Convergence of the formation energies of intrinsic point defects in wurtzite ZnO: first-principles study by projector augmented wave method

2006 J. Phys.: Condens. Matter 18 1495

(http://iopscience.iop.org/0953-8984/18/5/002)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **18** (2006) 1495–1508 doi:10.1088/0953-8984/18/5/002

Convergence of the formation energies of intrinsic point defects in wurtzite ZnO: first-principles study by projector augmented wave method

Jun-Liang Zhao[1,2], **Wenqing Zhang**[1], **Xiao-Min Li**[1], **Ji-Wei Feng**[1,2] **and Xun Shi**[1,2]

[1] State Key Laboratory of High Performance Ceramics and Superfine Microstructures, Shanghai Institute of Ceramics, Chinese Academy of Sciences, Shanghai 200050, People's Republic of China
[2] Graduate School of Chinese Academy of Sciences, Beijing 100039, People's Republic of China

Received 27 September 2005, in final form 25 November 2005
Published 17 January 2006
Online at stacks.iop.org/JPhysCM/18/1495

Abstract

Electronic structure, formation energies, transition levels, and concentration of intrinsic defects in wurtzite ZnO are investigated by the projector augmented wave method in the generalized gradient approximation. Interstitials, vacancies, and antisites at different charge states are considered. Convergence of the formation energies of various intrinsic point defects is carefully checked, and comparison with earlier results is made and discussed. Even though there exists a difference for the calculated formation energies of certain defects, our calculations also show that oxygen and zinc vacancies are the dominant intrinsic donor and acceptor defects in ZnO, indicating a consistency among results by different methods. The oxygen vacancy is not expected to be the main source of strong n-type conductivity in the unintentionally doped ZnO, due to its deep level in the bandgap, but it must be the origin of the experimentally observed visible photoluminescence band centred between 2.3 and 2.5 eV.

1. Introduction

Zinc oxide (ZnO) has long been receiving much attention for its applications in varistors, transparent high power electronics, surface acoustic wave devices, piezoelectric transducers, and gas-sensing devices [1, 2]. In recent years, ZnO has also acted as a novel wide-bandgap semiconductor material for short-wavelength optoelectronic devices, such as UV lasers, blue to UV light emitting diodes and UV detectors [3–6]. The notable properties of ZnO are its direct bandgap of 3.37 eV at room temperature (RT) and high exciton binding energy of 60 meV, much higher than the values of some widely used wide-bandgap materials, such as ZnSe (20 meV) and GaN (21 meV) [7].

ZnO crystal grown by various techniques always presents highly n-type conductivity. Although the intrinsic n-type conduction is very useful for some applications, it is more

appreciable to have better control over the conduction type for ZnO. However, it has been proved that p-type doping is difficult to realize in ZnO [8–10], leading to a bottleneck in development of ZnO based devices. To overcome the doping bottleneck, it is essential to achieve a fundamental understanding of the physics of various defects in ZnO. In spite of numerous studies on defects there is still controversy about the intrinsic defects in ZnO. Experimental studies on the intrinsic defects in ZnO have long been focused on the choice of predominant defect from zinc interstitial and oxygen vacancy [11–16]. In addition, the origin of the green photoluminescence band centred at about 2.3–2.5 eV [7, 17–20] for ZnO film is not understood yet. This green emission is relatively strong for ZnO film grown by some vacuum techniques, but very weak in ZnO film grown in the atmosphere [11–22]. Experimentally, it is speculated that the origin is oxygen vacancies [17, 20], zinc vacancies [23], zinc interstitials [19, 24], zinc antisites [25], or oxygen antisites [7]. Theoretical study on the intrinsic defects is also helpful to resolve these issues.

In recent years, theoretical investigations based on first principles calculations have been applied to study the defect physics of ZnO. Sun studied the intrinsic defects in ZnO, including the oxygen vacancy, zinc vacancy, and zinc interstitial, by the full potential linear muffin-tin orbital method, and concluded that the zinc interstitial is the dominant donor according to its shallower defect level [26]. Kohan and Van de Walle [23, 27] studied different possible native point defects in ZnO by the plane-wave pseudopotential method in the local density approximation (LDA). Their results indicated that the oxygen vacancy and zinc vacancies are the dominant donor and acceptor species, respectively. They also concluded that all native donors have deep transition levels and that n-type conductivity in undoped ZnO should not be attributed to native defects. Zhang *et al* [28] also calculated the formation energies of native defects in ZnO by the plane-wave pseudopotential method in the LDA. Oxygen and zinc vacancies also have the lowest formation energies in their calculations, but the formation energies of most defects have a considerable difference from Kohan's results. Oba *et al* [29] calculated the electronic structure of native defects in ZnO by the plane-wave pseudopotential method in the generalized gradient approximation (GGA), and showed that oxygen vacancies and zinc vacancies act as dominant donor and acceptor defects with deep levels, in agreement with Kohan's and Zhang's LDA studies, but they concluded that zinc interstitials and zinc antisites are shallow donors, which may be employed to explain the n-type conductivity of undoped ZnO.

Even with so many calculations as discussed above, there is still significant discrepancy between the data obtained by different groups. For example, the formation energies of the neutral zinc antisite (Zn_O) calculated by various groups [23, 28, 29] show a difference of 2–3 eV, while the formation energies of the charged Zn_O^{2+} defect exhibit an even larger difference of 3–5 eV. It is also noticed that there is no systematic check of the effect of different calculation parameters on defect formation energies. Kohan *et al* [23] reported a check for the convergence of defect formation energies with supercell size. They calculated the difference in formation energies of Zn_i at tetrahedral and octahedral sites at various supercell sizes, using an empirical pair-potential calculation approach, and found a decrease of $\sim$1 eV for the difference when the supercell is increased from $3\times3\times2$ to $4\times4\times4$. They suggested an internal relaxation approach to ensure the relative formation energies converge at a smaller supercell. However, the internal relaxation is in contradiction with the defect-induced structure optimization, in which all atoms should move to their energy favourable positions. Moreover, they only studied the convergence of formation energies of neutral point defects and did not consider that for charged defects.

