]

Erhart and Albe[43] have also calculated migration barriers for the oxygen vacancy in ZnO. Contrary to our results, they find differences as large as 0.7 eV between migration barriers involving oxygen atoms from the basal plane of the vacancy and from planes above or below the basal plane of the vacancy. They have reported energy barriers of 1.09 and 1.49 eV for migration of V_O^{2+} through out-of-plane and in-plane paths, respectively, and energy barriers of 2.55 and 1.87 eV for migration of V_O^0 through out-of-plane and in-plane paths, respectively. Such large anisotropies in the migration barriers are quite unexpected since the local geometry around the oxygen vacancy has almost tetrahedral symmetry. We suspect that the small supercells (32 atoms) used in Ref. 43 are responsible for the discrepancies. In fact, calculated migration barriers for the nitrogen vacancy in GaN using 32- and 96-atom supercells differ by as much as 0.6 eV due to the large relaxations around the nitrogen vacancy that are not properly described in the 32-atom supercell.[29] More details about the effects of supercell size can be found in Ref. 29.

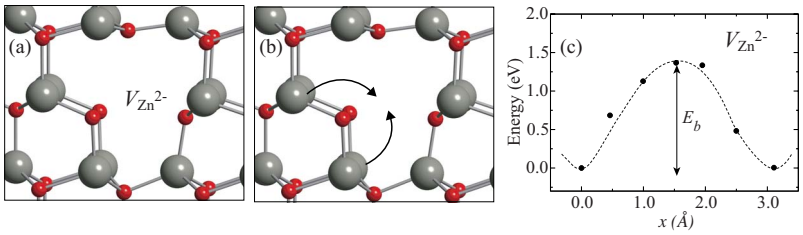

FIG. 6. (Color online) (a) Local atomic geometry of the zinc vacancy in the 2− charge state (V_{Zn}^{2-}). (b) Migration paths of V_{Zn}^{2-}: a zinc atom from an adjacent basal plane or from the same basal plane moves into the vacancy. (c) Calculated migration energy barrier for the path where a zinc atom from an adjacent basal plane moves to the vacant site. The two migration paths give energy barriers that differ by less than 0.1 eV.

D. Zinc vacancies

1. Formation energy and transition levels

Zinc vacancies in ZnO introduce partially occupied states in the band gap. These states are derived from the broken bonds of the four oxygen nearest neighbors and lie close to the VBM. They are partially filled and can accept additional electrons, thus causing V_{Zn} to act as an acceptor. Since the formation energy of acceptor-type defects decreases with increasing Fermi level, zinc vacancies can more easily form in *n*-type samples. They are also more favorable in oxygen-rich conditions. Figure 3 shows that zinc vacancies have exceedingly high formation energies in *p*-type ZnO, and therefore should exist only in very low concentrations. In *n*-type ZnO, on the other hand, zinc vacancies have the lowest formation energy among the native point defects. This energy is low enough for V_{Zn}^{2-} to occur in modest concentrations in *n*-type ZnO, acting as a compensating center. Positron annihilation experiments[35,84] have indeed identified the presence of zinc vacancies in *n*-type ZnO.

We find that zinc vacancies are deep acceptors with transition levels $\varepsilon(0/-)=0.18$ eV and $\varepsilon(-/2-)=0.87$ eV (see Fig. 2 and Table II). These levels are too deep for the zinc vacancy to act as a shallow acceptor. Combined with their very high formation energy in *p*-type materials, zinc vacancies are thus unlikely to play any role in *p*-type conductivity. Recently, it has been proposed that a complex formed by a column-V element (As or Sb) on a substitutional Zn site surrounded by two zinc vacancies could be responsible for *p*-type conductivity.[85] From our results, the formation energy of a single V_{Zn} is 3.7 eV under *p*-type conditions (E_F at the VBM), in the most favorable O-rich limit. The formation energy of an As_{Zn}-$2V_{Zn}$ complex would then amount to 2×3.7 eV plus the formation energy of substitutional As_{Zn} (or Sb_{Zn}) minus the binding energy of the complex. Assuming that the formation energy of As_{Zn} (or Sb_{Zn}) is about 1 eV (a very low value considering the chemical mismatch between Zn and As), a binding energy of more than 6 eV would be required to stabilize these complexes at equilibrium conditions. This value is too large to be attainable, and indeed much larger than the value calculated in Ref. 85. Therefore, we feel that these complexes cannot be responsible for *p*-type conductivity.

2. Atomic geometry

The oxygen atoms around the zinc vacancy exhibit large outward breathing relaxations of about 10% with respect to the equilibrium Zn-O bond length, as shown in Fig. 6. Similar relaxations are observed for the three possible charge-state configurations, V_{Zn}^{0}, V_{Zn}^{-}, and V_{Zn}^{2-}. This indicates that the overlap between the oxygen $2p$ orbitals surrounding the zinc vacancy is too small to result in significant chemical bonding between the oxygen atoms. Indeed, the calculated O-O distances are about 3.5 Å, much larger than the sum of the covalent radii, which is 2×0.73 Å$=1.46$ Å.

Our calculations, using either LDA or LDA+U, do not produce any low-symmetry Jahn-Teller distortions for V_{Zn} in the 1− or neutral charge states, even though such distortions are known to occur.[17,86,87] Jahn-Teller distortions around cation vacancies also occur in ZnSe,[88] and it is known that LDA calculations are unable to reproduce them.[89,90] In the case of ZnO, for a zinc vacancy in the 1− charge state, the hole is expected to be localized on one of the four nearest-neighbor oxygen atoms. The oxygen atom with the hole is expected to relax toward the vacancy, lowering the energy of the system. In the case of ZnSe, according to Vlasenko and Watkins,[17] the Jahn-Teller effect lowers the energy of V_{Zn}^{-} by 0.35 eV. Note that a lowering of the energy of the negative charge state would result in the $\varepsilon(-/2-)$ transition level shifting away from the VBM, i.e., the tendency is toward making the level deeper.

3. Electron paramagnetic resonance studies of zinc vacancies

Several EPR observations of zinc vacancies in ZnO have been reported. Taylor *et al.*[86] reported EPR signals with *g* factors in the range 2.0018–2.056 in irradiated single crystals. It was proposed that a subset of these lines would be due to Zn vacancies. Galland and Herve[87] observed lines with *g* factors between 2.0024 and 2.0165 in irradiated single crystals, also attributing them to Zn vacancies.

4. Green luminescence

ZnO often exhibits green luminescence, centered between 2.4 and 2.5 eV.[18,19,91,92] This green luminescence has been observed in samples prepared with a variety of growth techniques, and it is important to point out that there may not be a single source for this luminescence. For instance, Cu has been suggested as a potential cause.[93,94] Still, not all ZnO samples contain copper. Native defects have also been suggested as a potential source. Reynolds *et al.*[18,19] and Kohan *et al.*[37] have suggested that the Zn vacancy can give rise to green luminescence. Indeed, our calculated transition level

between the 1− and 2− charge states occurs at 0.9 eV above the VBM, and hence a transition between the conduction band (or a shallow donor) and the V_{Zn} acceptor level would give rise to luminescence around 2.5 eV, in good agreement with the observed transition energy. In addition to the agreement with the observed emission energy, the Zn vacancy is also a likely candidate because it is an acceptor-type defect: acceptor defects are more likely to occur in *n*-type material, and most ZnO materials to date have exhibited unintentional *n*-type conductivity. This proposed explanation for the green luminescence that is similar to the proposal that gallium vacancies are the source of the yellow luminescence in GaN.[95]

Other explanations have been proposed for the green luminescence. Several groups have suggested that oxygen vacancies are the source of green luminescence.[38,82,96–99] Vanheusden *et al.* reported a correlation between the intensity of the green luminescence and the concentration of oxygen vacancies.[96,99] However, their assessment of the presence of oxygen vacancies was based on the observation of a line with $g \sim 1.96$ in EPR measurements; as discussed in Sec. III C 4, this assignment is not correct, undermining the arguments made in Refs. 96 and 99. Our calculated configuration-coordinate diagrams for V_O also do not show any transitions consistent with green luminescence.[66] Leiter *et al.* associated the emission band around 2.45 with oxygen vacancies based on optically detected magnetic resonance experiments.[82,83] As discussed in Sec. III C 4, we consider it unlikely that the $S=1$ center that they observed represents the oxygen vacancy. Instead, we suggest that V_{Zn} may be the cause of the observed green luminescence in their experiments. Indeed, an $S=1$ EPR signal related to V_{Zn} has recently been proposed by Vlasenko and Watkins.[17] We suggest that the signal observed by Leiter *et al.* is associated with the spin-dependent process $V_{Zn}^{-}+EM^{0} \rightarrow V_{Zn}^{2-}+EM^{+}$, where the exchange interaction between the localized V_{Zn}^{-} center and the delocalized effective-mass (EM) donor is large enough to produce the combined $S=1$ spectrum.

