First-principles study of native point defects in ZnO

A. F. Kohan, G. Ceder, and D. Morgan
Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusettes 02139

Chris G. Van de Walle
Xerox Palo Alto Research Center, Palo Alto, California 94304

(Received 18 January 2000)

The characterization of native point defects in ZnO is still a question of debate. For example, experimental evidence for ZnO with an excess of Zn is inconclusive as to whether the dominant defects are metal interstitials or oxygen vacancies. This information is essential to understand the behavior of the material and to tailor its numerous technological applications. We use the first-principles pseudopotential method to determine the electronic structure, atomic geometry, and formation energy of native point defects in ZnO. Interstitials, vacancies, and antisites in their relevant charge states are considered and the effects of dopants are also discussed. The results show that both the Zn and O vacancies are the relevant defects in ZnO. We also propose a possible transition mechanism and defect center responsible for the experimentally observed green luminescence.

I. INTRODUCTION

Zinc oxide has been receiving much attention in recent years due to its many technological applications, particularly in varistors and in optical devices. A varistor is a polycrystalline electronic device with pronounced non-Ohmic current-voltage characteristics.[1] The green photoluminescence behavior of ZnO has been of interest for building flat panel displays.[2] Other applications include gas sensors,[3] solar cells,[4,5] catalysts,[5] UV blocking,[5] substrates[6] or buffer[7] layers for growth of GaN, or as a light-emitting material in its own right.[8,9]

To optimize the use of ZnO devices it is essential to obtain a basic physical understanding of its properties. In spite of numerous experimental studies, there is still controversy as to what are the relevant native defects of this oxide. This issue is critical since most of the properties of ZnO depend in one way or another on the defects that are present in the material.

Single-crystal ZnO has always been observed to contain metal excess (or oxygen deficiency) under experimentally attainable zinc and oxygen partial pressures.[10] The metal excess can be accommodated in part by the presence of zinc interstitials or oxygen vacancies. Experiments have been inconclusive as to which of these is the predominant defect. Results presented in the literature point towards both directions and different interpretations have even been taken on the same set of experimental data (see, for example, Ref. 11). Interstitial zinc atoms have been proposed as the dominant defect on the basis of ionic diffusion or size considerations.[10,12–14] Other authors, based on calculation of reaction rates,[11] diffusion experiments,[15] or electrical conductivity and Hall effect measurements[16] concluded that oxygen vacancies were the predominant defect.

It is the objective of this work to shed light on these issues. By using the first principles, plane-wave pseudopotential approach we determine the electronic structure, atomic geometry, and formation energy of native point defects in ZnO. We extensively study all possible interstitials, vacancies, and antisites in their relevant charge states as a function of doping and of zinc and oxygen chemical potentials.

Perfect crystalline ZnO has been the focus of many theoretical analyses. Its crystal and electronic band structure has been studied by various total-energy methods: Hartree-Fock,[17,18] local density approximation (linear muffin-tin orbital,[19] linear combination of atomic orbitals,[20] pseudopotential,[21] and full-potential linearized augmented-plane-wave[22]), self-interaction-corrected pseudopotentials,[23] and the GW approximation.[24]

Computational studies of defects in ZnO are more limited. The self-consistent-field X-α scattered wave cluster molecular orbital method has been used to calculate the position of defects levels.[25] However, to our knowledge only pair-potential models have been used to compute the formation energies of the crystal imperfections in ZnO.[26] The complexity of the calculations involved has prevented the use of more reliable techniques.

In recent years, the combination of increased computer power and improvements on the computational methods have allowed researchers to address, from first principles, very complex problems. Taking advantage of these new developments we study here all the relevant defects that may be present in ZnO. By using the plane-wave soft-pseudopotential technique together with the supercell approach we have found that the dominant native defects in ZnO are oxygen and zinc vacancies.

Our calculations also provide new insights to the study of the luminescent behavior of ZnO. The green luminescence, in particular, has been widely observed, and its origins have been the subject of many speculations: oxygen vacancies (V_O^{+1}),[2,27–30] zinc interstitials,[31,32] transitions from Zn interstitials to Zn vacancies,[33] Zn_O antisites,[34] and extrinsic impurities[35,36] have all been proposed. Our results indicate that the green emission may be attributed to transitions from electrons in the conduction band to a deep acceptor level due to zinc vacancies.

It should be noted that the effects of the association of impurities with intrinsic defects and the interaction between defects have not been considered here. The results should be interpreted within these limitations and are an indication of what the behavior would be if only noninteracting defects were present in the material.

In the following pages, we briefly review the formalism to compute defect concentrations and their levels (Sec. II). Then in Sec. III, we show the results obtained. An interpretation of these calculations follows in Sec. IV.

II. FORMALISM

The formalism to compute defect concentrations and defect levels from first principles is well understood.[37–40] In what follows, we briefly review this procedure. Special attention is given to the fact that we are dealing with compounds (the defect concentration depends on the chemical potentials set by the environment conditions) and to the use of the supercell technique. A brief description of the total-energy techniques to be used are also provided.

A. Defect concentration

At equilibrium and in the dilute limit, the concentration of a defect in a crystal depends upon its free energy of formation, ΔG_f in the following way:

$$C_d = N_{\mathrm{sites}} e^{-\Delta G_f / k_B T}. \tag{1}$$

N_{sites} is the number of sites in the crystal where the defect can occur (per unit volume), k_B is the Boltzmann constant, T the temperature, and ΔG_f corresponds to

$$\Delta G_f = \Delta E_f - T\Delta S_f + P\Delta V_f . \tag{2}$$

Here, ΔE_f is the change in total energy (including chemical potential terms), ΔS_f is mainly the change in vibrational entropy, and ΔV_f is the change in volume when the defect is introduced into the system. Since the contribution of volume changes is relatively small and the changes in entropy are of the same order when comparing different defects, we focus only on computing formation energy terms. In many cases, defect energies are defined as the energy of the system containing the defect minus the energy of a perfect system with the same number of atoms. This is not possible in this work as we will be studying defects that change the composition of the material. The defect formation energies will therefore be defined with respect to a set of external chemical potentials (for Zn and O) and the Fermi energy (if the defect carries a charge).