It is well known that the generalized gradient approximation for the exchange–correlation potential offers more accuracy and reliability for total energy than the LDA [30]. In addition, the projector augmented wave (PAW) method is also believed to be more reliable than those

methods based on pseudopotentials [30–32]. It is our purpose to study the intrinsic defects at all possible charge states using PAW methods in GGA. The convergence of formation energies of neutral and charged defects has been systematically checked in our calculations.

2. Theoretical approach

2.1. Details of total-energy calculations

Calculations on the total energies and electronic structure of defects in ZnO are based on the density-functional theory, using the PAW method as implemented in the Vienna *Ab Initio* Simulation Package (VASP) [33]. The GGA is employed for the exchange–correlation (XC) potential [34]. The Zn 3d electrons were treated as part of the valence band. The cut-off energy for the plane-wave basis is 400 eV. A 72-atom supercell ($3 \times 3 \times 2$) is used for the defect calculations, with four-k-point sampling in the irreducible Brillouin zone of the supercell. A 108-atom supercell ($3 \times 3 \times 3$) and a 256-atom supercell ($4 \times 4 \times 4$) are also employed for a few cases to check the convergence of the calculated defect formation energies. The projector augmented wave (PAW) method [31] is applied for all the self-consistent electronic structure calculations, with structure optimization based on the calculated forces on each atom. In all calculations, all atoms were allowed to relax until the Hellmann–Feynman forces acting on them became less than 0.01 eV Å^{-1}. A jellium background charge was used for charged systems.

2.2. Defect formation energy

For the ZnO system in equilibrium with a reservoir of Zn and O, the formation energy of a point defect D at charge state q is defined as [35, 36]

$$E_{\mathrm{f}}(D,q) = E_{\mathrm{tot}}(D,q) - E_{\mathrm{tot}}(\mathrm{ZnO,\ bulk}) - \sum_i n_i \mu_i + q(\varepsilon_{\mathrm{F}} + E_{\mathrm{v}} + \Delta V), \tag{1}$$

where $E_{\mathrm{tot}}(D,q)$ is the total energy of the supercell with one defect D, and E_{tot} (ZnO, bulk) is the total energy for an equivalent supercell containing only bulk ZnO. n_i indicates the number of atoms of type i (host atoms or impurity atoms) that have been added to ($n_i > 0$) or removed from ($n_i < 0$) the supercell when the defect is created, and μ_i are the corresponding chemical potentials of these species, which represent the energy of the reservoir with which atoms are being exchanged. ε_{F} is the Fermi level, referenced to the valence-band maximum in the bulk, E_{v}. ΔV is a correction term to align the reference potential in our defect supercell with that in the perfect cell.

The *ab initio* effective chemical potentials for Zn and O are defined as

$$\Delta\mu_{\mathrm{Zn}} = \mu_{\mathrm{Zn}} - \mu^0_{\mathrm{Zn}}, \qquad \Delta\mu_{\mathrm{O}} = \mu_{\mathrm{O}} - \mu^0_{\mathrm{O}}, \tag{2}$$

with the boundaries as

$$\Delta E_{\mathrm{f}}^{\mathrm{ZnO}} < \Delta\mu_{\mathrm{Zn}} < 0, \qquad \Delta E_{\mathrm{f}}^{\mathrm{ZnO}} < \Delta\mu_{\mathrm{O}} < 0, \tag{3}$$

where μ^0_{Zn} is the chemical potential of metallic Zn, and μ^0_{O} is the chemical potential of molecular oxygen (per oxygen atom). $\Delta E_{\mathrm{f}}^{\mathrm{ZnO}}$ is the formation enthalpy of ZnO.

2.3. Defect and carrier concentration

At thermodynamic equilibrium, the concentration of a defect in a crystal is given by the expression:

$$C_{\mathrm{d}} = N_{\mathrm{sites}} N_{\mathrm{config}} \mathrm{e}^{-E_{\mathrm{f}}/k_{\mathrm{B}}T}. \tag{4}$$

Table 1. Computed and experimentally measured lattice parameters for a defect-free ZnO wurtzite structure.

	a (Å)	c (Å)	u (internal)
Calculations	3.287	5.279	0.381
Experiments [16]	3.249	5.205	0.382

Here, E_f is the formation energy of the defect as defined in equation (1). N_{sites} is the number of sites in the lattice (per unit volume) where the defect can be incorporated; N_{config} is the number of equivalent configurations. For vacancies, antisites, and substitutional defects considered in this work, $N_{config} = 1$ due to no symmetry breaking occurrence.

For a perfect ZnO crystal, the carrier concentration is

$$n_e = n_h = (N_C N_V)^{1/2} \exp(-E_g/2kT), \tag{5}$$

where n_e and n_h are the concentration of free electrons and holes respectively. k is the Boltzmann constant and T the temperature. N_C (N_V) is the effective density of states of the conduction (valence) band.

For a ZnO crystal with defects, the carrier concentration depends upon the Fermi level in the following way:

$$n_e = N_C \exp[-(E_g - \varepsilon_F)/kT], \qquad n_h = N_V \exp(-\varepsilon_F/kT). \tag{6}$$

Here, the Fermi level ε_F is determined by requiring overall charge neutrality [37], i.e.

$$\sum_i q(i) N_s(i) \exp[-E_f^{(q)}(i)/kT] = N_C \exp[-(E_g - \varepsilon_F)/kT] - N_V \exp(-\varepsilon_F/kT). \tag{7}$$

Here, $E_f^{(q)}(i)$ is the formation energy of defect i at charge state q and $N_s(i)$ is the number of sites where defect i can be formed per unit volume.

3. Results and discussions

3.1. Perfect-crystal ZnO

The computed lattice constants are shown in table 1. According to the computed band structure for perfect ZnO, the predicted bandgap is 0.8 eV, far below the experimental value of 3.37 eV. The underestimation for the bandgap [30] is due to the problem of GGA itself, and was also found in other calculations based on LDA (0.91 eV [23], 0.6 eV [28]) and GGA using the ultrasoft pseudopotential method (0.96 eV) [29].

In ZnO bulk, the defect properties depend on the chemical potential of the species involved as shown in equation (1). To find the limiting value for the chemical potentials we need the energy of metallic Zn and molecular oxygen. Therefore, we compute the total energies for these two systems. The supercell approach used here is similar to that in [23], and the resulting formation energy for ZnO is −3.44 eV/ZnO, in good agreement with the corresponding experimental value −3.61 eV/ZnO [38]. The formation energy of ZnO calculated by the LDA method is −4.01 eV [23], less accurate than our result due to the overbinding error of LDA [30].

