First-principles calculations of solubilities and doping limits: Li, Na, and N in ZnSe

Chris G. Van de Walle
Xerox Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, California 94304

D. B. Laks*
National Renewable Energy Laboratory, Golden, Colorado 80401

G. F. Neumark
Division of Metallurgy and Materials Science, Columbia University, New York, New York 10027

S. T. Pantelides
IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598
(Received 1 September 1992)

We present a comprehensive theoretical approach to determine concentrations of dopant impurities in semiconductors. The formalism is applied to the problem of acceptor doping in ZnSe. Formation energies and concentrations of impurities and native defects are expressed as a function of chemical potentials, for which experimentally accessible ranges are calculated. We show that limitations in the achievable hole concentrations can be explained by two mechanisms: one is the competition between various substitutional and interstitial configurations (compensation), the other is the solubility limit imposed by formation of other phases. Nitrogen is most promising among the dopants examined.

I. INTRODUCTION

Limits to semiconductor doping have been widely discussed both in III-V and II-VI compounds. In wide-band-gap semiconductors the problem is particularly acute because typically one type of conduction (n-type or p-type) is very difficult to obtain. Detailed understanding of these phenomena has been lacking. In this paper we present a formalism that allows the determination of defect concentrations, impurity solubilities, and doping levels. It includes a unifying treatment of the various interactions of the dopant with the host lattice (in substitutional or interstitial sites), the role of native defects, and the factors that determine solubility. The key quantities that enter this formulation can be obtained from first-principles electronic-structure calculations.

Our formalism entails the following steps.

(1) Calculation of the total energies of all native defects and of the various configurations that can be assumed by the impurity in the crystal, including lattice relaxations and different charge states.

(2) Application of thermodynamics to express the relevant equilibrium concentrations at the temperature of interest, and determination of the resulting Fermi level from the condition of overall charge neutrality for all impurity configurations, native defects, and free carriers.

At this juncture the results remain functions of two chemical potentials (one for the host crystal, which controls the stoichiometry, and one for the impurity) which are free parameters to be fixed by growth conditions. The physical meaning of these chemical potentials and the way in which they enter the formalism will be discussed in detail. Thermodynamics imposes bounds on the experimentally accessible range of these chemical potentials; the bounds result from the last step of the formalism:

(3) Calculation of the heats of formation of competing phases that can be formed out of the constituents (i.e., the impurity and the component elements of the semiconductor).

By imposing these bounds we obtain limits on impurity concentrations, i.e. we can calculate solubilities.

We illustrate the approach with the technologically important example of ZnSe, in which n-type doping poses no difficulties, but well-conducting, reproducible p-type doping has been very hard to achieve.[1] Despite some impressive recent experimental advances,[2–5] the cause of the doping problem has remained unclear. Lithium was the first dopant to yield reproducible, well-conducting p-type ZnSe, with a net acceptor concentration of about 10^{17} cm^{-3}.[2,4] More recently, N doping up to 10^{18} cm^{-3} was achieved and led to the fabrication of a blue semiconductor laser.[6] In the case of Li, our results will demonstrate quantitatively the competition between substitutional and interstitial impurity configurations,[7] and identify a regime where the desired substitutional form dominates (earlier work[8] that proposed this competition as the source of compensation did not recognize the existence of such different regimes). Our results will also show that there is a second overriding cause that limits doping, namely the overall solubility which is constrained by the formation of a Li_2Se phase.[9] These conclusions agree with experimental observations on Li-doped ZnSe; more importantly, they provide guidelines for optimizing

growth conditions.

Our investigations of Na indicate qualitative similarities to Li, but significant quantitative differences which render Na unsuitable for p-type doping of ZnSe. Indeed, its low solubility explains the failure of doping attempts with Na.[10] Nitrogen, finally, does not exhibit a substitutional/interstitial competition and, in addition, has the highest solubility.

One of the strengths of the formalism is that it treats native point defects (vacancies, self-interstitials, and antisites) and dopant impurities on an equal footing, allowing us to investigate whether native defects can form a significant source of compensation.[11] We find that under appropriate growth conditions the native defect concentration is usually so low as to be unimportant. We previously arrived at this conclusion from a study of native defect concentrations as a function of Fermi-level position, in which the exact nature of the dopant impurities was left unspecified.[12] Our current results confirm that native defects do *not* form a generic source of compensation in ZnSe. We also present more detailed information on defect concentrations under various growth conditions.

The present investigation of acceptor impurities in ZnSe relates to an experimental problem of high current interest due to the impact on a blue semiconductor laser; however, we stress that the formalism is a general one that can be applied to the study of doping in any semiconductor system.

II. METHODS

A. Total-energy calculations

In this section we describe how to calculate concentrations of defects and impurities in the semiconductor. In order to obtain quantitative results, one needs reliable values for the total energies of defects and impurities; we have obtained such values from first-principles calculations. The calculations are based on density-functional theory in the local-density approximation,[13] and *ab initio* pseudopotentials.[14] Scalar relativistic effects are included in the pseudopotentials, but spin-orbit splitting is neglected; our calculated bands are therefore averages over the states which would be split due to spin-orbit interactions. The spin-orbit splitting can be introduced as a perturbation. The Fermi-level positions which we will discuss should still be interpreted as referred to the top of the valence band (Γ_8). We use a mixed-basis approach, ensuring an accurate description of the structural properties by explicitly including the d states of the Zn atoms. The basis set contains plane waves with kinetic energy up to 9 Ry, and pseudoatomic orbitals on the Zn and N atoms.[12] In order to achieve a proper description of Li and Na we implement a nonlinear core exchange-correlation correction.[15]

The defect calculations are performed in a supercell geometry, with 32-atom supercells providing adequate accuracy. Relaxations of up to two shells of neighbors are included. Additional details about the calculational approach are given in Ref. 16. We have used this approach to obtain total energies for the dopant impurities (Li, Na, and N) which are the subject of this study, in their various configurations in the lattice. Our calculations are typically carried out for the charged (positive or negative) state of the dopant. To avoid divergence of the long-range Coulomb terms, the $\mathbf{G}=\mathbf{0}$ terms in the total energy are always calculated for a neutral system. A justification of this self-consistent approach to treat charge states of impurities was given in Ref. 17. Because of the extended nature of the wave function the neutral charge state of a shallow impurity is difficult to treat within the supercell formalism; instead, we use experimental activation energies[18] to determine the formation energies of the neutral charge state. Finally, we make use of our previously calculated[12,16] energy values for native defects in all relevant charge states.