The formation energy of a charged point defect in ZnO, ΔE_f, is computed as

$$\Delta E_f = E(N_{\mathrm{Zn}}, N_{\mathrm{O}}) - N_{\mathrm{Zn}}\mu_{\mathrm{Zn}} - N_{\mathrm{O}}\mu_{\mathrm{O}} + q\epsilon_F , \tag{3}$$

where $E(N_{\mathrm{Zn}}, N_{\mathrm{O}})$ is the total energy of a system containing N_{Zn} and N_{O} zinc and oxygen atoms (arranged so that a defect is present), μ_{Zn} and μ_{O} are the external zinc and oxygen chemical potentials, q is the charge of the defect (including its sign), and ϵ_F is the Fermi energy. The Fermi level is taken as the energy of the reservoir (chemical potential) from (in) which an electron is removed (placed) to form a *charged* defect. The Zn and O chemical potentials are not independent but related to the total energy of perfect ZnO, E_{ZnO} (per ZnO pair). For ZnO material with only a small concentration of defects the energy to add one Zn and one O atom is the sum of the Zn and O chemical potential. But this should also approximately be equal to the molar energy of ZnO (since the system is enlarged by one ZnO molecule). The sum of the Zn and O chemical potentials therefore has to add to the energy of ZnO. Equation (3) can then be rewritten as

$$\Delta E_f = E(N_{\mathrm{Zn}}, N_{\mathrm{O}}) - (N_{\mathrm{Zn}} - N_{\mathrm{O}})\mu_{\mathrm{Zn}} - N_{\mathrm{O}} E_{\mathrm{ZnO}} + q\epsilon_F . \tag{4}$$

Clearly, from Eq. (4), ΔE_f for perfect stoichiometric ZnO is zero, indicating a proper reference state. In this first *ab initio* study of defect energies in ZnO, we focus on dilute native point defects. Although the existence of complexes cannot be ruled out *a priori*, a comprehensive treatment from a first-principles perspective is still difficult. Consequently we only compute total energies for isolated point defects in an infinite crystal. The infinite limit is approximated here by placing the defect in a large ZnO cell and periodically repeating this ‘‘supercell’’ in space (supercell approach). The dimensions of the supercell should be such that defects in neighboring supercells do not interact appreciably with each other. To calculate the energy of the supercell, we use the plane-wave soft pseudopotential method.

The chemical potentials determine the off stoichiometry of the system, and depend on different parameters such as partial pressures and growth conditions. ZnO can exist within a range of oxygen and zinc chemical potentials and it is our objective to study the defect formation energies within this range. The boundaries on μ_{Zn} in ZnO are determined by the stability limits of ZnO with respect to metallic Zn and molecular oxygen. To prevent pure Zn formation,

$$\mu_{\mathrm{Zn}}^{\mathrm{ZnO}} < \mu_{\mathrm{Zn}}^{o}, \tag{5}$$

where $\mu_{\mathrm{Zn}}^{\mathrm{ZnO}}$ is the chemical potential of Zn in ZnO and μ_{Zn}^{o} is the energy of Zn in the standard state (pure metal).

To prevent oxygen loss,

$$\mu_{\mathrm{O}}^{\mathrm{ZnO}} < \mu_{\mathrm{O}}^{o}, \tag{6}$$

where $\mu_{\mathrm{O}}^{\mathrm{ZnO}}$ is the chemical potential of O in ZnO and μ_{O}^{o} is the energy of pure O_2 gas (per oxygen atom). Equation (6) can be written in terms of a condition on the zinc chemical potential by using the formation energy of ZnO from metallic Zn and O_2 gas (assuming off-stoichiometry is small),

$$\Delta E_f^{\mathrm{ZnO}} = \mu_{\mathrm{Zn}}^{\mathrm{ZnO}} - \mu_{\mathrm{Zn}}^{o} + \mu_{\mathrm{O}}^{\mathrm{ZnO}} - \mu_{\mathrm{O}}^{o} . \tag{7}$$

Combining Eqs. (5), (6), and (7) leads to

$$\mu_{\mathrm{Zn}}^{o} + \Delta E_f^{\mathrm{ZnO}} < \mu_{\mathrm{Zn}}^{\mathrm{ZnO}} < \mu_{\mathrm{Zn}}^{o} . \tag{8}$$

Equation (8) provides a range of $\mu_{\mathrm{Zn}}^{\mathrm{ZnO}}$ for which the defect energies in Eq. (4) should be evaluated. For *intrinsic* ZnO, the value of the Fermi level is determined by requiring electroneutrality in the system. Fermi level values different from the intrinsic ones are also of importance as they represent conditions in *extrinsically* doped ZnO, the dopants being far from the defect we are dealing with (no dopant-defect interaction). We conventionally take ϵ_F to be zero at the top of the valence band and assume its value can vary from the valence band edge up to the conduction band.

B. Total energies and the pseudopotential approach

The computation of total energies from first principles requires finding solutions to the Schrödinger equation. We simplify the many-body electronic problem by making use of density-functional theory[41] in the local-density approximation[42] (LDA). A well-known consequence of using the LDA is the fact that the band gap is usually underpredicted.[43] This error can be critical when dealing with defects whose electronic levels are close to the conduction band. In the discussion section, we will consider a very simple correction consisting of a rigid shift of these levels and the conduction bands so that the bottom of the latter is at the exact same location as in the experimental results.

On the other hand, the use of the LDA considerably simplifies the solution of the Schrödinger equation by mapping it into a single-electron equation. Oxides present a computationally challenging case, due to the strongly localized oxygen potential and the size of the unit cells. Many different approximations have been tested in these systems.[44] The situation is aggravated in defect calculations because large supercells have to be used. Here, we apply the plane-wave soft-pseudopotential approach.[45–47] For the accuracy required in our calculations, the pseudopotential approach is one of the fastest currently available methods. Efforts are being made to develop faster techniques without sacrificing accuracy.[48–51]

In the plane-wave soft-pseudopotential method, the electronic wave functions are expanded in plane waves. The number of these basis functions is controlled by an energy cutoff E_c. Only plane waves with kinetic energies smaller than E_c are used in the expansion. Plane waves are not well suited to represent the rapid variations of the electronic wave functions close to the nuclei. These rapid changes are caused by the orthogonalization to the wave functions of the ionic core. The problem is solved by replacing the potential of the ionic core by a *pseudopotential* that not only incorporates the nuclear but also the electrostatic potential due to the core electrons. The pseudopotential is determined so that the solution of the Schrödinger equation provides valence electron wave functions (*pseudo* wave functions) that are smooth within the atomic core region and that are similar to the actual function everywhere else. In the *soft* pseudopotential approach, the pseudo wave function is not normalized, allowing for a lower E_c.[45]

The accuracy of the solution is not only controlled by E_c but by other factors such as the number of k points used during the integrations in reciprocal space. Tests were performed to find the appropriate values of these parameters as explained in Sec. III B.