3.2. Convergence of defect formation energies by supercell approach

To check the effect of the size of the supercell, we calculated the defect formation energies at various supercell size, as shown in figure 1. Only neutral states are considered, and charged

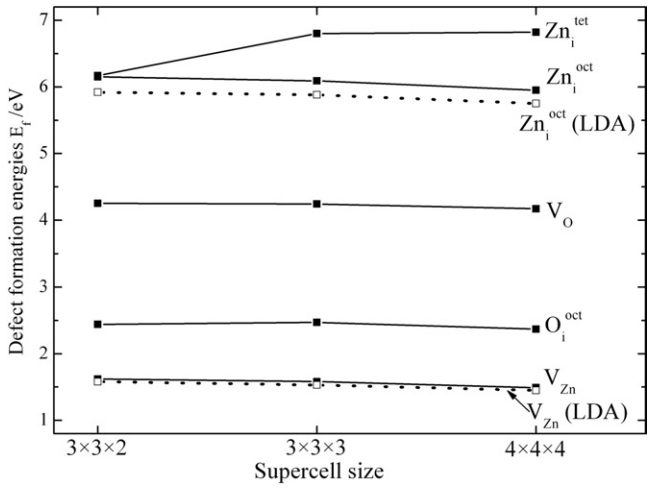

Figure 1. Dependence of the formation energies of various defects on supercell size at high oxygen chemical potential ($\Delta\mu_O = 0$). The solid lines show the data calculated by the PAW method in the GGA and the dotted lines indicate the data obtained by the PAW method in the LDA.

defects will be discussed later. Formation energies of vacancies always converge at relatively small supercells, due to little lattice relaxation induced by vacancies. It is clear that the formation energies of oxygen and zinc vacancies have small changes (∼0.1 eV) when the supercell size is increased from $3 \times 3 \times 2$ to $4 \times 4 \times 4$. Interstitials have two possible sites in wurtzite structure: tetrahedrally coordinated (tet) and octahedrally coordinated (oct). With the supercell size increased from $3 \times 3 \times 2$ to $4 \times 4 \times 4$, the formation energy for the zinc interstitial at the octahedral site changes from 6.15 to 5.95 eV, while a change up to 0.87 eV is found for the tetrahedrally coordinated zinc interstitial. Thus a reasonably reliable defect formation energy can be obtained for Zn_i at the octahedral site in the $3 \times 3 \times 2$ supercell, while that for Zn_i at the tetrahedral site needs at least a $3 \times 3 \times 3$ supercell to converge. The main reason is due to the larger relaxation induced by the tetrahedral Zn_i than that by the octahedral Zn_i. We found that the maximum change of Zn–O bond length is 1.2 Å induced by Zn_i at the tetrahedral site in the $3 \times 3 \times 2$ supercell, but only 0.4 Å induced by the Zn_i at octahedral site, which leads to the large change of the formation energy difference for Zn_i at tetrahedral and octahedral sites with increasing supercell size. This is mainly attributed to the unreliable formation energy of Zn_i at the tetrahedral site in the $3 \times 3 \times 2$ supercell (as shown in figure 1). Note that the Zn_i at the octahedral site always shows a lower formation energy than that at the tetrahedral site for both $3 \times 3 \times 3$ and $4 \times 4 \times 4$ supercells, indicating that the zinc interstitial prefers to be stable at the octahedral site. Thus we only need to consider the octahedrally coordinated zinc interstitial, which shows convergent formation energy in the $3 \times 3 \times 2$ supercell in the full relaxation approach. The oxygen interstitial is also shown to be more stable at the octahedral site. The formation energy of the oxygen interstitial at the octahedral site changes only from 2.44 to 2.37 eV with the supercell size increasing from $3 \times 3 \times 2$ to $4 \times 4 \times 4$, indicating reliable formation energy for the oxygen interstitial obtained from the $3 \times 3 \times 2$ supercell calculation. We also checked some other defects by the PAW method in the LDA [39], as shown in figure 1. Although LDA formation energies are a little different from GGA results, the formation energies also converge at the supercell size of $3 \times 3 \times 2$ in LDA calculation. Therefore, a $3 \times 3 \times 2$ supercell is used for defect formation energy calculations. Therefore, the point defects have a minimum separation of 9.9 Å. The estimated uncertainties for defect

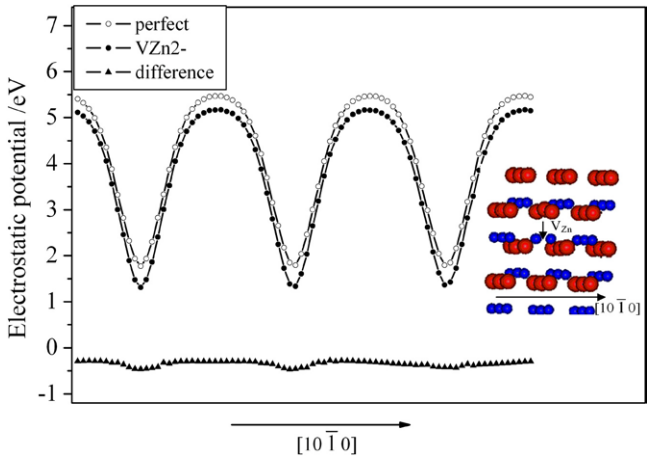

Figure 2. Average electrostatic potential along the $[10\bar{1}0]$ direction for perfect and V_{Zn}^{2-} defect cells. The difference between the defect cell and the perfect cell is also plotted, which represents the difference of their valence band maxima. Inset is the illustration of the V_{Zn}^{2-} defect cell for the electrostatic potential calculation. The larger (red) solid circles represent oxygen atoms while the smaller (blue) circles represent zinc atoms.

(This figure is in colour only in the electronic version)

formation energies in the $3 \times 3 \times 2$ supercell are expected to be *lower than 0.1 eV for most of the defects and lower than 0.3 eV for* Zn_i *at the interstitial sites* in our calculations.