B. Formation energies, concentrations, and chemical potentials

The equilibrium concentration of an impurity or defect D_i is given by

$$[D_i] = N_{\text{sites}} \exp\left(-\frac{E_{\text{form}}(D_i)}{kT}\right), \tag{2.1}$$

where N_{sites} is the appropriate site concentration [e.g., for substitutional Li (Li_{Zn}), N_{sites} is the number of substitutional Zn sites in the crystal, $2.2\times10^{22}\text{cm}^{-3}$], and E_{form} is the formation energy. The energy appearing in Eq. (2.1) is a Gibbs free energy, which should include a pressure-dependent term; however, this term can be neglected for the solid phase. The Gibbs free energy also contains an entropy contribution; these terms are generally small,[12] and they also tend to cancel when comparing relative free energies.[19] The assumption of thermodynamic equilibrium, which underlies the formalism, is expected to be satisfied, particularly in light of the high mobility of various defects and impurities studied here.[12]

Before we give a general definition of the formation energy of an impurity or defect in the compound semiconductor, we illustrate the concept with the example of a Li atom on a substitutional Zn site:

$$E_{\text{form}}(\text{Li}_{\text{Zn}}^{-}) = \mathcal{E}(\text{Li}_{\text{Zn}}^{-}) - \mu_{\text{Li}} + \mu_{\text{Zn}} - E_F. \tag{2.2}$$

$\mathcal{E}(\text{Li}_{\text{Zn}}^{-})$ is the calculated energy of a supercell containing the $\text{Li}_{\text{Zn}}^{-}$ impurity, minus the energy of a reference cell containing the pure bulk semiconductor. These energies are obtained from first-principles calculations, which are described in Sec. II A. The other terms in Eq. (2.2) contain chemical potentials, the physical significance of which we now discuss in some detail.

μ_{Li} is the chemical potential of Li. This term enters because the formation energy is the *difference* between the energy of Li as an impurity, and its energy in a reference state. The reference corresponds to a reservoir of Li atoms, whose energy (at $T=0$) by definition is the chemical potential. This chemical potential depends on the abundance of Li under the relevant growth conditions.[20] For an element in thermal equilibrium with the gas phase, the chemical potential can be related to the partial pressure of the gas;[20] for an ideal gas with partial pressure

p one has $\mu = \mu^0 - kT \ln p$. In the literature one often finds studies of defect concentrations as a function of partial pressures. We prefer to work with chemical potentials for the following reasons: (a) Chemical potentials are thermodynamically defined as energy values, which can be directly related to the energies which we calculate from first principles. (b) Although the assumption of thermodynamic equilibrium is likely to be satisfied within the solid, allowing the use of expressions such as Eq. (2.1), it is uncertain to what extent equilibrium is established between the solid and a surrounding gas under experimental conditions such as molecular-beam epitaxy (MBE). Knowledge of the chemical potential in the gas may therefore not necessarily reflect the relevant chemical potential for the solid. (c) Even if thermodynamic equilibrium with the gas is assumed, the relationship between chemical potential and gas pressure is not well known since the gas sources used in MBE do not obey simple ideal gas laws. While this precludes a quantitative determination of chemical potentials in terms of experimentally accessible quantities, we will see that the chemical potentials are subject to rigorous bounds that can be directly related to experimental conditions.

The Zn chemical potential μ_{Zn} appears in Eq. (2.2) because, in order to make room for the substitutional impurity, a Zn atom has to be removed to its reservoir. It is very important to realize that μ_{Zn} should be treated as a variable; indeed, in a compound semiconductor only the *sum* of the chemical potentials of the constituents is fixed, and equal (at T=0) to the energy of a two-atom unit of the material:

$$\mu_{\mathrm{Zn}} + \mu_{\mathrm{Se}} = \mu_{\mathrm{ZnSe}}. \tag{2.3}$$

In an elementary semiconductor, this condition would uniquely determine the value of the chemical potential; additional freedom exists, however, in a compound semiconductor. We will therefore explicitly present our results as a function of chemical potentials. Equation (2.3) fixes μ_{Se} once μ_{Zn} is chosen; alternatively, μ_{Se} could be chosen as the free variable, leading to a fixed μ_{Zn}.

The last term in Eq. (2.2) is the Fermi level E_F, i.e., the energy of the reservoir delivering the electron responsible for the negative charge on the impurity.

In general the total energy $E_{\mathrm{tot}}(D_i)$ for a defect D_i will be determined from a calculation for a supercell containing n_i^{Zn} Zn atoms, n_i^{Se} Se atoms, and n_i^{Li} Li atoms (we continue to use Li as a sample impurity, but the formulas are valid for a general impurity). The defect formation energy $E_{\mathrm{form}}(D_i)$ is then

$$\begin{aligned} E_{\mathrm{form}}(D_i) &= E_{\mathrm{tot}}(D_i) - n_i^{\mathrm{Zn}}\mu_{\mathrm{Zn}} - n_i^{\mathrm{Se}}\mu_{\mathrm{Se}} \\ &\quad - n_i^{\mathrm{Li}}\mu_{\mathrm{Li}} - n_i^e E_F \\ &= \mathcal{E}(D_i) - \Delta n_i \mu_{\mathrm{Zn}} - n_i^{\mathrm{Li}}\mu_{\mathrm{Li}} - n_i^e E_F, \end{aligned} \tag{2.4}$$

$$\mathcal{E}(D_i) = E_{\mathrm{tot}}(D_i) - n_i^{\mathrm{Se}}\mu_{\mathrm{ZnSe}}, \tag{2.5}$$

$$\Delta n_i = n_i^{\mathrm{Zn}} - n_i^{\mathrm{Se}}, \tag{2.6}$$

where n_i^e is the number of excess electrons in the defect, and Δn_i is the number of extra Zn atoms that must be added to form the defect (e.g., +1 for $\mathrm{Zn_i}$ and $\mathrm{V_{Se}}$, −2 for $\mathrm{Se_{Zn}}$, etc.). Here, we treat μ_{Zn} as an independent variable and use Eq. (2.3) to remove μ_{Se} from the expression for $E_{\mathrm{form}}(D_i)$; alternatively, we could treat μ_{Se} as independent and eliminate μ_{Zn}.