C. Supercell size

The strength of the interaction between a defect and its images determines the size of the supercell that needs to be used. The supercell can be determined by placing the defect in supercells with different sizes, the structural parameters relaxed, and the corresponding formation energies plotted against the number of atoms. The supercell to be used would be the one for which increasing its size would not change the formation values more than an allowed error. However, this procedure tends to converge slowly, resulting in computationally unaffordably large cells.

On the other hand, for single isolated defects, the volume of the supercell can be approximated by the one corresponding to the perfect crystal. We can use this information to fix the lattice constant of the supercell to the bulk value, and allow only internal relaxations. Since the magnitude of atomic relaxations is expected to decrease rapidly with distance from the defect, only atoms that are close to the defect are relaxed. This considerably speeds up the computations. We have performed a study of the effect on convergence and accuracy of relaxing different numbers of atoms close to the defect and of performing full relaxations. From this analysis, an appropriate supercell size is deduced. Unfortunately, the use of accurate total energy methods in these tests is out of the question, since the number of atoms in some of the cells is extremely large. We resorted to the use of simpler, but faster, energy models and, when possible, contrast the results against pseudopotential values.

A fast, commonly used technique in oxides is the pair potential model.[52,53] In this approach, the electronic density and ionic charges are replaced by point charges centered at ions (or point charges and non-concentric spherical shells to account for polarization effects). The energy is computed as the sum of the electrostatic interaction and a short-range repulsive pair-potential term representing the overlap of the electronic clouds. Although issues of internal relaxations are dealt with reasonably well within this model, electronic effects, such as the dispersion of defect-induced levels in the gap, are unaccounted for. Within this limitation, we study in Sec. III A the effect of supercell size.

Pair potentials allowed us to deal with supercells of the order of thousands of atoms. Internal relaxations are expected to be important in oxides, especially when there exists a large size mismatch between the ions. In this case, the general convergence behavior of the supercell calculations should be correctly captured.

III. RESULTS

A. Supercell size convergence

In this work, we are concerned with finding the relevant point defects in ZnO. Consequently, we focus on *differences* in concentrations, that in turn are determined by *differences* in defect formation energies. It makes sense then for us to study the convergence with supercell size by analyzing the *difference* in formation energies between two different types of defects. There are two open spaces to place an atom interstitially in the ZnO wurzite structure.[54] One is in an octahedral and the other is in a tetrahedral position. We computed the formation energy of a neutral Zn interstitial placed at both sites and analyzed the difference between these two formation energies ($\delta\Delta E_f$) for different supercells. We addressed both the issue of changes with the number of atoms in the supercell as well as the effect of adopting a constant volume (corresponding to the theoretical lattice constant of perfect ZnO) and relaxing only those atoms close to the defect.

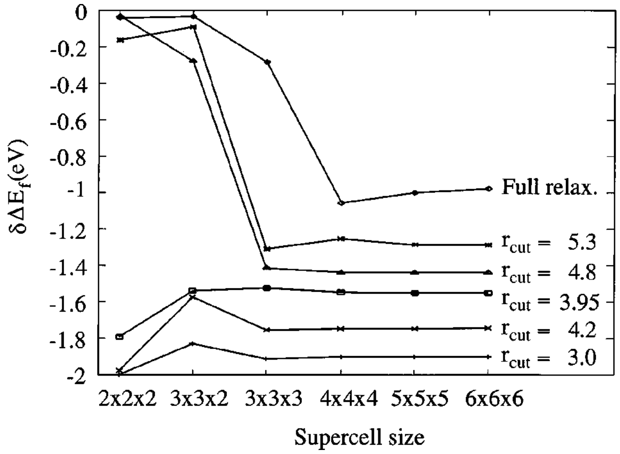

FIG. 1. Variation of the difference in formation energy of a zinc interstitial defect placed at an octahedrally and a tetrahedrally coordinated site ($\delta\Delta E_f$) as a function of supercell size and number of internal atoms that are allowed to relax. The energies are calculated with a pair-potential approach. The supercell sizes are specified as multiples of the primitive wurtzite cell along the a-b-c directions. Except for the fully relaxed case, r_{cut} is the radius of a sphere centered at the defect within which the atoms are relaxed.

As already mentioned, we used a pair potential approach to compute total energies. The parameters of the model were taken from Ref. 53. The atoms that are allowed to relax during the computations are identified as those within the radius of a sphere centered at the defect (here called r_{cut}).

The results obtained are summarized in Fig. 1 where we show $\delta\Delta E_f$ for different supercells and r_{cut}. The line labeled ''full relax.'' corresponds to fully relaxing the volume and internal coordinates. The other lines correspond to fixing the volume to the one computed for perfect ZnO and relaxing the atoms close to the defect. There are four atoms per unit cell in ZnO, so that a $2\times2\times2$ supercell contains 32 atoms, etc. For the $6\times6\times6$ supercell the defects are separated by 19.59 Å.

It is clear from Fig. 1 that *internal relaxations* play a significant role in computing accurate formation energies. By increasing r_{cut} from 3 Å to 5.3 Å, $\delta\Delta E_f$ was reduced from -1.92 eV to -1.29 eV. On the other hand, the effect of volume relaxations is much smaller. If the volume is allowed to relax (away from the perfect crystal volume) at $r_{\mathrm{cut}}=5.3$ Å, $\delta\Delta E_f$ only changes from -1.29 eV to -1.20 eV. Consequently, *using a constant volume supercell with limited internal relaxations constitutes a reliable approach to study defects properties in ZnO*.

Inspection of Fig. 1 and consideration of our available computer resources led us to choose a $3\times3\times2$ supercell with $r_{\mathrm{cut}}=3.95$ Å. In this case, the pair-potential analysis suggests that errors of the order of 0.3–0.4 eV should be expected in the relative formation energies. Note that if a fully relaxed supercell of the same size were used, the estimated errors would almost triple. In the $3\times3\times2$ supercell point defects have a minimum separation of 9.6 Å.