3.3. Defects at charged states

The formation energy of a charged defect is a function of the Fermi level, which is customarily referred to the top of the valence band. For a defect cell, we assign the energy zero to the top of the valence band of the perfect ZnO, resulting in an energy shift, or the correction term ΔV in equation (1). To calculate this shift, an additional procedure needs to be implemented to 'line up' the top of the valence band of defect cells with that of perfect ZnO. We computed the average electrostatic potentials for both doped and pure ZnO and extract the difference between them at points far from the defect and outside the pseudopotential areas of atoms. Figure 2 shows the average electrostatic potentials at such points along the $[10\bar{1}0]$ direction for pure ZnO and the doped system with a V_{Zn}^{2-} defect. No obvious fluctuation is observed for the potential difference. The average of the difference along the line is taken as the correction term ΔV for V_{Zn}^{2-}. We calculated ΔV for all possibly charged defects using the aforementioned method and the results for some representative defects are shown in figure 3. It is noted that the ΔV term is comparable to the formation energies for some defects, indicating that the correction term is necessarily taken into account in the expressions for formation energies to obtain reliable results.

Another issue related to the charged system is the use of a uniform, compensating jellium background charge to circumvent the Coulomb divergence of the supercell approximation. Due to the electrostatic interactions between the background charge and the localized charge, the Coulomb energy of the supercell will converge slowly as a function of supercell size [35, 40]. Makov and Payne [41] proposed to correct for the Coulomb energy error in the case of jellium compensation, and derived a total-energy correction formula for charge distributions in cubic supercells. Several different Coulomb correction methods for the supercell approximation have

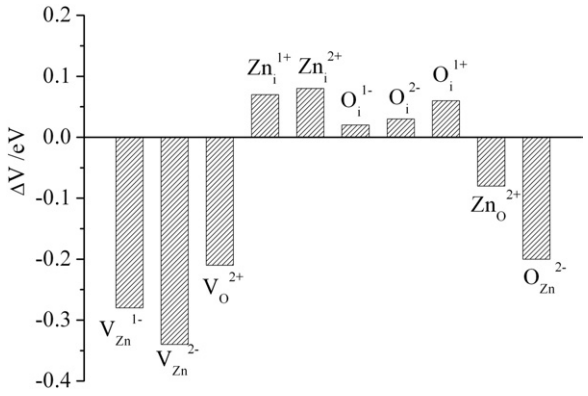

Figure 3. The difference of average electrostatic potential (ΔV) between some charged defect cells and the perfect cell, calculated by the method illustrated in figure 2.

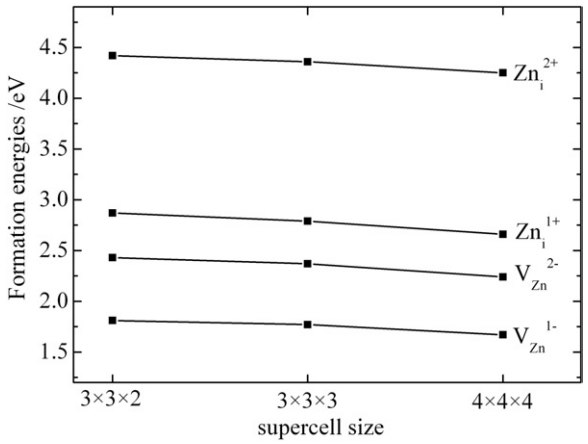

Figure 4. Dependence of the formation energies of various charged defects on supercell size at high oxygen chemical potential ($\Delta\mu_O = 0$).

been presented [40, 42, 43]. However, most of these approaches work well for ionic materials, while they lead to an overestimate of the correction term for defects in semiconductors [35]. Since more work is clearly needed to better understand these corrections, we do not employ any correction for jellium compensation and our results reported below do not include them. To reduce the effect of the jellium compensation error, we employ a enough large supercell, in which the interaction between background and localized charge can be negligible.

Now we discuss the convergence of formation energies for charged defects with the supercell size. We calculate the formation energies for some charged defects (V_{Zn}^{1-}, V_{Zn}^{2-}, Zn_i^{1+}, and Zn_i^{1+}) at various supercell sizes and the results are shown in figure 4. The difference between $4 \times 4 \times 4$ and $3 \times 3 \times 3$ cells is larger than the difference between $3 \times 3 \times 3$ and $3 \times 3 \times 2$ cells, which results from the latter difference coming from the increase of the cell size in one direction and the former coming from the increase in all three dimensions. The maximum error bar for the formation energies of charged defects in the $3 \times 3 \times 2$ supercell could be estimated to be about 0.3 eV in our calculations, which is small enough to not change the main conclusions of the remaining sections. Therefore, convergent formation energies can

be obtained for charged defects in a $3 \times 3 \times 2$ supercell, indicating that the 9.9 Å separation is sufficient for screening out the electrostatic interaction in charged supercells using jellium compensation.

Besides supercell size, k-point summation can also influence the convergence of defect formation energies. For a defect in the supercell approach, the interaction between defects in neighbouring supercells leads to dispersion of the defect level in the bandgap [35]. Thus certain charged states correspond to half-filled defect bands. When the defect level is partially occupied, the total energy is always a poor approximation to the value expected for an isolated defect, due to the unequal occupation of the defect level. Averaging over the defect band by special k-points produces the best approximation to the defect level of the isolated defect. We employed the Monkhorst–Pack scheme [44] to produce a regularly spaced mesh of $2 \times 2 \times 2$ points (four irreducible points) in the reciprocal unit cell. Since a large 72-atom supercell is used in our calculations, the defect–defect interactions are relatively weak, leading to limited influence of dispersion of the defect level on the convergence of the total energy. Therefore, it is believed that the $2 \times 2 \times 2$ k point mesh can produce converged defect formation energies. Increasing k-point sampling changes the energy by less than 0.01 eV.