C. Self-consistent solution

An expression based on Eq. (2.4) can be written down for all configurations of the impurity, in their various charge states, as well as for all native defects. Once the formation energy is known, the concentration of a specific defect or impurity can be obtained from Eq. (2.1). At this point, all concentrations are still functions of the chemical potentials (μ_{Zn} and μ_{Li}), as well as of the Fermi level (E_F). The chemical potentials, as explained above, are independent parameters; we will therefore express all our results as functions of these chemical potentials. The Fermi level, however, is *not* an independent variable, since it is determined by the condition of charge neutrality:

$$\text{net charge} = 0 = p - n - \sum_i n_i^e [D_i], \tag{2.7}$$

where p and n are the hole and electron densities, respectively. These free-carrier densities are determined from the standard semiconductor equations. The charge conservation equation provides for an interaction between the concentrations of all charged defects through their influence on the Fermi level. For example, a positively charged defect produces extra free electrons that raise the Fermi level; the higher Fermi level, in turn, increases the concentrations of all negatively charged defects and lowers the concentrations of all positively charged defects. As pointed out by Zhang and Northrup,[21] this "negative feedback" reduces the sensitivity of the final results to possible inaccuracies in our first-principles energies. Using this prescription, all of the defect formation energies, and hence the concentrations $[D_i]$, are unique functions of μ_{Zn}, μ_{Li}, and the temperature T.

The choice of the chemical potential μ_{Zn} also determines the stoichiometry; the stoichiometry parameter X can be defined as

$$X = \frac{N_{\mathrm{Se}} - N_{\mathrm{Zn}}}{N_{\mathrm{Se}} + N_{\mathrm{Zn}}} = \frac{-\sum_i \Delta n_i [D_i]}{2N_{\mathrm{sites}}}, \tag{2.8}$$

where N_{Zn} and N_{Se} are the total numbers of Zn and Se atoms in the crystal. Only deviations from stoichiometry due to native defects are included here. X is positive for Se-rich material and negative for Zn-rich material. In this paper, we express all our results in terms of chemical potentials. Alternatively, we could present the results as a function of the stoichiometry parameter, but because of the one-to-one correspondence between chemical potential and stoichiometry no new information would be obtained.

D. Bounds on the chemical potentials

Now we discuss how the relevant range of the chemical potentials is determined. For this purpose one has to consider the various phases that can be formed out

of the constituents.[19,20] For instance, μ_{Zn} is bounded from above by the energy of a Zn atom in Zn metal: $\mu_{Zn}^{max} = \mu_{Zn(bulk)}$. Indeed, if one would try to raise μ_{Zn} above this level, Zn metal would be preferentially formed. Similarly, μ_{Se} has an upper bound imposed by bulk Se. Furthermore,

$$\mu_{ZnSe} = \mu_{Zn(bulk)} + \mu_{Se(bulk)} + \Delta H_f(ZnSe), \quad (2.9)$$

where $\Delta H_f(ZnSe)$ is the heat of formation of ZnSe (ΔH_f is negative for a stable compound). Combined with Eq. (2.3) this expression can be used to impose a *lower* bound on the Zn chemical potential, given by $\mu_{Zn}^{min} = \mu_{Zn(bulk)} + \Delta H_f$ (ZnSe). A lucid discussion of similar arguments, in the context of surface recontructions, has been given in Ref. 19. The Zn chemical potential can thus vary over a range corresponding to the heat of formation of ZnSe.

To find an upper bound on the chemical potential of the dopant we explore the various compounds that the impurity can form in its interactions with the system. For Li, a possible upper bound on μ_{Li} is of course imposed by Li (bulk) metal. However, the most stringent constraint arises from the compound Li_2Se, which leads to the following constraint on the chemical potentials:

$$2\mu_{Li} + \mu_{Se} = \mu_{Li_2Se} = 2\mu_{Li(bulk)} + \mu_{Se(bulk)} + \Delta H_f(Li_2Se). \quad (2.10)$$

Numerical results for the heats of formation, as well as practical applications of the bounds on the chemical potentials, will be given in the following section.

III. RESULTS AND DISCUSSION

A. Lithium

1. Configurations of Li in the lattice

We have analyzed various possible configurations and charge states of the lithium impurity in the lattice. The substitutional acceptor Li_{Zn}^{-} induces virtually no relaxation of the surrounding host atoms. For the lithium interstitial (Li_i^+), which is a shallow donor, we find the T_d site surrounded by Se atoms (T_d^{Se}) to be 0.2 eV lower in energy than the T_d^{Zn} site. For the interstitials, the energy gained by relaxation of the host atoms is smaller than 0.1 eV. We have also studied other interstitial positions, allowing us to estimate that the barrier for migration of the interstitial is less than 0.5 eV (i.e., a Li interstitial can move readily, even at room temperature). Finally, we have also investigated Li on a substitutional Se site, but found this configuration to have a prohibitively large formation energy.

2. Contour plots of total Li concentration

Our results are presented in the form of contour plots, which allow us to explicitly show the dependence on the chemical potentials μ_{Zn} and μ_{Li}. As explained in Sec. II C, there is no explicit dependence on Fermi energy, since it is determined by charge neutrality. Figure 1(a) shows a contour plot for the total concentration of Li in ZnSe, at $T = 600$ K, which is a typical temperature in MBE growth of ZnSe:Li.[2,4]

We first discuss the contour lines themselves. The total Li concentration ([Li]) increases with increasing μ_{Li}, because it becomes more favorable for the impurity to dissolve in the semiconductor as the energy of the reservoir rises. Similarly, [Li] increases with *decreasing* μ_{Zn}, which is the energy of the reservoir to which Zn needs to be removed in order to accommodate Li on Zn sites.

3. Competition between interstitials and substitutionals

The formation energy for Li in a substitutional location was given in Eq. (2.2). For the interstitial site, where Li is a shallow donor, we have

$$E_{form}(Li_i^+) = \mathcal{E}(Li_i^+) - \mu_{Li} + E_F. \quad (3.1)$$

$\mathcal{E}(Li_i^+)$ is the calculated energy of an interstitial Li at its most stable site, which is at the tetrahedral interstitial site surrounded by Se atoms. Inspection of Eqs. (2.2) and (3.1) reveals that as the Fermi level moves down (i.e., as the material becomes increasingly p-type), the formation energy of the acceptor species rises, whereas the formation energy of the donor species goes down. This predicts the existence of a limiting Fermi-level position (maximum hole concentration), which can be obtained by equating the two formation energies. Attempts to push the Fermi level lower would result in preferential formation of donors, which would push the Fermi level back up. Incorporation of additional Li leaves the Fermi level unchanged, as each substitutional acceptor is immediately compensated by an interstitial donor.