We also investigated supercell-size convergence using the first-principles pseudopotential approach. Due to computational constraints this study was limited to smaller cells, namely $2\times2\times1$, $2\times2\times2$, and $3\times3\times2$ supercells. $\delta\Delta E_f$ changed from -3.0 eV to -0.9 eV to -1.19 eV for a $2\times2\times1$, $2\times2\times2$, and $3\times3\times2$ supercell respectively. On the other hand, in the pseudopotential case the difference in $\delta\Delta E_f$ between a *fully relaxed* and a *constant volume* calculation is smaller (0.4 eV and 0.11 eV for a $2\times2\times1$ and a $2\times2\times2$ supercell respectively). Consequently, the error on the formation values coming from supercell size effects can be estimated to be lower than in the pair potential case and of the order of 0.1–0.2 eV. This error is small enough to not change the conclusions of the remaining sections. (Even if the more conservative 0.3–0.4 eV error value estimated from pair potentials were adopted, the main conclusions would remain unchanged.)

TABLE I. Computed and experimentally measured lattice parameters for a defect-free ZnO wurzite structure.

	Calculations	Experiments (Ref. 57)
a	3.19 Å	3.25 Å
c	5.15 Å	5.21 Å
u (internal)	0.379	0.382

B. Perfect-crystal ZnO

We computed the properties of the defect-free ZnO wurzite structure using the pseudopotential method. We used ultrasoft[45] pseudopotentials with $2s^2 2p^4$ and $3d^{10}4s^2$ as the valence-electron configuration for the oxygen and zinc atoms respectively. There is a substantial interaction between the zinc *d* and *s* electrons which makes it necessary to consider the *d* electrons as part of the valence.[21] We also included core corrections for Zn.[55] The Perdew-Zunger[56] parametrization of the exchange and correlation potential was always used.

Convergence with respect to *k* points and energy cutoff was carefully checked. All calculations in this paper were done with $E_c=400$ eV. A $6\times6\times6$ uniform mesh for *k*-space integrations (28 independent *k* points) provided total energies converged within 1 meV/atom.

The computed lattice parameters are shown in Table I. These lattice parameters were then used to build the ZnO supercells.

As already mentioned, LDA tends to underpredict the band gap energy. This is certainly the case here. Figure 2 shows the computed band structure. The predicted band gap energy is 0.91 eV, well bellow the experimental[57] 3.4 eV value. This difference could affect the defect formation energies, as discussed in Sec. IV.

In a compound solid, the defect properties depend on the chemical potential of the species involved as shown in Eq. (3). To find the limiting value for the chemical potentials we need the energy of metallic Zn and molecular oxygen [see Eq. (8)]. Consequently, we compute the total energies for these two systems.

Zinc orders in a hexagonal closed packed structure. In this case, we used a $20\times20\times20$ uniform *k*-space mesh (equivalent to 484 irreducible *k* points) with a $\sigma=0.1$ eV temperature smearing. The computed lattice parameters are $a=2.57$ Å (2.66 Å) and $c=4.81$ Å (4.947 Å); the values between parenthesis are the experimental data from Ref. 58.

For oxygen, we considered the energy of an isolated molecule. For computational purposes, the oxygen molecule was

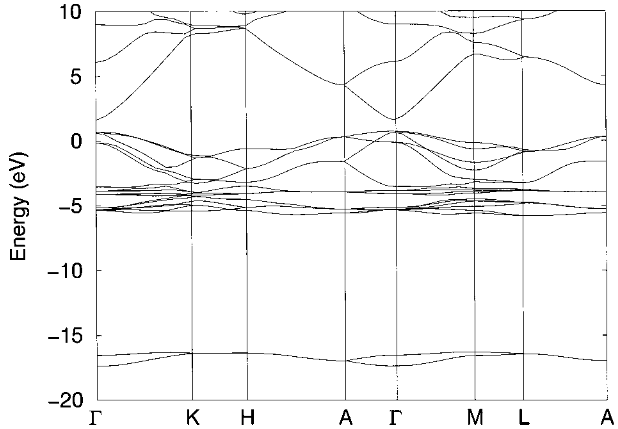

FIG. 2. Calculated pseudopotential band structure for a defect-free wurtzite crystal. The band gap is 0.91 eV.

placed within a 15 Å cubic cell (supercell approach). Only the Γ point was used. The computed bond length was 1.22 Å compared with an experimental value[59] of 1.21 Å.

The resulting formation energy for ZnO is −4.01 eV/ZnO. This value is comparable to the experimental −3.61 eV/ZnO enthalpy of formation;[60] the slight overbinding is typical of LDA calculations.

C. Defect formation energies

Using the formalism explained in Sec. II we computed the formation energies of all the native point defects in ZnO in their relevant charge states. These defects include oxygen and zinc vacancies, interstitials, and antisites. There are two possible interstitial sites in the wurtzite structure: one is tetrahedrally coordinated (tet) while the other is octahedrally coordinated (oct). A list of all the defects, and their charge states, considered in the present study is shown in Table II.

Based on the analysis in Sec. III A, we used a 3×3×2 wurtzite supercell in all defect calculations. Only those atoms that are present within a distance $r_{\mathrm{cut}}=3.95$ Å from the defect were relaxed. The volume of the cell was kept constant and corresponded to the computed lattice parameters shown in Table I.

The same pseudopotentials and cutoff energy employed in Sec. III B are used here. Because of the large size of the supercell, only one k point is necessary for k-space integrations. Selecting the Γ point, we find that the total energy is converged within 5 meV/ZnO for all defects. A different choice of special k point changes the energies, on average, by 5 meV/ZnO and the internal relaxations by less than 2.0% (relative to the amount of relaxation).

The formation energy of a charged defect is a function of the Fermi level [q is different from zero in Eq. (4)]. It is customary to assume this level to be zero at the top of the valence band. In this case, we assign the zero to the top of the *perfect-crystal* valence band in the *defect cell* (since we assume the electrons are placed or removed far away from the defect). Unfortunately, there is no absolute reference for the eigenvalues of different calculations.[61] The Fermi levels

TABLE II. Native-point defect formation energies (E_f) in ZnO. Values are given for the limiting chemical potentials indicated in Eq. (8) and for Fermi energies (ϵ_F) corresponding to the top and bottom of the valence and conduction bands, respectively. The different defects are identified by either their charge or the Kröger-Vink notation.