A strong argument in favor of zinc vacancies being the source of green luminescence has been provided by the experiments of Sekiguchi *et al.*, who have reported strong passivation of the green luminescence by hydrogen plasma treatment.[100] This observation is consistent with the green luminescence being caused by zinc vacancies, which act as acceptors; these acceptors can be passivated by hydrogen, which acts as a donor.[101,102] In fact, the same passivation effect was observed by Lavrov *et al.*, who simultaneously observed an increase in vibrational modes associated with hydrogenated zinc vacancies.[102] We note that the passivation of the green luminescence by hydrogen is very plausible if the green luminescence is caused by zinc vacancies; hydrogen atoms passivate zinc vacancies by forming strong O-H chemical bonds.[102]

5. Migration

Migration of zinc vacancies occurs when a nearest-neighbor zinc atom moves into the vacant site leaving a vacancy behind, as schematically shown in Fig. 6(b). For the calculation of the migration energy barrier, we have focused on the most relevant 2− charge state; zinc vacancies in the neutral and 1− charge states have very high formation energies and are therefore very unlikely to occur under equilibrium conditions. The calculated migration energy barrier for the zinc vacancy in the 2− charge state is 1.4 eV, as shown in Fig. 6(c). We also find that migration paths involving zinc atoms in adjacent basal planes or in the same basal plane of the zinc vacancy give almost the same result, differing by less than 0.1 eV. We thus predict that the migration of zinc vacancies in ZnO is isotropic, similar to the case of oxygen vacancies discussed above.

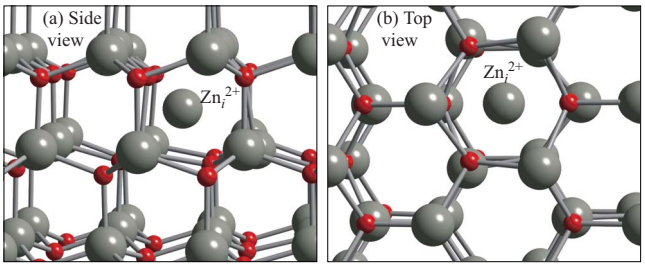

FIG. 7. (Color online) Local atomic geometry of the zinc interstitial in the 2+ charge state (Zn_i^{2+}) at the stable octahedral site. (a) Side view perpendicular to the *c* axis. (b) Top view parallel to the *c* axis (along the $[000\bar{1}]$ direction).

E. Zinc interstitials

1. Formation energy and atomic geometry

There are two distinct types of interstitial sites in the wurtzite structure: the tetrahedral site (tet) and the octahedral site (oct). The tetrahedral site has one zinc and one oxygen as nearest-neighbor atoms, at a distance of about $\sim 0.833 d_0$, where d_0 is the Zn-O bond length along the *c* axis. Thus, a zinc atom placed at this site will suffer severe geometrical constraints. The octahedral site is in the interstitial channel along the *c* axis. It is equidistant from three zinc and three oxygen atoms by $1.07 d_0$. From the distances above, one can expect the Zn_i interstitial to prefer the octahedral site where the geometrical constraints are less severe. Indeed, we find that the octahedral site is the stable site for Zn_i. The zinc interstitial at the tetrahedral site is 0.9 eV higher in energy and unstable: it spontaneously relaxes to the octahedral site, and therefore may play a rule in Zn_i migration, as discussed below. Moreover, instead of occupying the ideal octahedral site, we observe a large displacement of the Zn_i along the *c* axis. The resulting relaxed geometry has an increased Zn_i-Zn distance of $1.22 d_0$ and a decreased Zn_i-O distance of $1.02 d_0$; it is depicted in Fig. 7. Similar displacements along the *c* axis were also observed for Ga_i in GaN.[29]

Figure 3 shows that under *n*-type conditions, i.e., for Fermi-level positions near the conduction band, Zn_i has a high formation energy. This is true even under extreme Zn-rich conditions, where the value is ~6 eV. Zinc interstitials are thus unlikely to be responsible for unintentional *n*-type conductivity, since they will be present in very low concentrations in *n*-type ZnO. On the other hand, the formation energy of Zn_i^{2+} decreases rapidly when the Fermi level de-

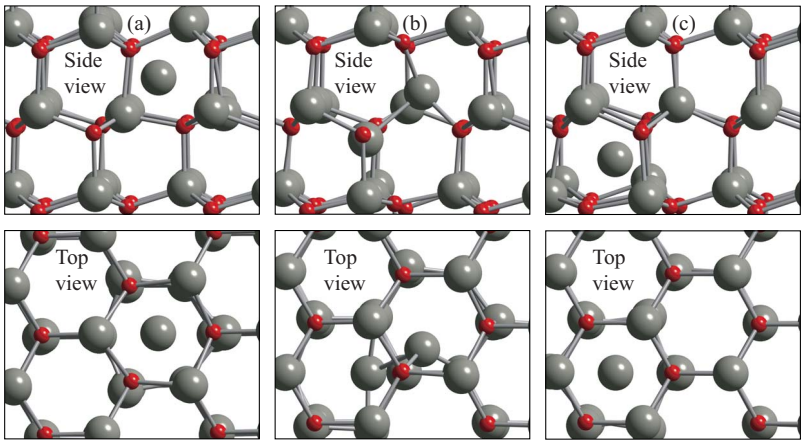

FIG. 8. (Color online) Migration path of Zn_i^{2+} in the kick-out mechanism, where a zinc interstitial exchanges places with a zinc atom on a substitutional lattice site. Upper panels show side views, whereas lower panels show top views. (a) Starting configuration, (b) saddle point, and (c) final configuration where the new zinc interstitial is in a basal plane adjacent to the basal plane of the original interstitial atom.

creases toward the VBM; zinc interstitials are thus a potential source of compensation in *p*-type ZnO.

2. Transition levels

The zinc interstitial introduces an a_1 state with two electrons above the CBM. These two electrons can be transferred to conduction-band states; indeed, we find that the zinc interstitial occurs exclusively in the 2+ charge state, with the $\varepsilon(2+/+)$ and $\varepsilon(+/0)$ levels above the conduction-band minimum, as listed in Table II. The zinc interstitial will always donate electrons to the conduction band, thus acting as a shallow donor. These electrons can, of course, be bound to the defect center in hydrogenic effective-mass states, so that effectively the observed defect levels would be the effective-mass levels below the CBM.

However, because of their high formation energy (see above), zinc interstitials will not be present in ZnO under equilibrium conditions. It has been suggested that zinc interstitials can be observed in *n*-type ZnO under nonequilibrium conditions. Thomas[6] reported the introduction of shallow donors when ZnO crystals were heated in Zn vapor followed by a rapid quench, and Hutson[5] observed the appearance of a shallow donor with ionization energy of 51 meV in Hall measurements. Look *et al.*[20] conducted high-energy electron irradiation experiments and identified a shallow donor with ionization energy of 30 meV. Based on the much higher production rate of this defect for the Zn (0001) face than for the O $(000\bar{1})$ face, the donor was suggested to be related to a Zn-sublattice defect, either the Zn interstitial itself or a Zn-interstitial-related complex. Because of the high formation energies of Zn_i and its high mobility even a low temperatures (as discussed in Sec. III E 3 below), we do not believe that isolated Zn_i were the observed species in the above experiments.

3. Migration

We have considered migration paths for Zn_i^{2+} parallel and perpendicular to the *c* axis. A "natural" migration path for the zinc interstitial would be along the hexagonal interstitial channel, parallel to the *c* axis [perpendicular to the plane of Fig. 7(b)]. Along this path, the interstitial has to pass through a plane containing three zinc atoms with relatively short Zn_i-Zn distances, leading to a barrier of 0.78 eV. At the saddle point, the Zn_i is in the same plane as the three Zn nearest neighbors, which are pushed apart, leading to a Zn_i-Zn distance of about $1.2d_0$ Å. This path can account for migration only along the *c* axis.

We have also considered a path where the interstitial Zn atom moves through the unstable tetrahedral site, resulting in diffusion perpendicular to the *c* axis. At the saddle point, the Zn_i atom is at the tetrahedral site, repelling the Zn nearest neighbor by increasing the Zn_i-Zn distance from $\sim 0.833d_0$ to $1.12d_0$. The migration barrier for this path is 0.90 eV, which is the energy difference between the octahedral and the tetrahedral configuration.