The position of the Fermi level (at 600 K) is shown in Fig. 1(b); Li interstitials are responsible for the flattening of the contour lines on the right-hand side of the plot. For a fixed value of μ_{Zn}, the Fermi level saturates as μ_{Li} is raised, even though the total Li concentration still increases [see Fig. 1(a)]. If no interstitials could form, the contour lines would continue to rise with the same slope as in the left-hand side of the plot. The interstitials cause compensation and limit the achievable hole concentration.[7] Their presence has been experimentally observed.[22,23] A contour plot of the Li interstitial concentration is shown in Fig. 1(c).

The position at which the Fermi level saturates due to interstitial compensation still depends on the Zn chemical potential, as can be noted in Fig. 1. Our results differ markedly from those of Ref. 8, where it was concluded that compensation by Li interstitials would always dominate. The authors of Ref. 8 did not recognize that the level of compensation depends on the Zn chemical potential, and hence on the growth conditions. This dependence explains the experimental observation that the degree of compensation by Li interstitials varies widely in different samples.[2] Our results actually provide a guideline for optimizing the growth conditions: low values of μ_{Zn} lead to lower compensation, as well as higher Li_{Zn} concentrations.

4. Bounds on chemical potentials — solubilities

In order to determine solubility limits, we need to use the information about bounds on the chemical potentials discussed in Sec. II D. The bounds on the Zn chemical potential are shown as the horizontal lines in Fig. 1. For Li, the chemical potential is limited by formation of the compound Li_2Se. Formation of Li_2Se on the growing ZnSe surface in MBE has actually been experimentally observed in the case of heavy Li doping.[24] The compound Li_2Se leads to the line with slope +2 in Fig. 1, which was defined in Eq. (2.10). The point where this line intersects the lower bound on μ_{Zn} is given by $\mu_{Li}^0 = \mu_{Li(bulk)} + \frac{1}{2}\Delta H_f(Li_2Se)$. Our calculated heats of formation for the various compounds are listed in Table I. For comparison, we also list experimental values. The deviations are in line with the expected accuracy of the method.

Our calculated contours, together with the bounds on the chemical potentials, provide important insights in the ability to dope ZnSe with Li. We note that, over much of the range of the Li and Zn potentials, the maximum Li concentration is slightly higher than 10^{18} cm^{-3}. The fact that the slope of the contours in this region coincides with the slope of the Li_2Se boundary in Fig. 1(a) is accidental, caused by the fact that in this region the removal of *one* Zn atom leads to the incorporation of *two* Li atoms (one substitutional and one interstitial). The highest Li concentration (and lowest Fermi level, i.e., highest hole concentration) occurs in the lower right-hand corner of the accessible region, for $\mu_{Zn} = \mu_{Zn}^{min}$ and $\mu_{Li} = \mu_{Li}^0$. The

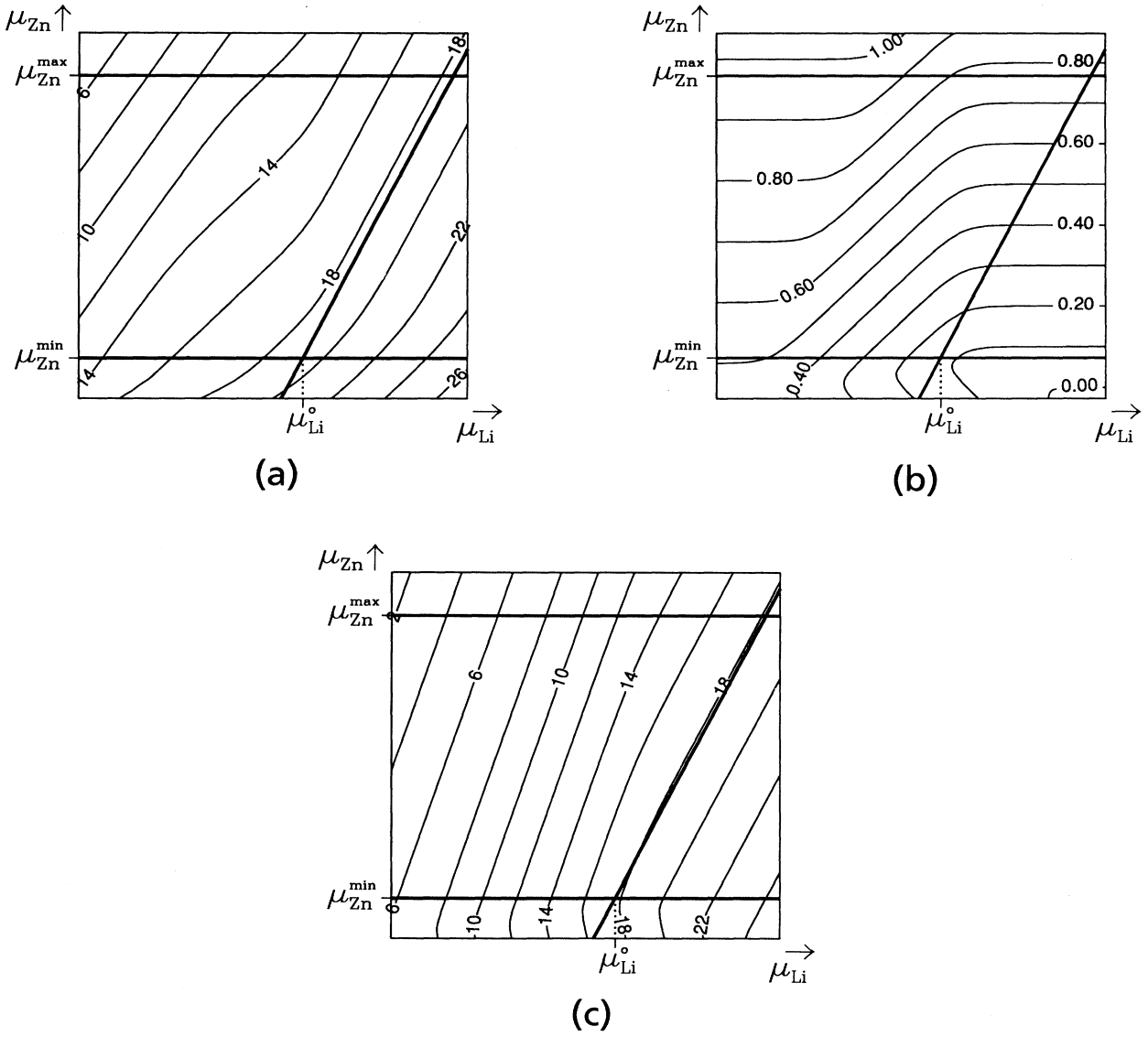

FIG. 1. Contour plots of (a) $\log_{10}$ [Li], where [Li] is the total Li concentration in cm^{-3}, (b) Fermi level (in eV, referred to the top of the valence band), and (c) $\log_{10}$ [Li_i], where [Li_i] is the interstitial Li concentration in cm^{-3}, at 600 K in ZnSe:Li, as a function of Zn and Li chemical potentials. Solid lines indicate bounds on μ_{Zn} and μ_{Li}.