3.4. Defect formation energies

Using equation (1) we computed the formation energies of all the intrinsic defect point defects, including oxygen and zinc vacancies, interstitials, and antisites, in a $3 \times 3 \times 2$ supercell. There are two possible interstitial sites in the wurtzite structure: tetrahedrally (tet) and octahedrally (oct) coordinated. As we discussed above, interstitials at octahedral sites are more stable than those at tetrahedral sites. Therefore, only the octahedrally coordinated zinc and oxygen interstitials are considered. All possible charge states for these defects are considered. The results are shown in table 2. The defect formation energies obtained by previous calculations [23, 28, 29] are also listed in table 2 for comparison. As discussed above, there is a significant difference between defect formation energies obtained by various previous calculations. The difference in the formation energies calculated by different methods is expected due to differences in computational details, especially the functionals for exchange–correlation potentials, pseudopotential types, k-point sampling, and relaxation approaches. Based on the aforementioned systematic check for the defect formation energies, we believe that our calculations using the full relaxation approach give relatively reliable results in the numerical values of defect formation energies.

The formation energies of dominant defects are shown in figure 5 as a function of Fermi level at two limiting oxygen chemical potentials, i.e. $\Delta\mu_O = -3.44$ eV (Zn-rich conditions) and $\Delta\mu_O = 0$ (oxygen-rich conditions), respectively. The kinks in the curves correspond to transitions between defects at different charge states, which can be applied to estimate thermodynamic defect levels [35]. In the oxygen-deficient condition, the most abundant native defect in ZnO is the oxygen vacancy, which acts as the dominant intrinsic donor. In the oxygen-rich condition, the zinc vacancy is the most abundant native defect and serves as the dominant intrinsic acceptor. It is also confirmed that the oxygen vacancy is a negative-U defect because the thermodynamic defect level for the transition from charge state $2+$ to the neutral state is higher than that for the transition from $1+$ to the neutral state. Note that this is in agreement with previous studies [23, 28, 29]. The zinc interstitial is also a candidate as the dominant intrinsic donor in some studies [14, 15]. However, our calculations indicate that the zinc interstitial is at least 1.0 eV higher in formation energy than the oxygen vacancy, as shown in figure 5(a). Meanwhile, the formation energy of the oxygen interstitial is about 1 eV higher than that of the zinc vacancy. The oxygen interstitial has a threefold-degenerated level

Table 2. Formation energies (E_f) of intrinsic point defects in the ZnO crystal at various charge states. Values are given at limiting chemical potentials and for Fermi energies (ε_f) corresponding to the top and bottom of the valence and conduction bands, respectively. Zinc and oxygen interstitials are all octahedrally coordinated. The values in parenthesis are the calculated data from [23, 28, 29], respectively.

		E_f ($\Delta\mu_O = 0$, $\Delta\mu_{Zn} = -3.44$)		E_f ($\Delta\mu_O = -3.44$, $\Delta\mu_{Zn} = 0$)	
Defect	Charge on defect	$\varepsilon_f = 0$ eV	$\varepsilon_f = 3.37$ eV	$\varepsilon_f = 0$ eV	$\varepsilon_f = 3.37$ eV
V_O	0	4.25 (4.02[a], 4.6[b], 3.2[c])	4.25	0.81	0.81
V_O	+1	3.60 (4.16[a], 3.9[b])	6.97	0.16	3.53
V_O	+2	1.97 (3.69[a], 2.6[b], 2.2[c])	8.71	−1.47	5.27
V_{Zn}	0	1.62 (1.46[a], 2.7[b], 2.0[c])	1.62	5.06	5.06
V_{Zn}	−1	1.81 (1.80[a], 2.6[b], 2.0[c])	−1.56	5.25	1.88
V_{Zn}	−2	2.43 (2.59[a], 2.7[b], 2.2[c])	−4.31	5.87	−0.87
O_i	−2	3.60 (4.3[b], 4.8[c])	−3.14	7.04	0.3
O_i	−1	2.82 (3.48[a], 3.3[b], 3.6[c])	−0.55	6.26	2.89
O_i	0	2.44 (2.49[a], 3.1[b], 3.3[c])	2.44	5.88	5.88
O_i	+1	2.36 (2.49[a])	5.73	5.80	9.17
O_i	+2	2.49	9.23	5.93	12.67
Zn_i	0	6.15 (5.74[a], 6.5[b], 4.5[c])	6.15	2.71	2.71
Zn_i	+1	4.42 (5.29[a], 4.6[b])	7.79	0.98	4.35
Zn_i	+2	2.87 (4.88[a], 2.9[b], 2.4[c])	9.71	−0.57	6.17
O_{Zn}	−2	5.05 (3.97[a])	−1.69	11.93	5.19
O_{Zn}	−1	3.69	0.32	10.57	7.2
O_{Zn}	0	2.70 (1.72[a], 3.7[c])	2.7	9.58	9.58
O_{Zn}	+1	2.51	5.88	9.39	12.76
O_{Zn}	+2	2.65	9.39	9.53	16.27
Zn_O	0	10.27 (10.43[a], 7.2[b], 8.4[c])	10.27	3.39	3.39
Zn_O	+1	7.99 (5.1[b])	11.36	1.11	4.48
Zn_O	+2	6.52 (8.56[a], 3.3[b], 5.4[c])	13.26	−0.35	6.39

[a] Reference [23].
[b] Reference [28].
[c] Reference [29].

in the gap, with possible charge states ranging from 2+ to 4−. It can serve as both donor and acceptor in ZnO. Furthermore, antisite defects, including both Zn_O and O_{Zn}, have also received considerable attention experimentally [7, 25]. However, the formation energies of the antisites are always higher than that of the vacancies and interstitials at both the oxygen-poor and oxygen-rich conditions, indicating the antisites hardly exist in practice.

It is well known that self-compensation by native donor-type defects plays an important role in the p-type doping bottleneck for ZnO. These donor-type defects always grow in oxygen-deficient environment. At the low oxygen chemical potential limit in figure 5(a), when the p-type condition that the Fermi energy is close to the VBM is assumed, the donor-type oxygen vacancy, zinc interstitial, and zinc antisite exhibit far lower formation energies than the acceptor-type zinc vacancy, oxygen interstitial, and oxygen antisite defects. This suggests that a heavy self-compensation effect for p-type doping could occur in the ZnO crystal grown in oxygen-deficient environment. On the other hand, when p-type conditions are assumed at the high oxygen chemical potential (figure 5(b)), the formation energies of zinc vacancies are only a little lower than those of the oxygen vacancies as the donor type, indicating that the effect of self-compensation should also be considered even for ZnO grown in oxygen-rich