In addition, we have considered a kick-out (or interstitialcy) mechanism, as shown in Fig. 8. In this mechanism, the interstitial Zn atom moves in the direction of a substitutional Zn and then replaces it. The former substitutional zinc atom moves to an adjacent octahedral interstitial site, which can be either in the same basal plane or in a basal plane adjacent to that of the original interstitial (Fig. 8). At the saddle point, the substitutional site is shared by two Zn_i atoms in a rather symmetric configuration as shown in Fig. 8(b). Surprisingly, we find that the kick-out mechanism gives the lowest migration barrier: 0.57 eV (see Fig. 9). The difference in the energy barriers for migration paths parallel and perpendicular to the *c* axis is smaller than 0.05 eV. Therefore, the migration of the zinc interstitial is also predicted to be isotropic. The low migration barrier implies that zinc in-

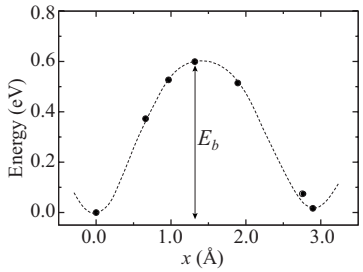

FIG. 9. Calculated migration energy barrier for Zn_i^{2+} in the kick-out mechanism illustrated in Fig. 8. The coordinate *x* is the sum of the distances between the intermediate positions of the two Zn atoms and their respective initial positions.

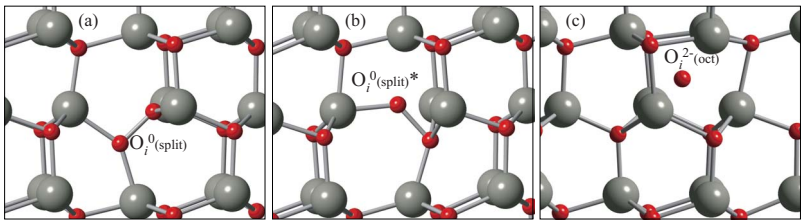

FIG. 10. (Color online) Local atomic geometry of electrically inactive oxygen split interstitial (a) in the most stable configuration $[O_i^0(\text{split})]$ and (b) in a metastable configuration $[O_i^0(\text{split})^*]$. (c) Local atomic geometry of electrically active oxygen interstitial at the octahedral site $[O_i^{2-}(\text{oct})]$.

terstitials are mobile below room temperature.

The calculated energy barrier of 0.57 eV is in very good agreement with experimental measurements by Thomas,[6] who reported a migration barrier of 0.55 eV for zinc interstitials based on experiments involving heating crystals in Zn vapor followed by rapid quenching. Despite this low migration barrier, the activation energy for Zn self-diffusion mediated by Zn interstitials is still quite high in *n*-type samples, since we have to add the formation energy of Zn_i [see Eq. (1)]. The fact that Zn interstitials migrate by exchanging positions with Zn atoms at regular lattice sites should be taken into account when interpreting the migration of Zn interstitials in ZnO crystals using Zn isotopes.[103,104] As a Zn isotope is introduced into ZnO, it promptly exchanges positions with Zn atoms in the lattice site via the kick-out mechanism. Further diffusion of these isotopes then has to occur through a self-diffusion mechanism, which can be mediated either by zinc interstitials or by zinc vacancies; these mechanisms have much higher activation energies than the migration barrier of zinc interstitials by themselves.

It is worth noting that their low migration barrier implies that zinc interstitials are very unlikely to occur as isolated interstitials; they will have a high tendency to either diffuse out of the sample or to bind with other defects or impurities. Combined with the high formation energy of Zn_i in *n*-type material, this renders it unlikely that zinc interstitials contribute to unintentional *n*-type conductivity in ZnO. Even when zinc interstitials are formed under nonequilibrium conditions, such as irradiation, they will be unlikely to remain present as isolated interstitials; the fact that they are mobile at temperatures below room temperature implies that they will either diffuse out of the sample, recombine with Zn vacancies, or form complexes with other defects or impurities.

It is conceivable that the high diffusivity of Zn_i would cause it to find sites (defects or impurities) to which it is attracted and with which it can form a complex with a positive binding energy. If this binding energy is sufficiently high, and if the resulting complex still acts as a shallow donor, this may yet provide a mechanism for the zinc interstitial to play a role in the unintentional *n*-type conductivity of ZnO. However, we feel that under equilibrium conditions, this is quite unlikely. Consider, for example, the seemingly favorable scenario of Zn_i binding to an impurity that acts as a single acceptor. Since Zn_i is a double donor, the resulting complex is expected to still act as a single donor. However, the formation energy of the isolated Zn interstitial is so high that the binding energy of the complex would need to be unreasonably high in order for the complex to occur in observable concentrations. We can illustrate this with the example of a zinc interstitial bound to a nitrogen acceptor, which was proposed in Ref. 105. Under *n*-type and Zn-rich conditions, Look *et al.*[105] calculated a formation energy for the $(N_O\text{-}Zn_i)^+$ complex of 3.4 eV. This high value renders it very unlikely that this complex would form during growth of the material. Such complexes would have a higher chance of occurring if the interstitials were created under nonequilibrium conditions such as irradiation, as discussed in Ref. 105.

F. Oxygen interstitials

1. Formation energy, atomic geometry, and transition levels

Excess oxygen atoms in the ZnO lattice can be accommodated in the form of oxygen interstitials. The oxygen atoms can occupy the octahedral or tetrahedral interstitial sites or form split interstitials. We find that the oxygen interstitial at the tetrahedral site is unstable and spontaneously relaxes into a split-interstitial configuration in which it shares a lattice site with one of the nearest-neighbor substitutional oxygen atoms. In Fig. 10(a), we show the local lattice relaxation around the oxygen split interstitial, O_i(split). The calculated O-O bond length is 1.46 Å, suggesting the formation of an O-O chemical bond, and the Zn-O distances are about 3% smaller than the equilibrium Zn-O bond length in ZnO. Recently, Limpijumnong *et al.* proposed that other first-row elements can also form "diatomic molecules" in ZnO.[106] Our results agree with theirs, and we also find two almost degenerate and completely filled states in the band gap that resemble the antibonding $pp\pi^*$ state from a molecular orbital description of the isolated O_2 molecule. However, in the isolated molecule, the $pp\pi^*$ molecular orbital (MO) is occupied by two electrons with parallel spins, resulting in a triplet $S=1$ ground state. In the solid, the four zinc nearest neighbors provide two additional electrons, and the $pp\pi^*$-like MO is completely occupied. This explains the significantly longer O-O bond length in the split interstitial (1.46 Å) compared to that of the isolated O_2 molecule (1.22 Å). Our calculations show that the oxygen split interstitial is electrically inactive, i.e., it is neutral for any Fermi-level position, with the calculated donor transition levels $\varepsilon(+/0)$ and $\varepsilon(+/2+)$ occurring below the VBM.

We also find a metastable configuration for the oxygen split interstitial, $O_i^0(\text{split})^*$, with formation energy ~0.2 eV higher than the lowest-energy configuration O_i^0(split). Its local atomic geometry is shown in Fig. 10(b). The O-O bond length is 1.51 Å. The existence of these two almost degenerate configurations with very different O-Zn-O bond angles reinforces the picture of the oxygen split interstitial as an O_2 molecule embedded in the ZnO crystal.

Erhart and Albe[43] also reported on the O_i(split) and its metastable configuration $O_i^0(\text{split})^*$. They referred to them as "dumbbell configurations," and they found that these defects

are electrically active with acceptor transition levels close to the CBM. Our results do not support this assignment. We suggest that the acceptor charge states found by Erhart and Albe are a result of occupying extended bulk states near the CBM, and not defect-induced states. This explains why Erhart and Albe find acceptor transition levels slightly above the CBM. According to our results, these two configurations are both dimers as previously reported in Ref. 106. The interaction between Zn-O_i is much less strong than that between O-O, resulting in similar formation energies despite the very different O-Zn-O bond angles in these two competing configurations.