TABLE I. Theoretical and experimental (Ref. 25) heats of formation (in eV per formula unit) for various materials containing Zn, Se, Li, Na, and N. Also listed is the minimum formation energy for the neutral substitutional acceptor in ZnSe, and the corresponding minimum Fermi-level position (in eV, referred to the top of the valence band), at 600 K.

	Solubility-limiting compound	$\Delta H_f^{\text{theor}}$	ΔH_f^{expt}	$E_{\text{form}}^{\text{min}}$	E_F
ZnSe		-1.39	-1.69		
ZnSe:Li	Li_2Se	-4.12	-3.96	0.46	0.13
ZnSe:Na	Na_2Se	-3.13	-3.54	1.08	0.44
ZnSe:N	Zn_3N_2		-0.24	0.38	0.09

corresponding formation energy of the *neutral* acceptor, and the self-consistently determined Fermi level are also listed in Table I. At this point of highest Li incorporation, the total Li concentration is 1.7×10^{19} cm^{-3}; fewer than 3% of these Li atoms occur in the form of interstitials.

5. Discussion

Our calculated differences in formation energies and heats of formation have an estimated error margin of ± 0.1 eV. At a temperature of 600 K, 0.12 eV roughly corresponds to an order of magnitude in concentration. Also, contours with values of [Li] higher than 10^{19} cm^{-3} are probably inaccurate because Eq. (2.1) is only valid for dilute concentrations; however, these contours fall outside the physically accessible range anyway. While these uncertainties should be kept in mind when considering plots such as Fig. 1, the qualitative and even quantitative insights are still clear. Some additional conclusions can be drawn. First, even though all native point defects were explicitly included in the calculations, their concentrations are very small over the whole of the accessible range in Fig. 1. The effect of native defects is noticeable for low μ_{Zn} values, causing bending of the contour lines; however, their concentration would only become important if $\mu_{\text{Zn}} < \mu_{\text{Zn}}^{\text{min}}$, which is physically not allowed. The dominant native defect is the Se_{Zn} antisite, which is a donor. Figure 2 shows a contour plot of the $\text{Se}_{\text{Zn}}^{2+}$ concentration. At the point of highest Li incorporation, the concentration is $[\text{Se}_{\text{Zn}}]=2.6 \times 10^{17}$ cm^{-3}, which is two orders of magnitude smaller than the Li concentration. Clearly the native defect concentration is too low to play any significant role in compensation. However, the concentration may be high enough to be detectable experimentally. Other native defects have concentrations significantly smaller (by more than four orders of magnitude) than the Se_{Zn} antisite.

The contour plots presented here were made for a temperature of 600 K, which is typical for MBE growth of ZnSe. The qualitative features of the plots do not change when we change the temperature (within physically reasonable limits). To illustrate the quantitative effect of temperature changes, as we lower the temperature from 600 to 500 K, we find that the total Li concentration is reduced by a factor of 5; the concentration of interstitial Li drops by more than an order of magnitude; and the concentration of the dominant native defect (Se_{Zn}) drops by almost two orders of magnitude.

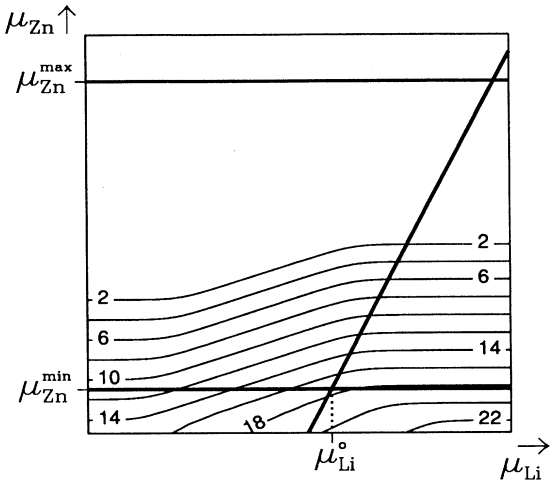

FIG. 2. Contour plot of $\log_{10}[\text{Se}_{\text{Zn}}^{2+}]$, the Se antisite concentration in cm^{-3}, at 600 K in ZnSe:Li, as a function of Zn and Li chemical potentials. Solid lines indicate bounds on μ_{Zn} and μ_{Li}.

A final point relates to doping of ZnSSe alloys with Li (alloys containing 6% S are commonly used to obtain lattice matching with GaAs substrates): since Li_2S is even more stable than Li_2Se (larger $|\Delta H_f|$), the bound on μ_{Li} in the ZnSSe:Li system will lie even lower, leading to reduced solubility in the alloy.

6. Complex formation

So far we have only talked about *isolated* point defects and impurities. In principle we should also consider complexes. Although our formalism is general enough to include any possible complexes, an exhaustive treatment is computationally prohibitive. Inspection of expressions for formation energies actually shows that a complex will only occur in appreciable concentrations (i.e., concentrations on the order of or larger than those of the individual defects out of which it is formed) if the binding energy exceeds the larger of the two formation energies of the individual components of the complex. This consideration makes it less likely that complexes would play an important role.

