Role of Native Defects in Wide-Band-Gap Semiconductors

D. B. Laks,[(1),(2),(a)] C. G. Van de Walle,[(3)] G. F. Neumark,[(1)] and S. T. Pantelides[(2)]
[(1)]*Columbia University, New York, New York 10027*
[(2)]*IBM T. J. Watson Research Center, Yorktown Heights, New York 10598*
[(3)]*Philips Laboratories, Briarcliff Manor, New York 10510*
(Received 4 October 1990)

Wide-band-gap semiconductors typically can be doped either *n*-type or *p*-type, but not both. Compensation by native defects has often been invoked as the source of this difficulty. Using first-principles total-energy calculations we show that, for ZnSe and diamond, native-defect concentrations are too low to cause compensation. For nonstoichiometric ZnSe, native defects compensate *both* *n*-type and *p*-type material; thus deviations from stoichiometry cannot explain why ZnSe can be doped only one way. In the absence of a generic mechanism, specific dopants should be examined case by case.

PACS numbers: 71.55.Gs, 61.70.Bv, 72.20.Jv

Wide-band-gap semiconductors (such as ZnSe, ZnS, CdS, ZnTe, BN, or diamond) have ideal band gaps for optical applications using blue or green light, including semiconductor lasers and light-emitting diodes. There is, however, a fundamental problem with these materials: It is difficult, if not impossible, to make diamond and ZnTe *n*-type, and to make the rest *p*-type.[1-3] The simplest explanation[1,4-7] suggested for this phenomenon is that native defects compensate, say, acceptors in ZnSe. Because of the wide band gap, some of the energy needed to form a native donor defect can be recouped when electrons from defect levels in the gap recombine with holes at the Fermi level in *p*-type material. The spontaneous formation of native defects would thus prevent the Fermi level from moving below a fixed value that is determined by the formation energies and electronic levels of the native defects, independent of the dopant and of how the material was prepared. This picture has some very appealing features. It would explain why doping problems occur in all wide-band-gap materials, and are less severe in medium-gap materials such as CdTe. It would also explain why the difficulty in producing *p*-type (or *n*-type) material is universal, appearing for all growth and doping techniques, and for all dopants. That these materials can be doped *n*-type and not *p*-type, or vice versa, can be explained if the native defects with the lowest formation energy are donors in some materials and acceptors in others. For example, Jansen and Sankey[7] have suggested that the native-defect mechanism can account for the difference between ZnSe, which can be made *n*-type, and ZnTe, which can be made *p*-type, even though the two materials are strikingly similar in other ways. There is, however, no direct evidence to either confirm or deny the role of native defects in wide-band-gap semiconductors.

In this Letter we report on theoretical determinations of native-defect concentrations in ZnSe. The underlying calculations attain for the first time the level of accuracy that has so far been practical only for materials like Si and GaAs. We find that (1) the native-defect concentrations are too low to be a significant source of compensation in stoichiometric ZnSe; (2) undetectably small deviations from stoichiometry can produce large concentrations of native defects. We find that the defects formed depend on whether the sample is *n*-type or *p*-type, but always compensate. Hence deviations from stoichiometry cannot explain why ZnSe can be doped *n*-type but not *p*-type, because they are as likely to compensate *n*-type material as *p*-type. We have further determined native-defect concentrations in diamond, and find again that compensation by native defects is insignificant. In the absence of a generic mechanism, potential dopants need to be examined case by case.

Our determination of defect concentrations is based on calculating the total energy of each defect using density-functional theory (DFT) and the local-density approximation (LDA), norm-conserving pseudopotentials and supercells.[8] These techniques have been very successful in elucidating defect properties in Si (Refs. 9 and 10) and GaAs.[11] Applying the same tools to ZnSe, however, presents a problem. Zinc contains a fully occupied band of 3*d* electrons, which are tightly bound to the nucleus, and yet fall within the valence bands of ZnSe. Standard defect calculations are performed with a plane-wave basis set, which would require far too many plane waves to represent the *d* states. If the *d* states are treated as core states of the pseudopotential, it is not necessary to represent them in the basis set; unfortunately, this procedure is unacceptable because it does not correctly represent the properties of ZnSe.[12] To treat the *d* states properly, and still be able to perform supercell calculations, we use an all-new mixed-basis-set program, similar in spirit to that of Louie, Ho, and Cohen.[13,14] The usual plane-wave (PW) basis set is supplemented by a set of pseudoatomic tight-binding (TB) functions situated on each zinc atom. Calculations of the bulk properties of ZnSe (and other semiconductors) show that this scheme describes material properties very well: The predicted lattice constant and bulk modulus agree with experiment to within 1% and 10%, respectively.