Oxygen interstitials can also exist as electrically active interstitials occupying the octahedral site, O_i(oct), as shown in Fig. 10(c). Interstitial oxygen at the octahedral site introduces states in the lower part of the band gap that can accept two electrons. These states are derived from oxygen p orbitals and result in deep acceptor transition levels $\varepsilon(0/-)$ and $\varepsilon(-/2-)$, at 0.72 and 1.59 eV above the VBM. Figure 3 shows that interstitial oxygen can exist either as electrically inactive split interstitials O_i^0(split) in semi-insulating and p-type materials or as deep acceptors at the octahedral interstitial site O_i^{2-}(oct) in n-type materials ($E_F > 2.8$ eV). Note that for both forms, the formation energies are very high (except under extreme O-rich conditions), and we do not expect oxygen interstitials to be present in significant concentrations under equilibrium conditions.

The relaxation pattern around the oxygen atom at the octahedral interstitial site is the reverse of the pattern around the zinc interstitial. The oxygen interstitial is displaced along the $[000\bar{1}]$ direction toward the basal plane formed by the three zinc nearest neighbors, as shown in Fig. 10(c). For the most relevant 2− charge state, the resulting O_i-Zn distances are almost equal to d_0 and the O_i-O distances are about $1.2d_0$. The relaxation pattern is opposite to that of the zinc interstitial, which is displaced slightly toward the oxygen plane.

2. Migration

The migration path and calculated energy barrier for the oxygen split interstitial O_i^0(split) are shown in Fig. 11(a), and for the interstitial oxygen at the octahedral site O_i^{2-}(oct) in Fig. 11(b). For the split interstitial O_i^0(split), we focused on migration of the lowest-energy configuration only [Fig. 10(a)]; the calculated migration barrier is 0.9 eV. For the octahedral-site interstitial O_i(oct), we considered only the migration in the most relevant 2− charge state. The migration barrier for O_i^{2-}(oct) through the hexagonal channel along the c axis is 1.1 eV. Migration through the kick-out mechanism is a high-energy path, because the saddle point of the kick-out process involves a configuration similar to the split interstitial; since the split interstitial is not stable in the 2− charge state, this is a high-energy configuration.

G. Zinc antisites

1. Formation energy, atomic geometry, and transition levels

Zinc antisite defects nominally consist of a zinc atom sitting on the "wrong" lattice site, i.e., on an oxygen site (Zn_O).

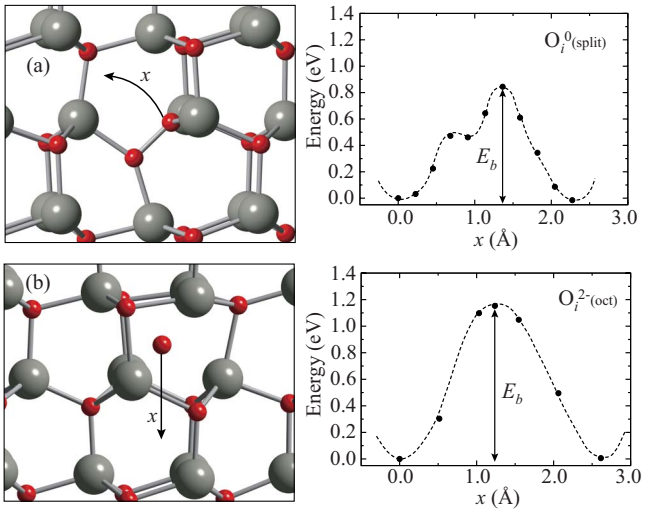

FIG. 11. (Color online) (a) Migration path and calculated energy barrier of oxygen split interstitial O_i^0(split). (b) Migration path and calculated energy barrier of interstitial oxygen at the octahedral site O_i^{2-}(oct).

We find that zinc on an oxygen site significantly lowers its energy by assuming a lower-symmetry off-site configuration. The zinc atom is displaced by more than 1 Å from the substitutional lattice site toward two next-nearest-neighbor oxygen atoms along the $[10\bar{1}0]$ direction, as shown in Fig. 12. The resulting Zn_O-O interatomic distances are only 8% larger than the equilibrium Zn-O bond length. At this equilibrium configuration, we find three Zn_O-Zn distances of ~2.4 Å and one Zn_O-Zn distance of ~2.8 Å. One can think of this low-symmetry configuration of Zn_O as a complex of a zinc interstitial and an oxygen vacancy. We find a doubly occupied a_1 state in the lower part of the band gap that resembles the a_1 state of the oxygen vacancy, and another doubly occupied state resonant in the conduction band that resembles the a_1 state of the zinc interstitial. The very different character of these defect states leads to quite different $\Delta\epsilon$ values for different charge states (see Table II), which are consistently taken into account in the corrections for formation energies in Table III.

We find that the 3+ charge state is not stable for any position of the Fermi level in the band gap, with the $\varepsilon(4$

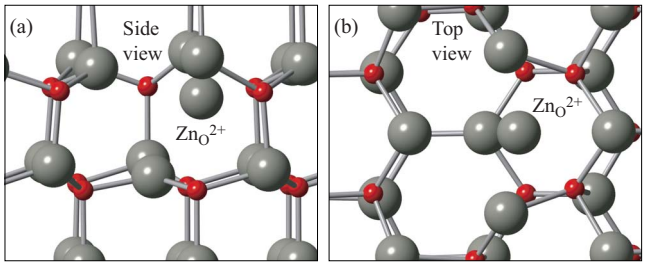

FIG. 12. (Color online) Local atomic geometry of the zinc antisite in the 2+ charge state (Zn_O^{2+}), showing a large displacement off the substitutional site. (a) Side view perpendicular to the c axis. (b) Top view parallel to the c axis (along the $[000\bar{1}]$ direction).

+/2+) at 1.45 eV above the valence-band maximum. The finite size of the supercell, in principle, requires corrections due to charged defect-defect interactions. For reasons discussed in Sec. II C, such corrections are not included here. Since such corrections scale with q^2, they could become significant for a charge state as high as $q=4$. These corrections can be expected to increase the formation energy of Zn_O^{4+} and therefore to lower the position of the $\varepsilon(4+/2+)$ in the band gap. Because of these uncertainties, we indicate the formation energy of Zn_i^{4+} as dotted lines in Fig. 3.

The two electrons in the a_1 state resonant in the conduction band can occupy the CBM and cause the zinc antisite to act as a shallow donor, i.e., it is stable in the 2+ charge state Zn_O^{2+} (see Fig. 3) and the $\varepsilon(2+/+)$ and $\varepsilon(+/0)$ transition levels are located above the CBM (Table II). However, zinc antisites have very high formation energies, as shown in Fig. 3; they are thus very unlikely to occur under equilibrium conditions in ZnO for any position of the Fermi level in the band gap.

First-principles calculations by Kohan *et al.*[37] and by Zhang *et al.*[38] have also found Zn_O antisites to be higher in energy than the other donor-type native defects (V_O and Zn_i). The calculations of Oba *et al.*[39] found Zn_O to be comparable in energy to V_O under Zn-rich conditions. All calculations agree that zinc antisites behave as shallow donors. However, none of the earlier studies revealed the off-site relaxation of Zn_O.

2. Migration

Migration of the antisite in effect involves splitting the defect into its Zn_i and V_O constituents. The question then is whether these constituents remain bound and move in concert or whether they move independently; in the latter case, it is unclear whether the antisite can actually reform. A comprehensive investigation of these mechanisms is beyond the scope of the present work, but a reasonable estimate of the dissociation barrier can be obtained as follows. In *n*-type ZnO, the calculated binding energy of Zn_O^{2+}, i.e., the energy required to separate Zn_O^{2+} into Zn_i^{2+} and V_O^0, is 0.7 eV (see Table III). If we add the migration barrier of interstitial zinc (0.6 eV), the estimated dissociation energy of Zn_O^{2+} is 1.3 eV. We therefore expect Zn_O^{2+} to be stable at temperatures of up to ~500 K. This suggests that Zn_O^{2+} may cause *n*-type conductivity if intentionally introduced under nonequilibrium conditions such as high-energy electron irradiation.[20] Indeed, recent experiments on electron-irradiated ZnO reported that the donor defects created by irradiation anneal out at 400 °C.[105] This result is consistent with the stability of Zn_O^{2+}.

H. Oxygen antisites

Formation energy, atomic geometry, and transition levels

Finally, we have investigated oxygen antisites, where an oxygen atom wrongly occupies a site on the zinc sublattice (O_{Zn}). O_{Zn} is an acceptor-type point defect with very high formation energy, even under the most favorable O-rich conditions (see Fig. 3). This makes it very unlikely that O_{Zn} would be present in equilibrium. However, oxygen antisites could potentially be created under nonequilibrium conditions such as under irradiation or ion implantation. We find that oxygen on the ideal zinc site is unstable and spontaneously relaxes to an off-site configuration, as shown in Fig. 13. The oxygen atom is displaced along the $[000\bar{1}]$ direction by more than 0.7 Å and forms a chemical bond with one of the oxygen nearest neighbors. The O-O bond length is 1.46 Å (exactly twice the covalent radius) in the 2− charge state and 1.42 Å in the neutral charge state.