Defect	Charge on defect	Kröger-Vink notation	$E_f(\mu_{\mathrm{Zn}}=\mu^o_{\mathrm{Zn}})$ [eV] $\epsilon_F=0$ eV	$\epsilon_F=3.4$ eV	$E_f(\mu_{\mathrm{Zn}}=\mu^o_{\mathrm{Zn}}+\Delta E_f^{\mathrm{ZnO}})$ (eV) $\epsilon_F=0$ eV	$\epsilon_F=3.4$ eV
Zn_i(oct)	0	$\mathrm{Zn}_i^{\mathrm{x}}$	1.73	1.73	5.74	5.74
Zn_i(oct)	+1	$\mathrm{Zn}_i^{\cdot}$	1.28	4.68	5.29	8.69
Zn_i(oct)	+2	$\mathrm{Zn}_i^{\cdot\cdot}$	0.87	7.67	4.88	11.68
Zn_i(tet)	0	$\mathrm{Zn}_i^{\mathrm{x}}$	2.92	2.92	6.93	6.93
Zn_i(tet)	+1	$\mathrm{Zn}_i^{\cdot}$	2.61	6.01	6.62	10.02
Zn_i(tet)	+2	$\mathrm{Zn}_i^{\cdot\cdot}$	2.40	9.20	6.41	13.20
O_i(oct)	−2	O_i''	7.76	0.96	3.75	−3.046
O_i(oct)	−1	O_i'	6.81	3.41	2.80	−0.59
O_i(oct)	0	$\mathrm{O}_i^{\mathrm{x}}$	6.43	6.43	2.42	2.42
O_i(oct)	+1	$\mathrm{O}_i^{\cdot}$	6.40	9.80	2.39	5.79
O_i(tet)	−1	O_i'	7.49	4.09	3.48	0.08
O_i(tet)	0	$\mathrm{O}_i^{\mathrm{x}}$	6.50	6.50	2.49	2.49
O_i(tet)	+1	$\mathrm{O}_i^{\cdot}$	6.50	9.90	2.49	5.89
V_{Zn}	−2	$\mathrm{V}_{\mathrm{Zn}}''$	6.60	−0.20	2.59	−4.21
V_{Zn}	−1	$\mathrm{V}_{\mathrm{Zn}}'$	5.82	2.42	1.80	−1.59
V_{Zn}	0	$\mathrm{V}_{\mathrm{Zn}}^{\mathrm{x}}$	5.47	5.47	1.46	1.46
V_{O}	0	$\mathrm{V}_{\mathrm{O}}^{\mathrm{x}}$	0.02	0.02	4.02	4.02
V_{O}	+1	$\mathrm{V}_{\mathrm{O}}^{\cdot}$	0.15	3.55	4.16	7.56
V_{O}	+2	$\mathrm{V}_{\mathrm{O}}^{\cdot\cdot}$	−0.32	6.47	3.69	10.49
Zn_{O}	0	$\mathrm{Zn}_{\mathrm{O}}^{\mathrm{x}}$	2.41	2.41	10.43	10.43
Zn_{O}	+2	$\mathrm{Zn}_{\mathrm{O}}^{\cdot\cdot}$	0.55	7.35	8.56	15.63
O_{Zn}	−2	$\mathrm{O}_{\mathrm{Zn}}''$	11.98	5.18	3.97	−2.83
O_{Zn}	0	$\mathrm{O}_{\mathrm{Zn}}^{\mathrm{x}}$	9.74	9.74	1.72	1.72

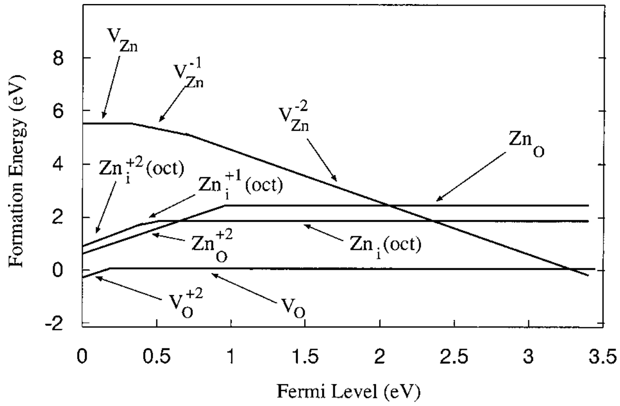

FIG. 3. Calculated defect formation energy for selected defects shown in Table II as a function of the Fermi level and for $\mu_{Zn}=\mu^0_{Zn}$ (high zinc partial pressure). Only the lowest formation-energy values are shown. The zero of the Fermi level is set to the top of the valence band.

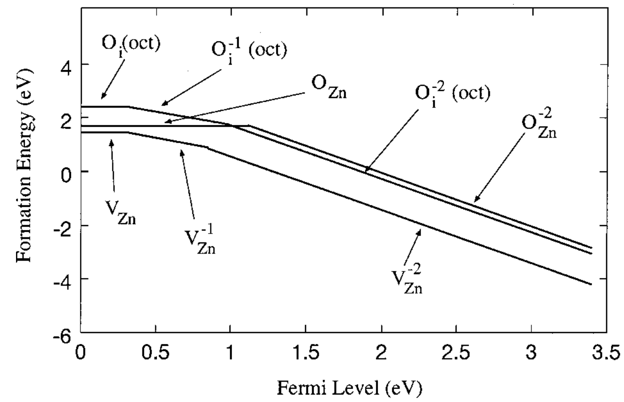

FIG. 4. Calculated defect formation energy for selected defects shown in Table II as a function of the Fermi level and for $\mu_{Zn}=\mu^0_{Zn}+\Delta E_f^{ZnO}$ (low zinc partial pressure). Only the lowest formation-energy values are shown. The zero of the Fermi level is set to the top of the valence band.

between different calculations cannot easily be compared. An additional procedure needs to be implemented in order to ''line up'' the bands of a perfect-crystal calculation with those of the defect supercells. To calculate this shift, we assumed that the potential in a perfect crystal is similar to the potential of the defect supercell *far* from the defect. We averaged these potentials along planes and plot their difference along the normal direction. Far from the defect, the difference between the potential in the perfect crystal and in the defect supercell is constant, and equal to the required shift.

In Table II, we show the defect formation energies at the limits of the zinc chemical potential and for Fermi levels corresponding to the top of the valence band and the bottom of the *experimental* conduction band. For the dominant defects, the formation energies are shown in Fig. 3 and Fig. 4 as a function of Fermi level (for the two limiting zinc chemical potential values). The slope of the lines in these figures corresponds to the charge state of the defect [see Eq. (4)]; for each defect, the line for a particular charge state has only been drawn over the range where this charge state has the lowest energy of all possible charge states. The kinks in the curves thus correspond to transitions between charge states (and hence to thermodynamic defect levels).