For defect calculations, the convergence of the results with respect to basis sets and supercell size was checked to ensure overall accuracy of better than 0.5 eV.[15] An additional uncertainty is introduced by the local-density approximation, which is well known to underestimate band gaps. In p-type material, this uncertainty is negligible because all levels in the energy gap are empty. The uncertainty in n-type material is larger and can be estimated from the error in the band gap itself. In our discussion of the results for n-type materials, we assumed the worst-case values.

Calculations were performed for all native point defects: Zn_i, Se_i (interstitials), V_{Zn}, V_{Se} (vacancies), Zn_{Se}, and Se_{Zn} (antisites) in a variety of charge states; 29 different cases were examined, and detailed results will be published elsewhere. Calculations for these native defects have been reported earlier by Jansen and Sankey,[7] using more approximate techniques.[16] We will refer to their results where appropriate.

For a compound semiconductor like ZnSe, the formation energies and hence the concentrations of native defects are a function of the stoichiometry of the material. The stoichiometry itself is related to the chemical potentials of the constituents of the compound, in our case Zn (μ_{Zn}) and Se (μ_{Se}) atoms. The two chemical potentials are constrained by the condition that (in equilibrium) their sum must equal the total energy of a two-atom unit of perfect ZnSe ($\mu_{ZnSe}=\mu_{Zn}+\mu_{Se}$). (We use the total energy of a perfect ZnSe cell at $T=0$ K for μ_{ZnSe}.) Given the Zn and Se chemical potentials, the formation energy of each native defect is well defined and can be derived from a supercell calculation as follows. The total energy of a supercell for the ith defect containing N Zn atoms and M Se atoms (E_i) is calculated. The defect formation energy is then

$$E_i - N\mu_{Zn} - M\mu_{Se} = E_i - (N-M)\mu_1 - (N+M)\mu_2 = \varepsilon_i - n_i\mu_1 ,$$

where $\mu_1=(\mu_{Zn}-\mu_{Se})/2$, $\mu_2=(\mu_{Zn}+\mu_{Se})/2$ (a constant), $n_i=N-M$, and $\varepsilon_i=E_i-(N+M)\mu_2$. n_i is the number of extra Zn atoms that must be added to form the defect (+1 for V_{Se}, −2 for Se_{Zn}, etc.), independent of the size of the supercell. Using this prescription, all of the defect formation energies, and hence their concentrations (C_i), are unique functions of μ_1. The concentrations, in turn, determine the stoichiometry. In practice, however, it is more convenient to fix the stoichiometry first, and then determine C_i. To do this we write C_i in terms of the total energies and entropies (S) of formation as

$$C_i = e^{S/k_B} e^{-(\varepsilon_i - n_i\mu_1)/k_BT} = e^{S/k_B - \varepsilon_i/k_BT} y^{n_i} = a_i y^{n_i} ,$$

where $y=\exp(\mu_1/k_BT)$. The stoichiometry parameter is

$$X = -\tfrac{1}{2}\sum_i n_i C_i = -\tfrac{1}{2}\sum_i n_i a_i y^{n_i}$$

($X=0$ for perfect stoichiometry, and $X>0$ for Se rich). To find defect concentrations as a function of stoichiometry, one simply chooses a value of X (and the temperature) and solves for y. (The problem is essentially finding a root of a polynomial, which can be done quickly and easily using standard algorithms.)

Defect concentrations are a function of formation energies and entropies. We have checked that our results are insensitive to the value of the entropies in the range $S=(0\text{–}10)k_B$. By comparison, a recent accurate calculation[17] of the formation entropy of the Si self-interstitial found a formation entropy of $(5\text{–}6)k_B$ for the ground state. The Si self-interstitial represents an extreme case in that the ground-state configuration has low symmetry, which accounts for half of the formation entropy. It is therefore highly unlikely that the entropies for native defects in ZnSe or diamond could be larger than $10k_B$. Similarly, the defect formation energies are high enough that, even with a generous estimate of the atomic relaxation energies, the concentrations remain very low. Relaxations are calculated explicitly for the dominant defects in p-type ZnSe and found to be less than 0.6 eV. The concentrations of other defects remain small even if relaxations up to 2 eV are assumed.