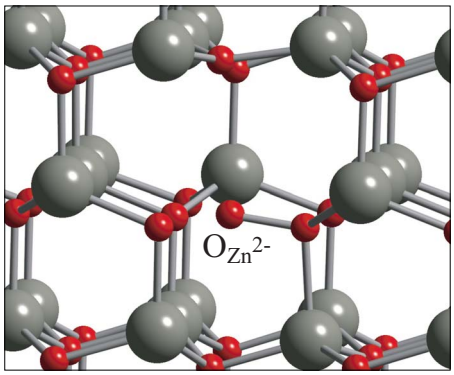

FIG. 13. (Color online) Local atomic geometry for the oxygen antisite O_{Zn} in the 2− charge state, showing a large displacement off the substitutional site.

The distances between the oxygen antisite and the other nearby oxygen atoms are ~2.0 Å, much larger than twice the oxygen covalent radius of 0.73 Å, thus indicating the absence of bonding. Our results indicate that oxygen antisites are deep acceptors with transition levels $\epsilon(0/-)$ and $\epsilon(-/2-)$ at 1.52 and 1.77 eV above the VBM.

Similar to zinc antisites, an investigation of migration paths for oxygen antisites would be quite complicated. We did not consider this effort warranted for a defect with such high formation energies. Qualitatively, we expect the migration barrier to be higher than that of vacancies or interstitials.

Lin *et al.*[92] have suggested that the green luminescence in ZnO corresponds to deep levels induced by oxygen antisites. They observed that the intensity of the green emission increases with oxygen partial pressure in thin films annealed at high temperatures. Based on previous LDA calculations of the defect transition levels, they ruled out zinc vacancies and oxygen interstitials. According to our results, oxygen antisites have much higher formation energies than oxygen interstitials and zinc vacancies, and are therefore unlikely to be present in significant concentrations under equilibrium conditions. As discussed in Sec. III D 4, we feel that zinc vacancies are more likely to be the cause of the observed green luminescence in ZnO, and equally consistent with the results of Lin *et al.*[85]

I. Comparison with previous calculations

First-principles calculations based on DFT-LDA or DFT-GGA for native point defects in ZnO have been reported by a number of groups.[37–43] However, the severe underestimation of the band gap in LDA or GGA calculations leads to

uncertainties in formation energies and transition levels, and the lack of a systematic way to interpret or correct the LDA or GGA results has led to significant differences in the conclusions drawn from such calculations. Zhang *et al.*[38] reported a strong asymmetry in the formation energies for donor-type (V_O, Zn_i, and Zn_O) and acceptor-type (V_{Zn}, O_i, and O_{Zn}) defects, with donor-type defects having significantly lower formation energies. No such asymmetry was observed by Kohan *et al.*[37] or Oba *et al.*,[39] and neither is it evident in our present work; in fact, acceptors seem to potentially exhibit lower formation energies in these studies (depending on stoichiometry conditions, of course). Lee *et al.*[40] and Kohan *et al.*[37] found that Zn_i is a deep donor with transition levels $\varepsilon(2+/+)$ and $\varepsilon(+/0)$ at ~0.5 eV above the VBM, whereas Oba *et al.*[39] found the donor transition level $\varepsilon(2+/0)$ at ~1.2 eV (i.e., above the calculated CBM), and Zhang *et al.*[38] found that Zn_i is a shallow donor.

In an attempt to correct for the LDA deficiency in predicting band gaps, Zhang *et al.*[38] tested a series of empirical and non-self-consistent approaches to correct their LDA results. The lack of self-consistency and the scattered data from different correction methods resulted in significant uncertainties in their calculated formation energies for different charge states of a given defect. One of their results was that zinc vacancies do not introduce transition levels in the band gap, in contrast with our present results as well as the results of other groups,[37,39,40] and also conflicting with recent experimental evidence.[23]

Erhart *et al.*[41,43] performed calculations for vacancy migration barriers and found a large anisotropy. As noted in Sec. III C 5, we feel that these large anisotropies are an artifact of their use of a fairly small (32-atom) supercell. Previous work by Limpijumnong and Van de Walle has shown that 32-atom supercells are not sufficient for calculations of migration barriers in wurtzite-phase semiconductors.[29]

Lany and Zunger[44] performed calculations using LDA and LDA+U. Although their LDA+U results are similar to ours, their analysis and interpretation differ significantly. Lany and Zunger assume that LDA+U only changes the position of the valence band in ZnO. We have previously shown that LDA+U actually affects both the valence band and conduction band.[46] Lany and Zunger also assume that when further corrections are applied to bring the band gap in agreement with experiment, transition levels remain unaffected, i.e., they remain pinned to the valence band. We strongly feel that this assumption is not justified. Indeed, most defect levels exhibit both conduction- and valence-band character. This should be clear on general grounds (e.g., based on the nature of the hybrid orbitals that combine to form the defect states), and it also directly follows from a comparison of the LDA and LDA+U results, as extensively discussed in Secs. II E and III B and illustrated in Tables II and III. The Kohn-Sham states and transition levels for most defects exhibit significant shifts (in fact, shifts that are larger than the valence-band offset $\Delta E_v = 0.34$ eV between LDA and LDA+U that was calculated in Ref. 46), clearly indicating that the states exhibit a finite amount of conduction-band character. Further corrections to the conduction-band position should therefore also lead to further shifts in the transition levels, and indeed we have taken advantage of the information contained in the shift from LDA to LDA+U to calculate the magnitude of this correction [Eq. (6)].

Lany and Zunger's approach[44] leads them to conclude that oxygen vacancies should exist in ZnO in concentration of up to 10^{17} cm^{-3} under equilibrium conditions. If this were the case, then oxygen vacancies should be readily observable in magnetic resonance experiments. In fact, as discussed in Sec. III C 4, an unambiguous EPR signature of the oxygen vacancy has been identified, and this signature has never been observed in as-grown material, but *only* in samples subjected to high-energy electron irradiation (where vacancies can be formed in nonequilibrium concentrations). This result is consistent with the high formation energy of oxygen vacancies as found in our present study and demonstrates the shortcomings of Lany and Zunger's approach for calculating formation energies.

Recently, Patterson[45] performed DFT calculations using a hybrid functional approach that yields a calculated band gap of 3.34 eV. While such calculations are valuable, Patterson's analysis of the position of transition levels is unfortunately flawed. Instead of calculating the transition levels from total-energy differences [see Eq. (5)], Patterson related their position directly to the position of the Kohn-Sham states. This procedure leads to large errors when the different charge states exhibit large and different local lattice relaxations as in the case of V_O. However, an analysis of Patterson's "raw data" is informative: the calculated band structure in Ref. 45 shows that the position of the a_1 state of V_O^0 is 1 eV higher in the hybrid functional calculation than in LDA. This shift is in very good agreement with our extrapolation scheme.

J. Activation energies for self-diffusion

Now that we have presented our results for formation energies, migration paths, and energy barriers for point defects in ZnO, we can discuss possible mechanisms for self-diffusion and their respective activation energies. Although activation energies Q can be easily computed based on the calculated formation energies and migration barriers using Eq. (1), the comparison with experimental measurements is far from straightforward. Formation energies of point defects and, consequently, also activation energies depend on the chemical potential μ_{Zn} (or μ_O) and can also depend on the position of the Fermi level, i.e., $Q=Q(\mu_{Zn},E_F)$. μ_{Zn} can vary from $\mu_{Zn}=E_{tot}(\mathrm{Zn})$ (extreme Zn-rich conditions) to $\mu_{Zn}=E_{tot}(\mathrm{Zn})+\Delta H^f(\mathrm{ZnO})=E_{tot}(\mathrm{Zn})-3.6$ eV (extreme O-rich conditions), and E_F can vary from $E_F=0$ (p type) to $E_F=E_g=3.4$ eV (n type). On the experimental side, it is usually not straightforward to assign specific chemical potential values to the growth or annealing conditions, and often the exact position of the Fermi level has not been established. A direct comparison between experimental and theoretical values of self-diffusion activation energies thus, in principle, requires a major experimental effort to perform measurements under controlled and well defined conditions. Nevertheless, it is possible to draw some important conclusions based on the calculated activation energies and relate them to the available experimental data.