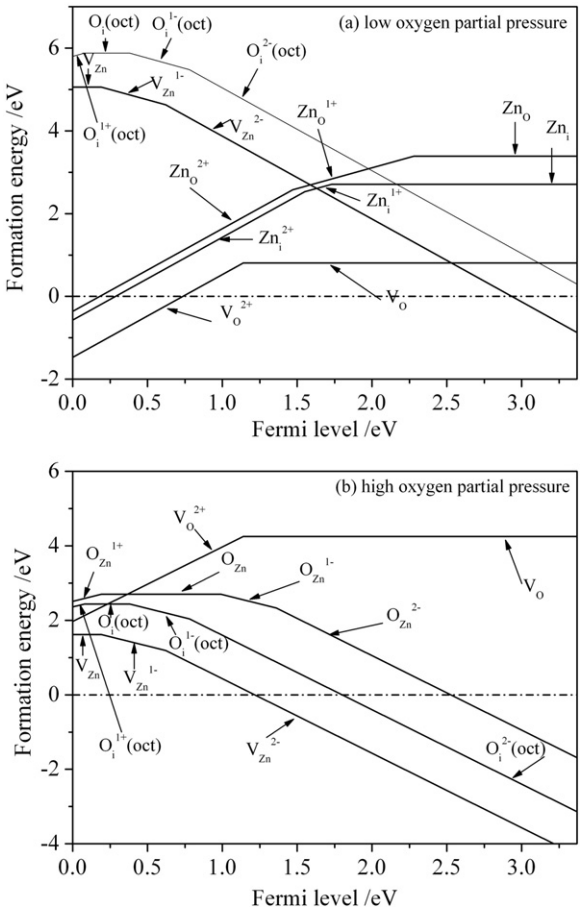

Figure 5. Calculated defect formation energies for selected defects shown in table 2 as a function of the Fermi level at (a) $\Delta\mu_O = -3.44$ (low oxygen chemical potential) and (b) $\Delta\mu_O = 0$ (high oxygen chemical potential), respectively.

environment. However, it is expected that the self-compensation effect could be significantly reduced by increasing the oxygen partial pressure in the crystal growth environment. This could possibly be employed to explain the p-type conduction in ZnO film grown by the ultrasonic spray pyrolysis method in atmospheric environment in our group [21, 22, 45, 46].

3.5. Defects/carrier concentration at finite temperature

Now we discuss the dependence of concentration of intrinsic defects and carriers on environmental conditions. The concentration was computed using equations (4) and (6) with the formation energies from figure 5. The Fermi level was obtained by requiring overall charge neutrality according to equation (7). Figure 6 *shows the concentration of intrinsic defects and carriers as function of oxygen chemical potential, using theoretical and experimental bandgap, respectively*. The temperature is given at 1000 K. From equation (7), we can see that different bandgap energy results in a different Fermi level, leading to different concentrations of charged defects and carriers, while the neutral defect concentration does not change with bandgap since the formation energy is not related to the Fermi level.

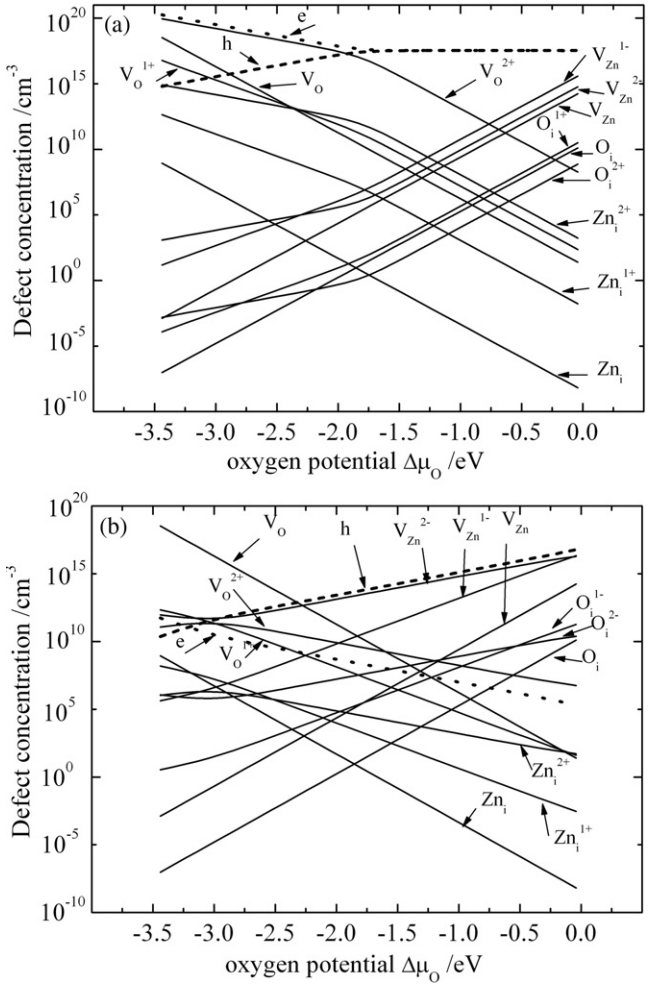

Figure 6. Concentration of intrinsic defects and carriers at 1000 K. The concentration was computed using equation (6) with the formation energy values from figure 4. The bandgap is given at (a) experimental and (b) theoretical bandgap, respectively. The Fermi level was obtained by requiring overall charge neutrality based on equation (12). Conduction electrons are identified as 'e' and holes in the valence band as 'h' with dotted and dashed lines, respectively.