As will be discussed below in Sec. IV, the calculations suggest that the two most common defects are likely to be oxygen and zinc vacancies. We have therefore included a summary of the relaxed local bond lengths for the neutral states of these defect types in Table III (the corresponding bond lengths in the bulk ZnO are also included for comparison). The changes in bond lengths associated with the defect can be understood in terms of size and charge effects. A zinc defect involves removing an ion which is both fairly small and positively charged. It is therefore to be expected that other positively charged zinc will move closer to the vacant site, as there is more open space and reduced electrostatic repulsion. On the other hand, the oxygen neighbors are no longer electrostatically attracted to the vacancy and consequently move farther away. In general it is found that both zinc and oxygen move closer to an oxygen site when the oxygen is removed (except for a slight increase in the third nearest-neighbor zinc). This suggests that the size of the oxygen ion dominates in keeping the other ions away, and the whole region contracts when the oxygen vacany is created. For all the relaxations it can be seen that they are very small at distances near $r_{cut}=3.95$ Å, which further suggests that relaxations outside this cutoff are inessential.

IV. DISCUSSION

From Figs. 3 and 4 we deduce that, depending on the partial pressure of Zn, the most abundant native defects in ZnO are the oxygen and zinc vacancies. As we mentioned in

TABLE III. Distances to nearest-neighbor atoms for bulk and defected ZnO. d_i gives the distance to the ith nearest neighbor in angstroms. The values in parentheses are the number of neighbors at that distance. Distances outside r_{cut} are not included as they are identical in the bulk and defected structures.

Defect	Neighbor type	d_1	d_2	d_3	d_4	d_5
Zn (in bulk ZnO)	Zn	3.17(6)	3.19(6)			
V_{Zn}	Zn	3.12(3),3.13(3)	3.17(6)			
Zn (in bulk ZnO)	O	1.94(3)	1.95(1)	3.20(1)	3.74(3)	3.74(6)
V_{Zn}	O	2.11(3)	2.12(1)	3.21(1)	3.74(3)	3.75(6)
O (in bulk ZnO)	Zn	1.94(3)	1.95(1)	3.20(1)	3.74(3)	3.74(6)
V_O	Zn	1.79(3)	1.79(1)	3.21(1)	3.73(3)	3.74(6)
O (in bulk ZnO)	O	3.17(6)	3.19(6)			
V_O	O	3.12(3),3.13(3)	3.15(6)			

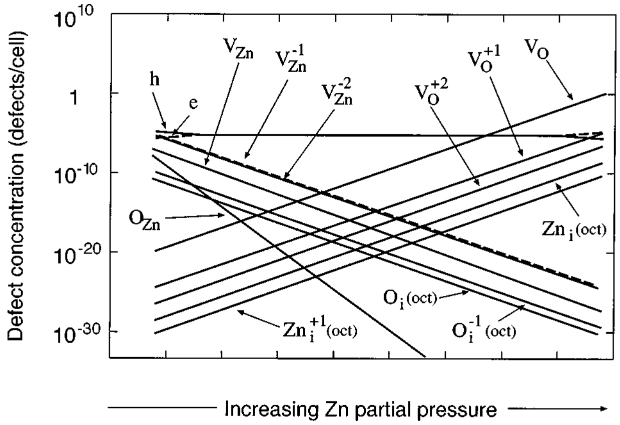

FIG. 5. Intrinsic defect concentration at 1000 K. The concentration was computed using Eq. (1) with the formation energy values from Figs 3 and 4, but using the *theoretical* band gap. The Fermi level was obtained by requiring overall charge neutrality. Conduction electrons are identified as ''e'' and holes in the valence band as ''h.''

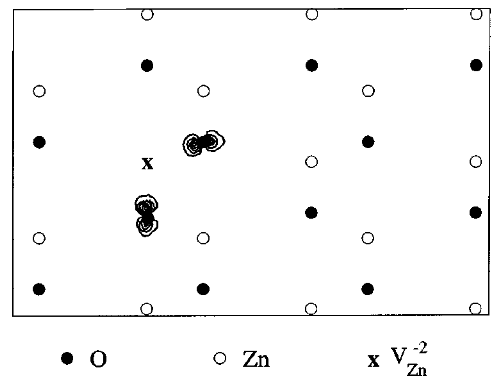

FIG. 6. Contour plot, in the $(1\bar{2}10)$ plane, of the electronic density corresponding to the V_{Zn}^{-2} level at the ZnO band gap. The ''x'' indicates the position of the vacancy. Contour lines are separated by 0.07 $e/Å^3$.

Sec. I, most of the literature agrees on the presence of Zn vacancies. On the other hand, there is no consensus on whether the other majority defect that is detected experimentally corresponds to oxygen vacancies or zinc interstitials. Our work indicates that oxygen vacancies have a lower formation energy than the zinc interstitial defects and hence should be more abundant. Mahan's[11] interpretation of Hagemarks[12] experimental measurements agrees well with our results. When there is zinc excess the native donors are the oxygen vacancies (see Fig. 3), and when there is Zn depletion Zn vacancies are present (see Fig. 4).

Experiments to distinguish between Zn interstitials and O vacancies are usually difficult to perform and interpret. These difficulties are due, in part, to the use of indirect evidence and to the sensitivity of the results to critical variables such as dopant concentration, processing, atmosphere, and temperature. Some of these effects can be analyzed using our results.

By changing the Fermi level we can simulate the effect of changes in dopant concentration. It is clear from Fig. 3 (Zn-rich conditions) that oxygen vacancies are lower in energy than Zn interstitials for all Fermi-level positions (the Zn_i formation energy value being at least 1.2 eV higher than that of V_O). Similarly, Fig. 4 (for oxygen-rich conditions) shows that zinc vacancies dominate over the whole range of Fermi-level values.

Changes in the environmental conditions should not change the conclusions either. In Fig. 5, we show the *intrinsic* defect concentration as a function of the Zn partial pressure. In this case the Fermi level is determined by requiring charge neutrality. Oxygen and zinc vacancies remain as the predominant defects in the chemical potential range specified by Eq. 8. Note that we are not considering the presence of defect complexes that can considerably change the present analysis. These complexes can be formed by the association between defects or dopants and defects.