TABLE V. Calculated migration barriers E_b and activation energies $Q=E^f+E_b$ for possible self-diffusion mechanisms of zinc and oxygen species in ZnO. The formation energies E^f are taken from Fig. 3. Activation energies are given for Zn-rich conditions and for three possible values of the Fermi level: $E_F=E_g$ for n-type, $E_F=E_g/2$ for semi-insulating, and $E_F=0$ for p-type material. Absence of a value indicates that the particular charge state is not stable under those conditions.

Mediating defect	E_b (eV)	Q (eV) $E_F=E_g$	$E_F=E_g/2$	$E_F=0$
Zn_i^{2+}	0.57	6.98	3.55	
V_{Zn}^{2-}	1.40	2.97	6.40	
V_O^0	2.36	6.08		
V_O^{2+}	1.70		4.53	
O_i^0(split)	0.87	6.11	6.11	6.11
O_i^{2-}(oct)	1.14	5.14	8.57	

Most of experiments so far have been performed on n-type or semi-insulating ZnO. In Table V, we list the calculated diffusion activation energies for the vacancy- and interstitial-mediated mechanisms for three representative doping conditions: n type, where we assume the Fermi level to be at the CBM ($E_F=E_g$), semi-insulating, where we assume the Fermi level to be midgap ($E_F=E_g/2$), and p type, where we assume the Fermi level to be at the VBM ($E_F=0$). We have omitted cases where the formation energy is negative because such negative energies indicate that the defect would form in such abundance that the Fermi level would be moved to a different position where the formation energy is positive. By inspecting Fig. 3, we believe that most experiments are performed under conditions that are close to Zn rich. Otherwise, the samples would be heavily compensated by zinc vacancies. Therefore, we assume Zn-rich conditions in the following analysis. Values for other conditions or other doping levels can be easily obtained by referring back to Eq. (3).

Zinc self-diffusion in ZnO can be mediated by Zn vacancies or Zn interstitials. In n-type ZnO, we note from Table V that despite the low-energy barrier for the interstitial-mediated mechanism, Zn self-diffusion is predicted to be mediated by Zn vacancies due to the significantly lower formation energy of the vacancies. This result agrees with experiments by Tomlins *et al.*[107] who reported an activation energy of 3.86 eV for Zn self-diffusion in ZnO and suggested that it is controlled by a vacancy-mediated mechanism. Similar to other compound semiconductors,[29,32,33] the migration barrier for the cation vacancies is higher than for interstitials. However, the formation energy of the zinc vacancy (an acceptor) is much lower than that of the zinc interstitial (a donor) under n-type conditions. Table V shows that zinc interstitials will play a more important role in Zn self-diffusion in semi-insulating and p-type ZnO where the formation energy of Zn interstitials is quite low and that of Zn vacancies is quite high.

Oxygen self-diffusion in ZnO can be mediated by O vacancies or O split interstitials [O_i(split)]. Although oxygen interstitials at the octahedral site [O_i(oct)] have a low migration barrier, they do not contribute to self-diffusion. In order to contribute to self-diffusion, the interstitials would ultimately need to become substitutional again, by exchanging positions with oxygen atoms at the regular lattice sites. As already discussed in Sec. III F 2, our calculations indicate that this process involves a high energy barrier.

In n-type ZnO, our calculations indicate that both V_O^0 and O_i(split) will "equally" contribute to oxygen self-diffusion with activation energies of 6.08 and 6.11 eV, respectively. On the other hand, in semi-insulating and p-type ZnO, oxygen self-diffusion will be mediated by V_O^{2+}, since the formation energy of V_O^{2+} decreases with the Fermi-level position. Our calculated activation energy in this case is 4.53 eV, in reasonable agreement with the experiments by Tomlins *et al.*,[16] who reported an activation energy between 3.6 and 4.2 eV for oxygen self-diffusion in semi-insulating ZnO crystals.

K. Defect-annealing temperatures

Based on our results for migration paths and energy barriers for point defects in ZnO, we can also discuss the temperatures at which we expect individual defects to become mobile. Comparison of these temperatures with experimental observations of changes brought about by annealing can aid in the identification of specific defect species. According to transition state theory,[36] an atom near a vacancy can jump to the vacancy (or an interstitial can jump to the next interstitial site) over an energy barrier E_b with a frequency

$$\Gamma = \Gamma_0 \exp\left(-\frac{E_b}{k_B T}\right), \tag{9}$$

where the prefactor Γ_0 is the ratio of the vibrational frequencies at the initial configuration to the frequencies at the saddle point, k_B is the Boltzmann constant, and T is the temperature. A reasonable estimate of the temperature at which a defect becomes mobile can be obtained by taking the usual definition of the activation temperature, i.e., the temperature at which the jump rate Γ in Eq. (9) is $1/s$. Detailed calculations of the prefactor are beyond the scope of the present investigation, but to a good approximation Γ_0 can be taken as a typical phonon frequency, i.e., 10^{13} s^{-1}. Therefore we can use $\Gamma=1$ s^{-1}, $\Gamma_0=10^{13}$ s^{-1}, and the calculated E_b values for each defect to estimate an annealing temperature based on Eq. (9). Comparison with established cases indicates that this method for estimating the annealing temperature typically slightly overestimates the temperature at which defects become mobile.

In Table VI, we list the estimated annealing temperatures for the various point defects in ZnO for which we computed the migration barriers. Only the lowest migration barrier for a given defect is considered here. We see that the low migration barrier of zinc interstitials results in very low annealing temperatures, well below room temperature, in good agreement with recent experimental data from low-temperature irradiation experiments.[17,23,34] During low-temperature electron irradiation, Frenkel pairs (zinc interstitial+zinc vacancy) are created; these have been observed to annihilate at temperatures between 100 and 200 K in optically detected

TABLE VI. Estimated annealing temperature T_{anneal} for vacancies and interstitials in ZnO based on transition state theory as described in the text.

Defect	E_b (eV)	T_{anneal} (K)
Zn_i^{2+}	0.57	219
V_{Zn}^{2-}	1.40	539
V_O^{2+}	1.70	655
V_O^0	2.36	909
O_i^0(split)	0.87	335
O_i^{2-}(oct)	1.14	439

electron paramagnetic resonance (ODEPR) measurements by Watkins and co-workers.[17,23] Our results clearly indicate that the zinc interstitial is the mobile species responsible for the healing of the crystal at low temperatures.[108]

Our results also indicate that oxygen interstitials can diffuse at modest temperatures and are the mobile species responsible for defect recombination on the oxygen sublattice. Oxygen and zinc vacancies become mobile only at higher temperatures, in agreement with experimental observations.[17,24] Indeed, Vlasenko and Watkins have reported that oxygen vacancies are stable up to ~670 K.[17] Nikitenko *et al.*[110] reported annealing of oxygen vacancies (created by irradiation and identified by EPR) in the range 530–630 K. According to Table VI, this is consistent with motion of V_O^{2+}. Indeed, we expect the 2+ charge state to be stable in the Li-doped hydrothermal ZnO crystals used in that study. Tuomisto *et al.*[35] identified both V_O and V_{Zn} using positron annihilation spectroscopy in irradiated samples and observed them annealing out at temperatures between 500 and 600 K, again consistent with our calculated barriers. Look *et al.*,[111] finally, observed that a Zn-sublattice related acceptor anneals out starting at around 300 °C; this is in good agreement with our finding for the zinc vacancy.

We note that our results support the fact that the observed radiation hardness of ZnO (Refs. 20 and 111) can be attributed to rapid defect annihilations that can take place already at temperatures around 200 K for the Zn sublattice and around 400 K for the oxygen sublattice. This property can be advantageous for electronic and photonic applications in high-radiation environments.

IV. SUMMARY

We have performed a detailed first-principles study of the electronic structure and diffusivity of all native point defects in ZnO. We have discussed a method to correct defect transition levels and formation energies based on the LDA and LDA+U calculations. Contrary to the conventional wisdom, we find that isolated native point defects are unlikely to be the cause of the frequently observed unintentional n-type conductivity in ZnO. Our results show that oxygen vacancies are deep donors and have high formation energies in n-type samples; however, they can compensate p-type doping. Zinc interstitials are shallow donors, but have high formation energies under n-type conditions; moreover, they are fast diffusers and hence unlikely to be stable as isolated point defects. Zinc antisites are also shallow donors, but they have high formation energies in n-type samples. Zinc antisites show a large off-site displacement and induce a large local lattice relaxation.