The only complex we have investigated as part of the current study is one consisting of a Li interstitial and a Li substitutional.[26] Formation of such complexes seems plausible, since the interstitial is quite mobile, and the acceptor and donor are Coulombically attracted. Details of the structure will be published elsewhere.[27] The binding energy of this complex is ~ 0.3 eV. This value is small enough so that these complexes are largely dissociated at a growth temperature of 600 K (in other words, their concentration is small compared to the concentration of the individual components, as discussed in the preceding paragraph). If we assume, however, that the con-

centration of Li substitutional and Li interstitial atoms is determined at the growth temperature, and remains fixed as the sample is cooled down, then the concentration of Li_{Zn}-Li_i pairs will increase as the temperature is lowered. The presence of such complexes should be taken into account in analyses of Fermi level positions and carrier concentrations at room temperature and below.[23,27]

B. Sodium

We now address Na, another column-I impurity which has been considered as an acceptor dopant in ZnSe.[10] The contour plots for the ZnSe:Na system are shown in Fig. 3. They are qualitatively similar to those in Fig. 1, but exhibit important quantitative differences. The relevant bound on the Na chemical potential is imposed by the compound Na_2Se. The most important result is that the solubility of substitutional Na is significantly lower than that of Li — the maximum concentration obtained from the contour plot is lower than 10^{16} cm^{-3}. At these lower concentrations, very few Na interstitials are present; we also find that the barrier for migration of the Na interstitial is much higher than for Li_i. Experimental doping attempts with Na have been unsuccessful;[10] our results clearly show that the solubility limit is the culprit, rather than, e.g., compensation due to foreign impurities in the source.

C. Nitrogen

Finally, we discuss N in ZnSe. Nitrogen on a substitutional Se site (N_{Se}) is a shallow acceptor. The surround-

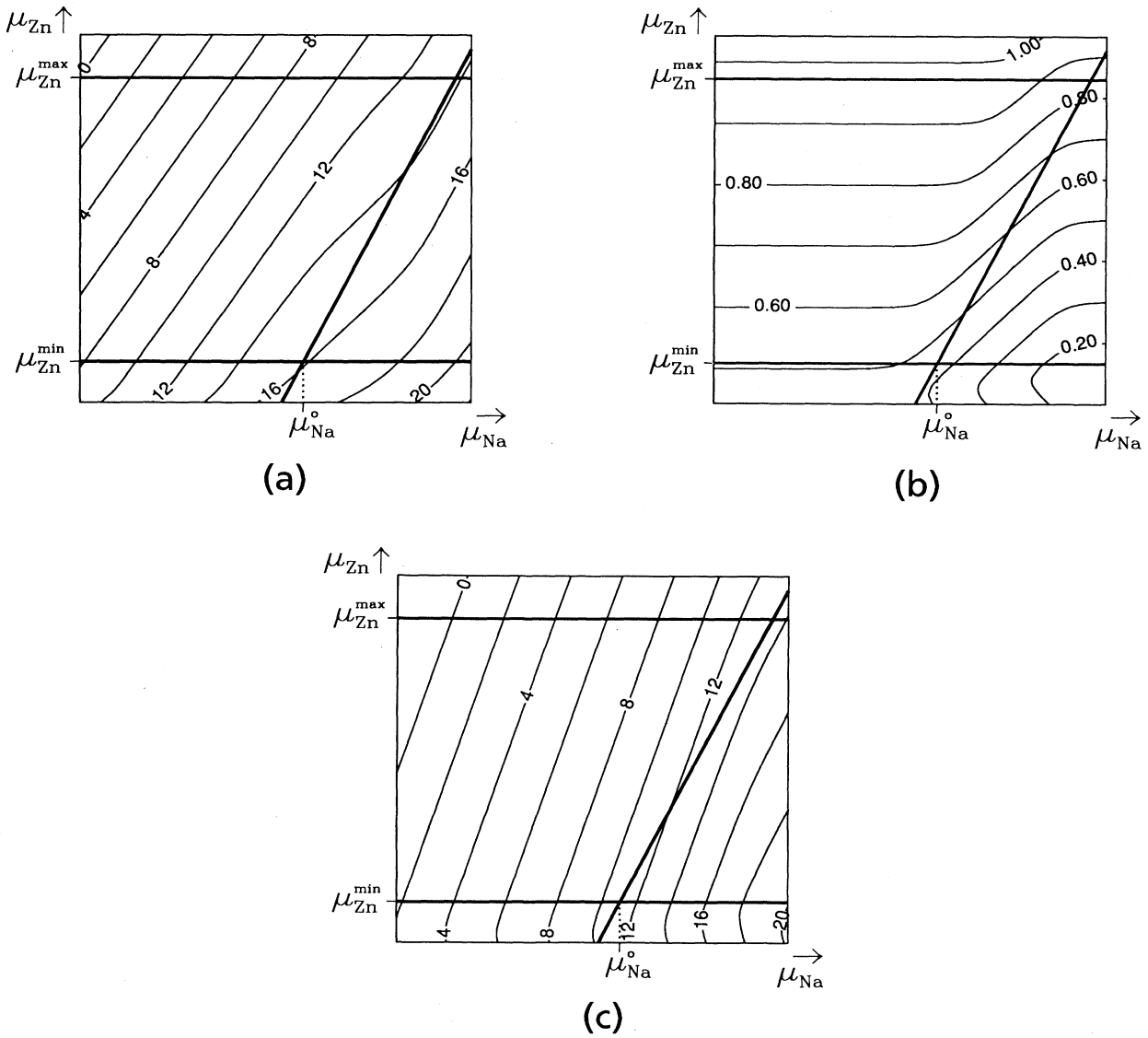

FIG. 3. Contour plots of (a) $\log_{10}$ [Na], where [Na] is the total Na concentration in cm^{-3}, (b) Fermi level (in eV, referred to the top of the valence band), and (c) $\log_{10}$ [Na_i], where [Na_i] is the interstitial Na concentration in cm^{-3}, at 600 K in ZnSe:Na, as a function of Zn and Na chemical potentials. Solid lines indicate bounds on μ_{Zn} and μ_{Na}.

ing Zn atoms undergo a significant inward relaxation, reducing the Zn-N distance to 2.1Å. This distance is very close to the Zn-N distance in the compound Zn_3N_2.[28] We have also investigated other configurations, such as the substitutional Zn site and interstitial sites, and found those to be much higher in energy than the substitutional Se site. Thus, N does not suffer from the substitutional/interstitial competition associated with the column-I elements, so that the saturation of the Fermi level which we observed in Fig. 1(b) does not occur here. In this work, we have not investigated any relaxation of the impurity away from the ideal lattice site.[29] According to Ref. 29, in the case of N such relaxations would not interfere with the shallow acceptor character of the dopant; if any relaxations do occur, they would therefore simply lead to a lower formation energy (and hence enhanced concentration) of the shallow acceptor state.