Now we examine the sensitivity of our results to corrections of the LDA band gap. The LDA *underpredicts* the ZnO band gap by 2.5 eV. Consequently, we can expect the calculated defect levels in the gap to be lower than they would be with a correct band gap. Any correction we introduce will consistently shift the formation values upwards. A crude way to account for this problem consists in increasing the formation values by $2.5n_{elec}$, where n_{elec} corresponds to the occupation of the defect level. However, one expects levels that exhibit more valence-band character to be less affected than those that exhibit more conduction-band character. The top of the valence band is mainly formed by oxygen $2p$ levels while the bottom of the conduction band is mainly Zn $4s$ in character.[18,21] In Fig. 6, we show a contour plot, in a $(1\bar{2}10)$ plane in the wurzite structure, of the electronic density corresponding to the wave function of the highest occupied defect state for a V_{Zn}^{-2} defect. This wave function is dominated by oxygen p orbitals. This means that the defect wave function is largely made up out of orbitals that have valence-band character; the defect state is therefore unlikely to shift upwards when the band gap is ''opened up.'' Consequently Zn vacancies will continue to have the lowest formation energy, since the formation energies of other defects will either stay the same or shift upwards in Fig. 4. Our conclusions regarding the dominance of Zn vacancies in oxygen-rich material are therefore robust.

The situation is more complex for Zn-rich material, in which the LDA results show oxygen vacancies to have the lowest energies. The oxygen vacancy defect level is a combination of Zn s and O p orbitals, as shown in Fig. 7. Because the defect state exhibits some conduction-band character, its level may be shifted up when the band-gap correction is applied. The formation energy of V_O^{+2} will remain unchanged since no electrons occupy the level in this case. In p-type Zn-rich material, oxygen vacancies will therefore remain the dominant defect. The formation energy of V_O^0, however, may be affected, and shift to higher values. While it is in principle possible that the formation energy of V_O^0 would become larger than that of Zn_i, this seems unlikely since the Zn_i^0 formation energy will also be pushed to higher values. (An analysis of the defect electronic density for Zn_i reveals conduction-band character and consequently the level is likely to shift upwards.) However, as can be seen from Fig. 3, for extreme n-type conditions Zn vacancies may actually become dominant.

Irrespective of the LDA errors, we conclude from our calculations that the oxygen vacancy is a so-called

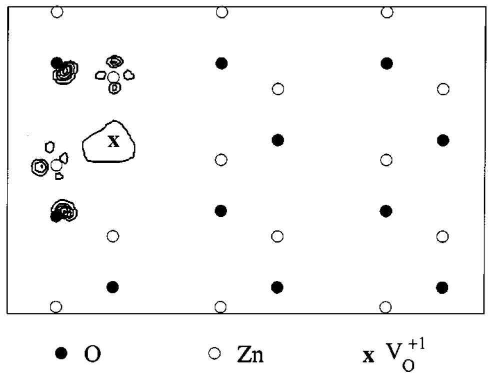

FIG. 7. Contour plot, in the $(1\bar{2}10)$ plane, of the electronic density corresponding to the V_O^{+1} level at the ZnO band gap. The ''x'' indicates the position of the vacancy. Contour lines are separated by 0.01 $e/Å^3$.

''negative-U'' defect; indeed, with increasing Fermi level a transition occurs directly from the $+2$ to the neutral charge state, with the $+1$ charge state always being higher in energy than the other two.

Because of limited computational resources, we did not sample all the relevant charge states for all defects. Only for zinc interstitials and zinc and oxygen vacancies were all charge states taken into account. For oxygen interstitials, we found that the more positive the defect charge is the less stable the defect is. Consequently, only the lowest charges were sampled. Also in the same way as with the zinc interstitials, oxygen interstitials were always more stable in the octahedral site (we computed a few tetrahedrally positioned defects to check that this was actually the case). Finally, for zinc and oxygen antisites, we tested a few cases and verified their formation energies were always higher than the ones for Zn and O vacancy defects. Based on these calculations we do not expect our conclusions to change if all possible charged states are considered.

As mentioned in the introduction, ZnO often exhibits green luminescence, centered between 2.4 and 2.5 eV.[34,62] Our results for defect levels suggest a possible mechanism for this emission. Defect levels correspond to transitions between charge states; these transitions in turn correspond to the kinks in the curves shown in Figs. 3 and 4. The transition level between the -1 and -2 charge states of V_{Zn} occurs around 0.8 eV above the valence band. We argued above that this level would be largely unaffected by the LDA band-gap error. A transition between the conduction band and the V_{Zn} acceptor level would thus give rise to luminescence around 2.6 eV, in reasonable agreement with the observed transition energy. In addition to the agreement with the observed transition 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 material exhibits high n-type conductivity. This proposed explanation for the green luminescence is similar to the proposal that gallium vacancies are the source of the yellow luminescence in GaN.[63]

Our calculations indicate that oxygen vacancies have a $+2/0$ transition at an energy within the gap. As discussed above, the position of this level is rather uncertain due to the LDA band gap error; the fact that the level occurs within the band gap does indicate, however, that oxygen vacancies can also give rise to luminescence in ZnO.

V. CONCLUSIONS

The characterization of native point defects in ZnO has important technological implications. In this work, we applied a first principles pseudopotential approach to this problem.

We found the most abundant native defects to be Zn and O vacancies depending on the Zn partial pressure. The validity of our conclusions was carefully checked with regard to the LDA band-gap error, effects of the supercell approximation, and convergence of pseudopotential method parameters. Our results also suggest that the green luminescence observed in ZnO may originate from transitions between electrons in the conduction band and zinc vacancy levels.

We did not analyze here the effects of the association between different defects. They certainly play an important role in other systems and we cannot discard *a priori* their influence on the ZnO properties. Estimating this effect will be the objective of future work.

ACKNOWLEDGMENTS

This work was supported by the Department of Energy, Office of Basic Energy Sciences under Contract No. DE-FG02-96ER45571. Dr. J. Hafner and Dr. G. Kresse are gratefully acknowledged for providing us with the VASP pseudopotential code. We thank the San Diego Supercomputer Center for the use of their computing facilities.

[1] T. K. Gupta, J. Am. Ceram. Soc. **73**, 1817 (1990).
[2] K. Vanheusden, W. L. Warren, C. H. Seager, D. R. Tallant, J. A. Voight, and B. E. Gnade, J. Appl. Phys. **79**, 7983 (1996).
[3] F.-C. Lin, Y. Takao, Y. Shimizu, and M. Egashira, J. Am. Ceram. Soc. **78**, 2301 (1995).
[4] K. L. Chopra and S. R. Das, *Thin Film Solar Cells* (Plenum, New York, 1983).
[5] K. L. Chopra, S. Major, and D. K. Pandya, Thin Solid Films **102**, 1 (1983).
[6] F. Hamdani, A. E. Botchkarev, H. Tang, W. Kim, and H. Morkoç, Appl. Phys. Lett. **71**, 3111 (1997).
[7] T. Detchprohm, K. Hiramatsu, H. Amano, and I. Akasaki, Appl. Phys. Lett. **61**, 2688 (1992).
[8] D. C. Reynolds, D. C. Look, and B. Jogai, Solid State Commun. **99**, 873 (1996).
[9] D. M. Bagnall, Y. F. Chen, Z. Zhu, T. Yao, S. Koyama, M. Y. Shen, and T. Goto, Appl. Phys. Lett. **70**, 2230 (1997).
[10] M. H. Sukkar and H. L. Tuller, in *Advances in Ceramics*, edited by M. F. Yan and A. H. Heuer (American Ceramic Society, Columbus, 1982), Vol. 7, p. 71.
[11] G. D. Mahan, J. Appl. Phys. **54**, 3825 (1983).
[12] K. I. Hagemark, J. Solid State Chem. **16**, 293 (1976).