As an alternative explanation for the observed background n-type conductivity in ZnO, we suggest the unintentional incorporation of donor impurities. Among the possible impurities, hydrogen stands out as a likely candidate.[101] Recently, we have shown that hydrogen can substitute on an oxygen site and form a multicenter bond with the four nearest-neighbor Zn atoms.[112] Substitutional hydrogen (H_O) has a low formation energy, acts as a shallow donor, and can explain the variation of conductivity with oxygen partial pressure, consistent with a wide range of experimental results.[14,15,107,113,114]

Zinc vacancies are deep acceptors and have low formation energies under n-type conditions; they can therefore occur as compensating defects in n-type samples. We suggest that zinc vacancies are a possible source of the often-observed green luminescence in ZnO. Oxygen interstitials have high formation energies and are not expected to exist in significant concentrations. They can exist as electrically inactive split interstitials or as deep acceptors at the octahedral site in n-type samples. In p-type samples, the neutral split interstitials are predicted to be more favorable. Oxygen antisites have the highest formation energies among the acceptor-type native point defects. They are deep acceptors and also show large off-site displacements, in which the oxygen atom bonds chemically to only one of the oxygen nearest neighbors.

The migration barriers of all the point defects are modest, explaining why radiation damage can be annealed out at relatively low temperatures (with some recovery already taking place below room temperature). Zinc interstitials diffuse through the kick-out mechanism with a rather low migration barrier of 0.57 eV, in agreement with experimental observations, and are responsible for the observed fast recovery of the electrical properties in irradiated ZnO. The migration barrier of oxygen interstitials in the octahedral configuration (relevant for n-type samples) is 1.1 eV, whereas zinc and oxygen vacancies diffuse with somewhat higher migration barriers of 1.4 and 2.4 eV, respectively. We note that these low migration barriers imply that most point defects will be highly mobile at the temperatures at which ZnO crystals and epilayers are commonly grown, indicating that such growth processes can be considered to be near equilibrium. Finally, our results provide a guide to more refined experiments to probe the influence of individual point defects on the electronic properties of ZnO.

ACKNOWLEDGMENTS

We are grateful to G. D. Watkins, L. Halliburton, and D. Segev for valuable discussions. This work was supported by the NSF MRSEC Program under Grant No. DMR05-20415. It made use of the CNSI Computing Facility under NSF Grant No. CHE-0321368, and also of DataStar at the San Diego Super Computer Center.

[1] M. Lannoo and J. Bourgoin, *Point Defects in Semiconductors I: Theoretical Aspects* (Springer-Verlag, Berlin, 1981); *Point Defects in Semiconductors II: Experimental Aspects* (Springer-Verlag, Berlin, 1983).
[2] *Deep Centers in Semiconductors: A State-of-the-Art Approach*, 2nd ed., edited by S. T. Pandelides (Gordon and Breach Science, Yverdon, 1992).
[3] *Identification of Defects in Semiconductors*, edited by M. Stavola, Semiconductors and Semimetals Vol. 51B (Academic, San Diego, 1999).
[4] S. E. Harrison, Phys. Rev. **93**, 52 (1954).
[5] A. R. Hutson, Phys. Rev. **108**, 222 (1957).
[6] D. G. Thomas, J. Phys. Chem. Solids **3**, 229 (1957).
[7] G. P. Mohanty and L. V. Azároff, J. Chem. Phys. **35**, 1268 (1961).
[8] A. Hausmann, Z. Phys. **237**, 86 (1970).
[9] A. Hausmann and W. Tuerle, Z. Phys. **259**, 189 (1973).
[10] K. Hoffmann and D. Hahn, Phys. Status Solidi A **24**, 637 (1974).
[11] B. Utsch and A. Hausmann, Z. Phys. B **21**, 27 (1975).
[12] A. Hausmann and B. Utsch, Z. Phys. B **21**, 217 (1975).
[13] K. I. Hagemark, J. Solid State Chem. **16**, 293 (1976).
[14] G. Neumann, in *Current Topics in Materials Science*, edited by E. Kaldis (North-Holland, Amsterdam, 1981), Vol. 7, p. 152.
[15] F. A. Kröger, *The Chemistry of Imperfect Crystals* (North-Holland, Amsterdam, 1974).
[16] G. W. Tomlins, J. L. Routbort, and T. O. Mason, J. Am. Ceram. Soc. **81**, 869 (1998).
[17] L. S. Vlasenko and G. D. Watkins, Phys. Rev. B **72**, 035203 (2005).
[18] D. C. Reynolds, D. C. Look, B. Jogai, and H. Morkoç, Solid State Commun. **101**, 643 (1997).
[19] D. C. Reynolds, D. C. Look, B. Jogai, J. E. Van Nostrand, R. Jones, and J. Jenny, Solid State Commun. **106**, 701 (1998).
[20] D. C. Look, J. W. Hemsky, and J. R. Sizelove, Phys. Rev. Lett. **82**, 2552 (1999).
[21] W. E. Carlos, E. R. Glaser, and D. C. Look, Physica B **308-310**, 976 (2001).
[22] N. Y. Garces, N. C. Giles, L. E. Halliburton, G. Cantwell, D. B. Eason, D. C. Reynolds, and D. C. Look, Appl. Phys. Lett. **80**, 1334 (2002).
[23] Yu. V. Gorelkinskii and G. D. Watkins, Phys. Rev. B **69**, 115212 (2004).
[24] L. S. Vlasenko and G. D. Watkins, Phys. Rev. B **71**, 125210 (2005).
[25] G. Neumann, in *Current Topics in Materials Science*, edited by E. Kaldis (North-Holland, Amsterdam, 1981), Vol. 7, p. 279.
[26] M. E. Glicksman, *Diffusion in Solids: Field Theory, Solid-State Principles and Applications* (Wiley, New York, 2000).
[27] C. S. Nichols, C. G. Van de Walle, and S. T. Pantelides, Phys. Rev. Lett. **62**, 1049 (1989).
[28] A. Janotti, A. Fazzio, R. Mota, and P. Piquini, Solid State Commun. **110**, 457 (1999).
[29] S. Limpijumnong and C. G. Van de Walle, Phys. Rev. B **69**, 035207 (2004).
[30] C. G. Van de Walle and J. Neugebauer, J. Appl. Phys. **95**, 3851 (2004).
[31] A. Janotti, M. Krčmar, C. L. Fu, and R. C. Reed, Phys. Rev. Lett. **92**, 085901 (2004).
[32] G. D. Watkins, in *Radiation Effects in Semiconductors, 1976*, edited by N. B. Urli and J. W. Corbett, IOP Conf. Proc. No. 31 (Institute of Physics, London, 1976), p. 95.
[33] K. H. Chow and G. D. Watkins, Phys. Rev. Lett. **81**, 2084 (1998).
[34] C. Coskun, D. C. Look, G. C. Farlow, and J. R. Sizelove, Semicond. Sci. Technol. **19**, 752 (2004).
[35] F. Tuomisto, K. Saarinen, D. C. Look, and G. C. Farlow, Phys. Rev. B **72**, 085206 (2005).
[36] G. H. Vineyard, J. Phys. Chem. Solids **3**, 121 (1957).
[37] A. F. Kohan, G. Ceder, D. Morgan, and Chris G. Van de Walle, Phys. Rev. B **61**, 15019 (2000).
[38] S. B. Zhang, S.-H. Wei, and A. Zunger, Phys. Rev. B **63**, 075205 (2001).
[39] F. Oba, S. R. Nishitani, S. Isotani, H. Adachi, and I. Tanaka, J. Appl. Phys. **90**, 824 (2001).
[40] E.-C. Lee, Y.-S. Kim, Y.-G. Jin, and K. J. Chang, Phys. Rev. B **64**, 085120 (2001).
[41] P. Erhart, A. Klein, and K. Albe, Phys. Rev. B **72**, 085213 (2005).
[42] C. G. Van de Walle, Physica B **308-310**, 899 (2001).
[43] P. Erhart and K. Albe, Phys. Rev. B **73**, 115207 (2006).
[44] S. Lany and A. Zunger, Phys. Rev. Lett. **98**, 045501 (2007).
[45] C. H. Patterson, Phys. Rev. B **74**, 144432 (2006).
[46] A. Janotti, D. Segev, and C. G. Van de Walle, Phys. Rev. B **74**, 045202 (2006).
[47] A. Janotti and C. G. Van de Walle, Phys. Rev. B **75**, 121201(R) (2007).
[48] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[49] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[50] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994); G. Kresse and D. Joubert, *ibid.* **59**, 1758 (1999).
[51] G. Kresse and J. Hafner, Phys. Rev. B **47**, R558 (1993); G. Kresse, thesis, Technische Universität Wien, 1993; G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996); Comput. Mater. Sci. **6**, 15 (1996).
[52] G. Herzberg, *Spectra of Diatomic Molecules*, 2nd ed. (Van Nostrand, Princeton, NJ, 1950).
[53] J. A. Dean, *Lange's Handbook of Chemistry*, 14th ed. (McGraw-Hill, New York, 1992).
[54] G. Makov and M. C. Payne, Phys. Rev. B **51**, 4014 (1995).
[55] J. Shim, E.-K. Lee, Y.-J. Lee, and R. M. Nieminen, Phys. Rev. B **71**, 035206 (2005).
[56] P. M. Mooney, in *Identification of Defects in Semiconductors*, edited by M. Stavola, Semiconductors and Semimetals Vol. 51B (Academic, San Diego, 1999), p. 93.
[57] S.-H. Wei and A. Zunger, Phys. Rev. B **37**, 8958 (1988).
[58] V. I. Anisimov, I. V. Solovyev, M. A. Korotin, M. T. Czyzyk, and G. A. Sawatzky, Phys. Rev. B **48**, 16929 (1993); A. I. Liechtenstein, V. I. Anisimov, and J. Zaanen, *ibid.* **52**, R5467 (1995).
[59] *Semiconductors—Basic Data*, 2nd revised ed., edited by O. Madelung (Springer, Berlin, 1996).
[60] C. G. Van de Walle (unpublished).
[61] H. Schulz and K. H. Thiemann, Solid State Commun. **23**, 815 (1977).
[62] S. Lany and A. Zunger, Phys. Rev. B **72**, 035215 (2005).
[63] X. Zhu and S. G. Louie, Phys. Rev. B **43**, 14142 (1991).
[64] A. Janotti and C. G. Van de Walle, J. Cryst. Growth **287**, 58 (2006).
[65] F. Tuomisto, K. Saarinen, and D. C. Look, Phys. Status Solidi A **201**, 2219 (2004).
[66] A. Janotti and C. G. Van de Walle, Appl. Phys. Lett. **87**, 122102 (2005).