Since the theoretical bandgap is severely underestimated, the carrier concentration is significantly overestimated by the calculation using the theoretical bandgap (figure 6(a)), with the maximum electron concentration of more than 10^{20} cm^{-3} at low oxygen chemical potential. To obtain a realistic description of electron concentrations in pure ZnO, the CBM is shifted upwards to the value of the experimental bandgap. The resulting carrier and defect concentrations are shown in figure 6(b). ZnO exhibits weak n-type behaviour at low oxygen chemical potential and converts into p type with the increase of oxygen chemical potential. ZnO shows strong p-type conduction at high oxygen chemical potential. From the dependence of the concentration of carriers on the temperature (figure 7(a)), it can be seen that at low oxygen chemical potential the Fermi level locates near the middle of the bandgap throughout a wide temperature range. Thus the conduction electron concentration (5.5×10^{11} cm^{-3} at 1000 K) is only a little higher than the hole concentration (2.3×10^{10} cm^{-3} at 1000 K), leading to weak

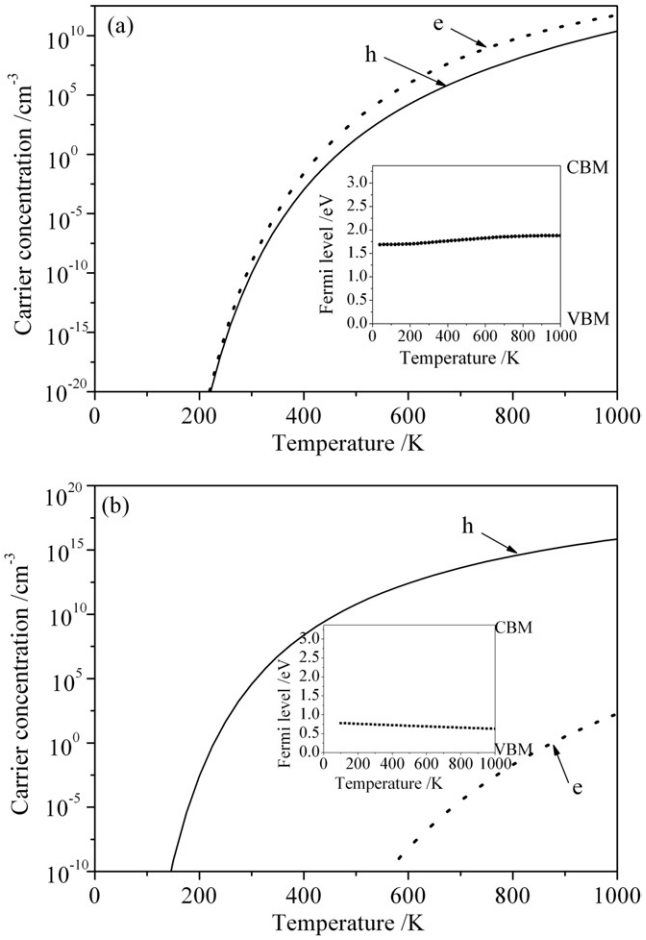

Figure 7. Calculated carrier concentration as a function of temperature at (a) $\Delta\mu_O = -3.44$ eV (low oxygen chemical potential limit) and (b) $\Delta\mu_O = 0$ (high oxygen chemical potential limit), using the experimental bandgap. Conduction electrons are identified as 'e' and holes in the valence band as 'h', using dotted and solid lines, respectively. The insets show the dependence of Fermi level on the temperature.

n-type conduction. On the other hand, the Fermi level is close to the VBM at high oxygen chemical potential (figure 7(b)), resulting in much higher hole concentration (7.2×10^{16} cm^{-3} at 1000 K) than electron concentration (1.8×10^5 cm^{-3} at 1000 K), which gives the system p-type behaviour. It is also noted from figure 6(a) that the conversion from n type to p type occurs at the oxygen chemical potential as low as -3.2 eV, indicating that the intrinsic ZnO can easily show p-type conduction, which contradicts the experimental findings, i.e., the strong n-type conduction in undoped ZnO. Two possible reasons can account for this. Firstly, the low conduction electron concentration results from the deep transition level of the oxygen vacancy, about 2.3 eV $(0/2+)$ below the CBM (figure 5), while the transition level of the zinc vacancy is only 0.2 $(0/1-)$ and 0.6 eV $(1-/2-)$ above the VBM. Experimental studies also confirmed that the oxygen vacancy forms a deep level in the bandgap of ZnO [47, 48]. Such low conduction electron concentration indicates that the intrinsic defects cannot play an important role in the n-type conductivity of unintentionally doped ZnO, which is in agreement with Kohan and Van de

Walle's conclusions [23, 27]. Secondly, the defect formation energies and transition levels can be affected by the GGA underestimation for the bandgap. Although many correction methods for the defect formation energies have been reported [9, 23, 28, 29], the reliability of these corrections needs further confirmation. Thus we refrain from using any corrections of defect formation energies for bandgap error. However, it is necessary for further study on the reliable correction for the bandgap error to obtain more reliable concentrations for defects and carriers in wide-bandgap semiconductors by first principle calculations.

Despite the difference in calculated concentration of defects and carriers between using theoretical and experimental gaps, the general conclusion can be extracted from both calculations that oxygen and zinc vacancies always show the highest concentration over the oxygen chemical potential range defined in equation (5). Oxygen and zinc interstitials also exhibit a considerable concentration, while oxygen and zinc antisites have much lower concentrations than the vacancies (not shown in figure 6).

3.6. Origin of green photoluminescence

As mentioned in the introduction, ZnO often exhibits a photoluminescence band in the visible range, centred between 2.3 and 2.5 eV. The most likely origin for this emission is attributed to the oxygen and zinc vacancies, as the defects with highest concentration in ZnO. Reference [16] suggested that the zinc vacancy should give rise to green photoluminescence according to the transition level between the -1 and -2 charge states. However, few experimental results can confirm their conclusion. Many experimental studies indicated that the green emission band always occurs in ZnO grown in an oxygen-deficient environment, while the emission becomes very weak in ZnO grown at high oxygen partial pressure [21, 22]. Our calculations show that at low oxygen chemical potential the oxygen vacancy has a concentration about 10^5 times higher than the zinc vacancy. In addition, the transitional level between 0 and $+2$ charge states of V_O is 1.1 eV above the valence band maximum, very close to the value (1.2 eV above VBM) from nonlinear spectroscopy measurements [48]. The resulting transition energy between the conduction band and the defect level of the oxygen vacancy (2.3 eV) corresponds to the observed green emission. The fact that V_O has a deep level within the bandgap and is the most abundant defect in an oxygen-deficient environment indicates the high possibility of the oxygen vacancy as the source of green luminescence.