Two bounds on the N chemical potential arise in this case: N_2 molecules and the Zn_3N_2 compound. The compound Zn_3N_2 has the bixbyite structure,[28] which contains 80 atoms in the unit cell. This exceeds the capabilities of state-of-the-art first-principles calculations; we have therefore resorted to calculating a higher-symmetry structure, whose energy closely approximates that of the real compound. With regard to the other bound, our application of N_2 molecules as a solubility-limiting phase does not imply that we assume equilibrium between ZnSe:N and N_2 gas outside. Rather, we envision formation of some condensed phase involving N_2, such as in a void or in a chemisorbed state. Because of the difficulty in obtaining converged results for the N_2 molecule with an acceptable basis set, the energy difference between N_2 and Zn_3N_2 was taken from experiment.

Our results are displayed in Fig. 4. The bending of the contour lines in the upper part of Fig. 4(b) is due to native defects. Indeed, the N concentration is very low here [less than 10^{12} cm^{-3}; see Fig. 4(a)], and a small concentration of native defects suffices to pin the Fermi level. However, native defects play only a minor role if the right conditions (chemical potentials) are present for high N dopant concentrations. At the point of highest N incorporation, the calculated N concentration is 6.4×10^{19} cm^{-3}.

The native defect concentration once again increases as we approach the lower end of the accessible region (low μ_{Se}, i.e., Zn-rich conditions). The dominant native defect is the Zn interstitial; its concentration as a function of chemical potentials is shown in Fig. 5. The compensation due to this native defect is still small enough not to pose any threat to the doping. We have verified that this conclusion remains true even if our calculated formation energy for the native defect would be off by several 0.1 eV. The reason the results are not very sensitive to such inaccuracies is the "negative feedback" mechanism discussed in Sec. II C, acting through the coupling of all defect and impurity concentrations via the charge neutrality condition. In addition, the Zn_i concentration falls off rapidly (faster than the N concentration) as the Se chemical potential is raised, away from its lower bound. Other native defects have concentrations four orders of magnitude smaller than the Zn interstitial. Although our calculations indicate Zn interstitials should be present in N-doped samples in concentrations high enough for experimental observation, other factors have to be taken into account. One such factor is the high mobility of the Zn interstitial,[30] which may cause it to move into the substrate or towards the surface. It is also conceivable that Zn interstitials (donors) would form complexes with substitutional N acceptors.

Once again, we have investigated the effect of temperature on our results. Lowering the temperature from 600 to 500 K decreases the total N concentration by a factor of 4; simultaneously, the concentration of Zn interstitials drops by a factor of 20.

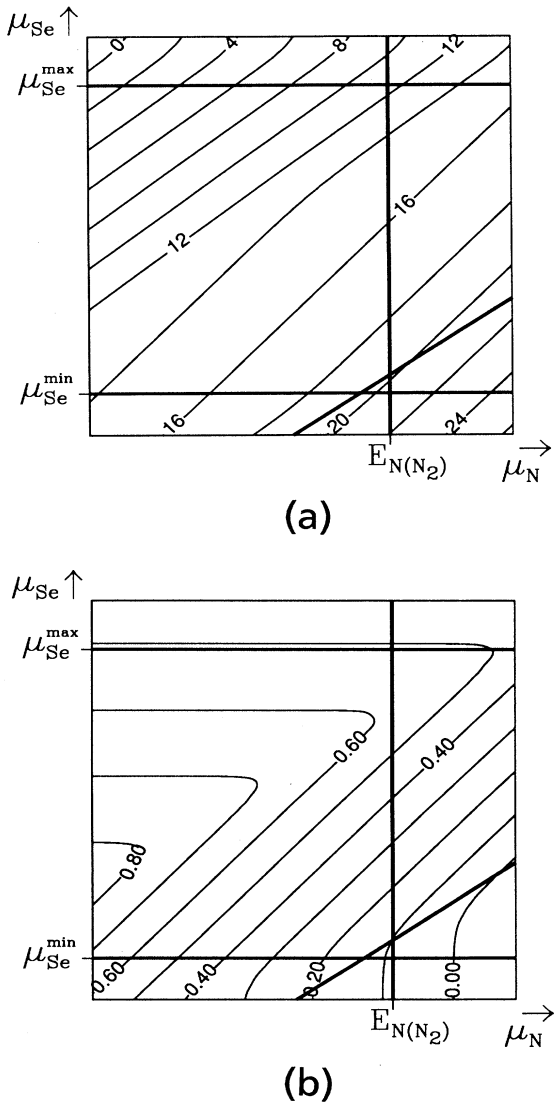

FIG. 4. Contour plots of (a) $\log_{10}$ [N], where [N] is the total N concentration in cm^{-3}, and (b) Fermi level (in eV, referred to the top of the valence band) at 600 K in ZnSe:N. Since N is substitutional on a Se site, μ_{Se} (rather than μ_{Zn}) is chosen as the variable here. Solid lines indicate bounds on μ_{Se} and μ_N.

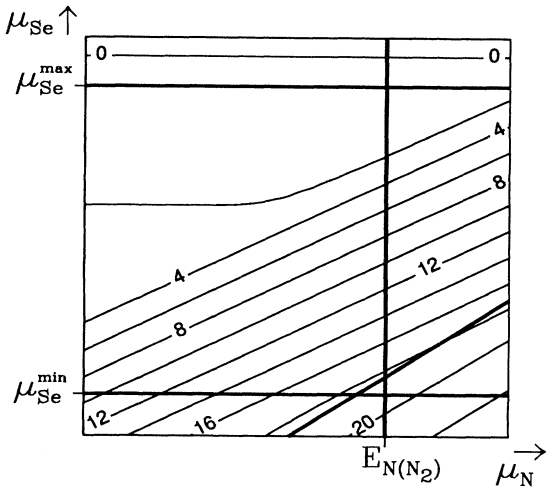

FIG. 5. Contour plot of $\log_{10}[Zn_i^{2+}]$, the Zn interstitial concentration in cm^{-3}, at 600 K in ZnSe:N, as a function of Se and Li chemical potentials. Solid lines indicate bounds on μ_{Se} and μ_N.

A comparison with Fig. 1 show that N has a solubility significantly higher than Li, which is consistent with experimental results. The failure of nitrogen doping starting from N_2 is due to the large kinetic barrier for breaking up the molecule; a plasma source or other technique for obtaining N in an atomic state, or at least N_2 in an excited state, is required.[5] Once one succeeds in incorporating atomic (as opposed to molecular) nitrogen into the lattice, N should act as a good acceptor, allowing hole concentrations high enough for useful device applications.