[13] J. S. Choi and C. H. Yo, J. Phys. Chem. Solids **37**, 1149 (1976).
[14] A. W. Sleight and R. Wang, in *Solid-State Chemistry of Inorganic Materials*, edited by P. K. Davies, A. J. Jacobson, C. C. Torardi, and T. A. Vanderah MRS Symposia Proceedings No. 453 (Materials Research Society, Pittsburgh, 1997), pp. 323–330.
[15] J. W. Hoffman and I. Lauder, Trans. Faraday Soc. **66**, 2346 (1970).
[16] E. Ziegler, A. Heinrich, H. Oppermann, and G. Stover, Phys. Status Solidi A **66**, 635 (1981).
[17] T. Kamiya, J. Phys. Chem. Solids **44**, 365 (1983).
[18] J. E. Jaffe and A. C. Hess, Phys. Rev. B **48**, 7903 (1993).
[19] C. K. Yang and K. S. Dy, Solid State Commun. **88**, 491 (1993).
[20] Y. N. Xu and W. Y. Ching, Phys. Rev. B **48**, 4335 (1993).
[21] P. Schröer, P. Krüger, and J. Pollmann, Phys. Rev. B **47**, 6971 (1993).
[22] A. Dal Corso, M. Posternak, R. Resta, and A. Baldereschi, Phys. Rev. B **50**, 10 715 (1994).
[23] D. Vogel, P. Krüger, and J. Pollmann, Phys. Rev. B **52**, 14 316 (1995).
[24] S. Massidda, R. Resta, M. Posternak, and A. Baldereschi, Phys. Rev. B **52**, 16 977 (1995).
[25] M. H. Sukkar, K. H. Johnson, and H. L. Tuller, Mater. Sci. Eng., B **6**, 49 (1990).
[26] W. C. Mackrodt, R. F. Stewart, J. C. Campbell, and I. H. Hillier, J. Phys. (Paris), Colloq. **41**, 64 (1980).
[27] F. A. Kröger and H. J. Vink, J. Chem. Phys. **22**, 250 (1954).
[28] P. H. Kasai, Phys. Rev. **130**, 989 (1963).
[29] K. Vanheusden, C. H. Seager, W. L. Warren, D. R. Tallant, and J. A. Voigt, Appl. Phys. Lett. **68**, 403 (1996).
[30] H. J. Egelhaaf and D. Oelkrug, J. Cryst. Growth **161**, 190 (1996).
[31] M. Liu, A. H. Kitai, and P. Mascher, J. Lumin. **54**, 35 (1992).
[32] V. F. I and G. A. Konivalow, Zh. Eksp. Teor. Fiz. **23**, 712 (1952).
[33] E. G. Bylander, J. Appl. Phys. **49**, 1188 (1978).
[34] D. C. Reynolds, D. C. Look, B. Jogai, and H. Morkoç, Solid State Commun. **101**, 643 (1997).
[35] R. Dingle, Phys. Rev. Lett. **23**, 579 (1969).
[36] W. Lehman, J. Electrochem. Soc. **115**, 538 (1968).
[37] D. B. Laks, C. G. Van de Walle, G. F. Neumark, P. E. Blochl, and S. T. Pantelides, Phys. Rev. B **45**, 10 965 (1992).
[38] C. G. Van de Walle, D. B. Laks, G. F. Neumark, and S. T. Pantelides, Phys. Rev. B **47**, 9425 (1993).
[39] G. A. Baraff and M. Schlüter, Phys. Rev. B **30**, 1853 (1984).
[40] S. T. Pantelides, Rev. Mod. Phys. **50**, 797 (1978).
[41] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[42] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[43] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).
[44] A. F. Kohan and G. Ceder, Comput. Mater. Sci. **8**, 142 (1997).
[45] D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).
[46] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).
[47] M. C. Payne, M. P. Jeter, D. C. Allan, T. A. Arras, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[48] F. Gygi, Europhys. Lett. **19**, 617 (1992).
[49] A. F. Kohan and G. Ceder, Phys. Rev. B **54**, 805 (1996).
[50] K. Cho, T. A. Arias, J. D. Joannopoulos, and P. K. Lam, Phys. Rev. Lett. **71**, 1808 (1993).
[51] H. T. Stokes, L. L. Boyer, and M. J. Mehl, Phys. Rev. B **54**, 7729 (1996).
[52] I. M. Torrens, *Interatomic Potentials* (Academic, New York, 1972).
[53] G. V. Lewis and C. R. A. Catlow, J. Phys. C **18**, 1149 (1985).
[54] W. D. Kingery, H. K. Bowen, and D. R. Uhlmann, *Introduction to Ceramics* (Wiley, New York, 1975).
[55] S. G. Louie, S. Froyen, and M. L. Cohen, Phys. Rev. B **26**, 1738 (1982).
[56] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[57] O. Madelung, M. Schulz, and H. Weiss, *Numerical Data and Functional Relationships in Science and Technology* (Springer-Verlag, Berlin, 1982), Vol. 17.
[58] ASM International Handbook Committee, *Metals Handbook*, 10th ed. (ASM, Materials Park, OH, 1990), Vol. 2.
[59] D. R. Lide, *Handbook of Chemistry and Physics*, 78 ed. (CRC, Boca Raton, 1997).
[60] J. A. Dean, *Lange's Handbook of Chemistry*, 14th ed. (McGraw-Hill, New York, 1992).
[61] L. Kleinman, Phys. Rev. B **24**, 7412 (1981).
[62] R. B. Lauer, J. Phys. Chem. Solids **34**, 249 (1973).
[63] J. Neugebauer and Chris G. Van de Walle, Appl. Phys. Lett. **69**, 503 (1996).