4. Conclusion

In conclusion, first-principles calculations with the PAW method in the GGA were applied to study the intrinsic defects in wurtzite ZnO. Systematic checks for the reliability of our calculations were discussed in detail, and the error bar for each calculated defect formation energies is estimated carefully. Comparisons with previous studies were also made. Our calculations show that oxygen and zinc vacancies are the dominant intrinsic donor and acceptor defects in ZnO, depending on the oxygen chemical potential. We also suggest that intrinsic defects are not the main source of strong n-type conductivity in unintentionally doped ZnO. The deep donor oxygen vacancy is proposed to be the origin of the experimentally observed visible photoluminescence band between 2.3 and 2.5 eV. Our calculations are useful for further understanding the physics of intrinsic defects and some interesting phenomena in unintentionally doped ZnO.

Acknowledgments

This work is supported by the Ministry of Science and Technology of China through the 973 project under grant No 2002CB613306, and the National Natural Science Foundation of China under grant No 90401010.

References

[1] Levinson L M, Castleberry D E and Becker C A 1982 *J. Appl. Phys.* **53** 3859
[2] Ono S, Yamazaki O, Ohji K, Wasa K and Hayakawa S 1978 *Appl. Phys. Lett.* **33** 217
[3] Service R F 1997 *Science* **276** 895
[4] Norton D P, Heo Y W, Ivill M P, Ip K, Pearton S J, Chisholm M F and Steiner T 2004 *Mater. Today* **7** 34
[5] Pearton S J, Norton D P, Ip K, Heo Y W and Steiner T 2005 *Prog. Mater. Sci.* **50** 293
[6] Özgür Ü, Alivov Y I, Liu C, Teke A, Reshchikov M A, Doğan S, Avrutin V, Cho S J and Morkoç H 2005 *J. Appl. Phys.* **98** 041301
[7] Lin B, Fu Z and Jia Y 2001 *Appl. Phys. Lett.* **79** 943
[8] Zhang S B 2002 *J. Phys.: Condens. Matter* **14** R881
[9] Wei S H 2004 *Comput. Mater. Sci.* **30** 337
[10] Yam Park C H, Zhang S B and Wei S H 2002 *Phys. Rev.* B **66** 073202
[11] Mahan G D 1983 *J. Appl. Phys.* **54** 3825
[12] Hoffman J W and Lauder I 1970 *Trans. Faraday Soc.* **66** 2346
[13] Ziegler E, Heinrich A, Oppermann H and Stover G 1981 *Phys. Status Solidi* a **66** 635
[14] Hagemark K I 1976 *J. Solid State Chem.* **16** 293
[15] Look D C, Hemsky J W and Sizelove J R 1999 *Phys. Rev. Lett.* **82** 2552
[16] Tuomisto F, Ranki V, Saarinen K and Look D C 2003 *Phys. Rev. Lett.* **91** 205502
[17] Jin B J, Im S and Lee S Y 2000 *Thin Solid Films* **366** 107
[18] Kang H S, Kang J S, Pang S S, Shim E S and Lee S Y 2003 *Mater. Sci. Eng.* B **102** 313
[19] Fang Z, Wang Y, Xu D, Tan Y and Liu X 2004 *Opt. Mater.* **26** 239
[20] Ma Y, Du G T, Yang S R, Li Z T, Zhao B J, Yang X T, Yang T P, Zhang Y T and Liu D L 2004 *J. Appl. Phys.* **95** 6268
[21] Bian J M, Li X M, Chen L D and Yao Q 2004 *Chem. Phys. Lett.* **393** 256
[22] Bian J M, Li X M, Zhang C Y, Yu W D and Gao X D 2004 *Appl. Phys. Lett.* **85** 4070
[23] Kohan A F, Geder G, Morgan D and Van de Walle C G 2000 *Phys. Rev.* B **61** 15019
[24] Liu M, Kitai A H and Mascher P 1992 *J. Lumin.* **54** 35
[25] Reynolds D C, Look D C, Jogai B and Morkoc H 1997 *Solid State Commun.* **101** 643
[26] Sun Y, Xu P, Shi C, Xu F, Pan H and Lu E 2001 *J. Electron Spectrosc. Relat. Phenom.* **114–116** 1123
[27] Van de Walle C G 2001 *Physica* B **308–310** 899
[28] Zhang S B, Wei S-H and Zunger A 2001 *Phys. Rev.* B **63** 075205
[29] Oba F, Nishitani S R, Isotani S, Adachi H and Tanaka I 2001 *J. Appl. Phys.* **90** 824
[30] Martin R M 1992 *Electronic Structure: Basic Theory and Practical Methods* (Cambridge: Cambridge University Press) pp 153–9, 258–63
[31] Blochl P E 1994 *Phys. Rev.* B **50** 17953
[32] Kresse G and Joubert D 1999 *Phys. Rev.* B **59** 1758
[33] Kresse G and Hafner J 1993 *Phys. Rev.* B **47** R558
Kresse G and Furthmuller J 1996 *Phys. Rev.* B **54** 11169
[34] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
[35] Van de Walle C G and Neugebauer J 2004 *J. Appl. Phys.* **95** 3851
[36] Zhang S B 2002 *J. Phys.: Condens. Matter* **14** R881
[37] Lee E C and Chang K J 2004 *Phys. Rev.* B **70** 115210
[38] Dean J A 1992 *Lange's Handbook of Chemistry* 14th edn (New York: McGraw-Hill)
[39] Perdew J P and Zunger A 1981 *Phys. Rev.* B **23** 5048
[40] Lento J, Mozos J L and Nieminen R M 2002 *J. Phys.: Condens. Matter* **14** 2637
[41] Makov G and Payne M C 1995 *Phys. Rev.* B **51** 4014
[42] Schultz P A 1999 *Phys. Rev.* B **60** 1551
Schultz P A 2000 *Phys. Rev. Lett.* **85** 1942
[43] Nozaki H and Itoh S 2000 *Phys. Rev.* E **62** 1390
[44] Monkhorst H J and Pack J D 1976 *Phys. Rev.* B **13** 5188
[45] Bian J M, Li X M, Gao X D, Yu W D and Chen L D 2004 *Appl. Phys. Lett.* **84** 541
[46] Bian J M, Li X M, Zhang C Y, Chen L D and Yao Q 2004 *Appl. Phys. Lett.* **84** 3783
[47] Kasai P 1963 *Phys. Rev.* **130** 989
[48] Gavryushin V, Raciukaitis G, Juodzbalis D, Kazlauskas A and Kubertavicius V 1994 *J. Cryst. Growth* **138** 924