IV. SUMMARY

We have presented a formalism that enables us to calculate impurity concentrations and doping levels in semiconductors. The technologically important case of acceptor doping in ZnSe was discussed in detail; however, the formalism is quite general in nature and can be applied to any semiconductor and any impurity for which reliable first-principles calculations can be carried out. The computed total energies of impurities and defects allow us to write down formation energies as a function of the atomic chemical potentials and of the Fermi level; the latter is then determined by imposing charge neutrality.

The results are presented in the form of contour plots, which reflect the dependence on chemical potentials. Although the latter are free parameters, which vary with the growth conditions, they are subject to thermodynamic bounds corresponding to formation of other phases (e.g., formation of Li_2Se in the case of ZnSe:Li). Imposing these bounds determines the maximum achievable impurity incorporation. In addition, our results provide insight in how variations in growth conditions can promote incorporation of the dopant in the desirable configuration.

For acceptors in ZnSe, we have reached the following conclusions: Although Li suffers from a competition between interstitial and substitutional configurations, appropriate growth conditions can be chosen to suppress interstitial formation. The limited solubility of Li (imposed by formation of Li_2Se) is a more severe obstacle to the success of Li as a p-type dopant. Sodium suffers from this type of solubility problem to an even greater extent. Nitrogen, finally, emerges as the best choice among the dopants examined here.

ACKNOWLEDGMENTS

Part of this work was carried out while one of us (C.G.V.d.W.) was at Philips Laboratories, Briarcliff Manor, NY. We are grateful to D. A. Cammack, J. Gaines, F. Greidanus, T. Marshall, R. M. Martin, and J. Tersoff for suggestions and discussions. We are indebted to D. Vanderbilt for his iterative diagonalization program. This work was supported in part by ONR Contract No. N00014-84-0396.

*Present address: IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598.

[1]R. N. Bhargava, J. Cryst. Growth **86**, 873 (1988).
[2]M. A. Haase, H. Cheng, J. M. DePuydt, and J. E. Potts, J. Appl. Phys. **67**, 448 (1990).
[3]J. Ren, K. A. Bowers, B. Sneed, D. L. Dreifus, J. W. Cook, Jr., J. F. Schetzina, and R. M. Kolbas, Appl. Phys. Lett. **57**, 1901 (1990).
[4]T. Marshall and D. A. Cammack, J. Appl. Phys. **69**, 4149 (1991).
[5]R. M. Park, M. B. Troffer, C. M. Rouleau, J. M. DePuydt, and M. A. Haase, Appl. Phys. Lett. **57**, 2127 (1990).
[6]M. A. Haase, J. Qiu, J. M. DePuydt, and H. Cheng, Appl. Phys. Lett. **59**, 1272 (1991).
[7]G. F. Neumark, J. Appl. Phys. **51**, 3383 (1980).
[8]T. Sasaki, T. Oguchi, and H. Katayama-Yoshida, Phys. Rev. B **43**, 9362 (1991).
[9]Recently, a similar phenomenon has been experimentally observed in Zn-doped InP, where the solubility is limited by formation of Zn_3P_2; L. Y. Chan, K. M. Yu, M. Ben-Tzur, E. E. Haller, J. M. Jaklevic, W. Walukiewicz, and C. M. Hanson, J. Appl. Phys. **69**, 2998 (1991).
[10]H. Cheng, J. M. DePuydt, J. E. Potts, and M. A. Haase, J. Cryst. Growth **95**, 512 (1989).
[11]G. Mandel, Phys. Rev. **134**, A1073 (1964).
[12]D. B. Laks, C. G. Van de Walle, G. F. Neumark, and S. T. Pantelides, Phys. Rev. Lett. **66**, 648 (1991).
[13]P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, *ibid.* **140**, A1133 (1965).
[14]D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).
[15]S. G. Louie, S. Froyen, and M. L. Cohen, Phys. Rev. B **26**, 1738 (1982).
[16]D. B. Laks, C. G. Van de Walle, G. F. Neumark, and S. T. Pantelides, Phys. Rev. B **45**, 10965 (1992).
[17]C. G. Van de Walle, P. J. H. Denteneer, Y. Bar-Yam, and

S. T. Pantelides, Phys. Rev. B **39**, 10 791 (1989).
[18]Activation energies: Li, 114 meV [J. L. Merz, K. Nassau, and J. W. Shiever, Phys. Rev. B **8**, 1444 (1973)]; Na, 128 meV [H. Tews, H. Venghaus, and P. J. Dean, *ibid.* **19**, 5178 (1979)]; N, 110 meV [K. Shahzad, B. A. Khan, D. J. Olego, and D. A. Cammack, *ibid.* **42**, 11 240 (1990)].
[19]G.-X. Qian, R. M. Martin, and D. J. Chadi, Phys. Rev. B **38**, 7649 (1988); N. Chetty and R. M. Martin, *ibid.* **45**, 6089 (1992).
[20]F. A. Kröger, *The Chemistry of Imperfect Crystals* (North-Holland, Amsterdam, 1964), pp. 136 and 628.
[21]S. B. Zhang and J. E. Northrup, Phys. Rev. Lett. **67**, 2339 (1991).
[22]M. A. Haase, J. M. DePuydt, H. Cheng, and J. E. Potts, Appl. Phys. Lett. **58**, 1173 (1991).
[23]T. Marshall, in *Proceedings of the 7th Trieste Semiconductor Symposium on Wide-Band-Gap Semiconductors*, edited by C. G. Van de Walle (North-Holland, Amsterdam, in press).
[24]Z. Zhu, H. Mori, M. Kawashima, and T. Yao, J. Cryst. Growth **117**, 400 (1992).
[25]*Lange's Handbook of Chemistry*, 12th ed., edited by J. A. Dean (McGraw-Hill, New York, 1979).
[26]G. F. Neumark and C. R. A. Catlow, J. Phys. C **17**, 6087 (1984).
[27]C. G. Van de Walle (unpublished).
[28]R. W. G. Wyckoff, *Crystal Structures*, 2nd ed. (Interscience, New York, 1964), Vol. 2.
[29]D. J. Chadi and K. J. Chang, Appl. Phys. Lett. **55**, 575 (1989); D. J. Chadi, *ibid.* **59**, 3589 (1991).
[30]G. D. Watkins, in *Defect Control in Semiconductors*, edited by K. Sumino (Elsevier Science, Amsterdam, 1990), p. 933.