[67] D. Galland and A. Herve, Solid State Commun. **14**, 953 (1974).
[68] C. Gonzalez, D. Galland, and A. Herve, Phys. Status Solidi B **72**, 309 (1975).
[69] J. M. Smith and W. E. Vehse, Phys. Lett. **31A**, 147 (1970).
[70] V. Soriano and D. Galland, Phys. Status Solidi B **77**, 739 (1976).
[71] K. A. Müller and J. Schneider, Phys. Lett. **4**, 288 (1963).
[72] K. M. Sancier, Surf. Sci. **21**, 1 (1970).
[73] G. Neumann, in *Current Topics in Materials Science*, edited by E. Kaldis (North-Holland, Amsterdam, 1981), p. 269.
[74] J. Schneider and A. Räuber, Z. Naturforsch. A **16a**, 712 (1961).
[75] P. H. Kasai, Phys. Rev. **130**, 989 (1963).
[76] C. H. Geisler and G. L. Simmons, Phys. Lett. **11**, 111 (1964).
[77] R. B. Lal and G. M. Arnett, J. Phys. Soc. Jpn. **21**, 2743 (1966).
[78] G. K. Born, A. B. Hofstaetter, A. O. Scharmann, G. M. Arnett, R. l. Kroes, and U. E. Wegner, Phys. Status Solidi A **4**, 675 (1971).
[79] T. Mookherji, Phys. Status Solidi A **13**, 293 (1972).
[80] A. Pöppl and G. Völkel, Phys. Status Solidi A **115**, 247 (1989).
[81] A. Pöppl and G. Völkel, Phys. Status Solidi A **121**, 195 (1990).
[82] F. H. Leiter, H. R. Alves, A. Hofstaetter, D. M. Hofmann, and B. K. Meyer, Phys. Status Solidi B **226**, R4 (2001).
[83] F. Leiter, H. Alves, D. Pfisterer, N. G. Romanov, D. M. Hofmann, and B. K. Meyer, Physica B **340-342**, 201 (2003).
[84] F. Tuomisto, V. Ranki, K. Saarinen, and D. C. Look, Phys. Rev. Lett. **91**, 205502 (2003).
[85] S. Limpijumnong, S. B. Zhang, S. H. Wei, and C. H. Park, Phys. Rev. Lett. **92**, 155504 (2004).
[86] A. L. Taylor, G. Filipovich, and G. K. Lindeberg, Solid State Commun. **8**, 1359 (1970).
[87] D. Galland and A. Herve, Phys. Lett. **33A**, 1 (1970).
[88] G. D. Watkins, in *Defect Control in Semiconductors*, Proceedings of the International Conference on the Science and Technology of Defect Control in Semiconductors, Yokohama, 1989, edited by K. Sumino (North-Holland, Amsterdam, 1990), p. 933.
[89] D. B. Laks, C. G. Van de Walle, G. F. Neumark, P. E. Blöchl, and S. T. Pantelides, Phys. Rev. B **45**, 10965 (1992).
[90] S. Pöykkö, M. J. Puska, and R. M. Nieminen, Phys. Rev. B **57**, 12174 (1998).
[91] R. B. Lauer, J. Phys. Chem. Solids **34**, 249 (1973).
[92] B. Lin, Z. Fu, and Y. Jia, Appl. Phys. Lett. **79**, 943 (2001).
[93] R. Dingle, Phys. Rev. Lett. **23**, 579 (1969).
[94] K. C. Mishra, P. C. Schmidt, K. H. Johnson, B. G. DeBoer, J. K. Berkowitz, and E. A. Dale, Phys. Rev. B **42**, 1423 (1990).
[95] J. Neugebauer and C. G. Van de Walle, Appl. Phys. Lett. **69**, 503 (1996).
[96] K. Vanheusden, C. H. Seager, W. L. Warren, D. R. Trallant, J. Caruso, M. J. Hampden-Smith, and T. T. Kodas, J. Lumin. **75**, 11 (1997).
[97] F. A. Kröger and H. J. Vink, J. Chem. Phys. **22**, 250 (1954).
[98] S. A. Studenikin, N. Golego, and M. Cocivera, J. Appl. Phys. **84**, 2287 (1998).
[99] K. Vanheusden, W. L. Warren, C. H. Seager, D. R. Trallant, and J. A. Voigt, J. Appl. Phys. **79**, 7983 (1996).
[100] T. Sekiguchi, N. Ohashi, and Y. Terada, Jpn. J. Appl. Phys., Part 2 **36**, L289 (1997).
[101] C. G. Van de Walle, Phys. Rev. Lett. **85**, 1012 (2000).
[102] E. V. Lavrov, J. Weber, F. Börrnert, C. G. Van de Walle, and R. Helbig, Phys. Rev. B **66**, 165205 (2002).
[103] R. Linder, Acta Chem. Scand. (1947-1973) **6**, 457 (1952).
[104] E. A. Secco and W. J. Moore, J. Chem. Phys. **26**, 942 (1957).
[105] D. C. Look, G. C. Farlow, P. Reunchan, S. Limpijumnong, S. B. Zhang, and K. Nordlund, Phys. Rev. Lett. **95**, 225502 (2005).
[106] S. Limpijumnong, X. Li, S.-H. Wei, and S. B. Zhang, Appl. Phys. Lett. **86**, 211910 (2005).
[107] G. W. Tomlins, J. L. Routbort, and T. O. Mason, J. Appl. Phys. **87**, 117 (2000).
[108] K. Lorenz, E. Alves, E. Wendler, O. Bilani, W. Wesch, and M. Hayes, Appl. Phys. Lett. **87**, 191904 (2005).
[109] K. Tarkpea, A. Ots, and V. Nikitenko, J. Phys. Chem. Solids **55**, 1353 (1994).
[110] V. A. Nikitenko, K. É. Tarkpea, I. V. Pykanov, and S. G. Stoyukhin, J. Appl. Spectrosc. **68**, 502 (2001).
[111] D. C. Look, D. C. Reynolds, J. W. Hemsky, R. L. Jones, and J. R. Sizelove, Appl. Phys. Lett. **75**, 811 (1999).
[112] A. Janotti and C. G. Van de Walle, Nat. Mater. **6**, 44 (2007).
[113] S. J. Jokela and M. D. McCluskey, Phys. Rev. B **72**, 113201 (2005).
[114] G. A. Shi, M. Stavola, S. J. Pearton, M. Thieme, E. V. Lavrov, and J. Weber, Phys. Rev. B **72**, 195211 (2005).