Figure 1 shows the concentrations of minority carriers produced by native defects for p-type stoichiometric ZnSe. The results shown are for material with 10^{18}-cm^{-3} dopants. The dominant native defects are Zn_i^{2+}, V_{Zn}^{0}, and Se_{Zn}^{2+}. At molecular-beam-epitaxy- (MBE-) growth temperatures ($T=600$ K) the concentration of minority carriers produced is less than 10^{12} cm^{-3}. For material grown at higher temperatures, excess native defects will recombine during cooling, unless the sample is rapidly quenched.[18]

We have also calculated native-defect concentrations

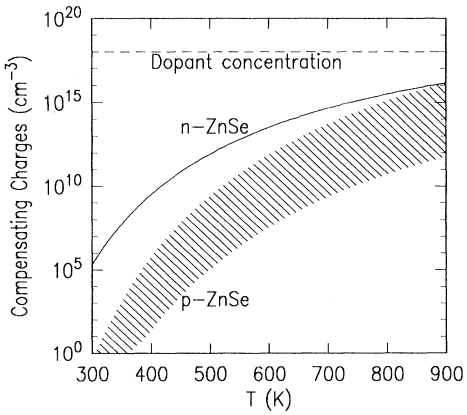

FIG. 1. Concentration of minority carriers produced by all native defects in stoichiometric p-type and n-type ZnSe. (The range of values shown for p-type ZnSe is bounded by assuming an entropy of $10k_B$ per defect for an upper bound and $0k_B$ for a lower bound.)

in *n*-type ZnSe (Fig. 1). The dominant defects are V_{Zn}^{2-} and Zn_{Se}^{-}. Well-conducting *n*-type ZnSe can be easily produced; thus it is an experimental fact that native defects do not compensate *n*-type doping. As shown in Fig. 1, native-defect concentrations in *n*-type ZnSe are comparable to, if not greater than, defect concentrations in *p*-type.[19] This is additional proof that native-defect compensation cannot explain why *p*-type ZnSe is harder to grow than *n*-type.

To further support our conclusions, we have derived native-defect concentrations for diamond from the first-principles defect energies of Bernholc *et al.*[20] The doping level is again 10^{18} cm^{-3}. At a chemical-vapor-deposition-growth temperature of 1100 K, the number of holes produced in *n*-type diamond by native defects is at most 2×10^{13} cm^{-3} (Fig. 2). Clearly, the concentrations of native defects in both stoichiometric ZnSe and diamond are far too low to produce significant compensation.

Jansen and Sankey have estimated native-defect concentrations in ZnSe and ZnTe.[7] They concluded that native-defect compensation could explain why ZnSe prefers to be *n*-type and ZnTe prefers to be *p*-type. However, their results were reported for a very high temperature ($T=1658$ K), and thus do not apply to the question of compensation for material that is grown at 600 K and never thermally annealed at higher temperatures. At lower temperatures, their results also show a low concentration of native defects. Furthermore, for ZnSe their numbers indicate that compensating native-defect concentrations are *lower* in *p*-type material than in *n*-type.

Our conclusion that the concentrations of native defects in stoichiometric ZnSe are very low does not mean

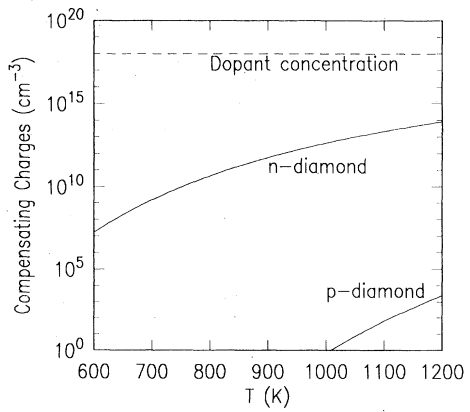

FIG. 2. Concentration of minority carriers produced by all native defects in *n*-type diamond. (The results for *n*-type diamond use the worst-case correction for the LDA band-gap error: The conduction-band edge is taken at the experimental value without shifting up any of the defect levels. True defect concentrations are probably much lower.)

that native-defect compensation in ZnSe never occurs. If the sample is grown with even a slight deviation from perfect stoichiometry, the concentration of native defects will necessarily be very large, even at $T=0$ K.[21] Because the density of atomic sites in ZnSe is 4×10^{22} cm^{-3} a deviation from stoichiometry as small as 10^{-4} implies a defect concentration of about 10^{18} cm^{-3}. We find that the native defects that accommodate deviations from stoichiometry are always those that compensate the majority carriers. For *p*-type ZnSe, the dominant defect is Zn_i in Zn-rich material, and Se_{Zn} in Se-rich material; we find that both are double donors. For *n*-type ZnSe, the dominant (acceptor) defects are Zn_{Se} and V_{Zn} for Zn and Se rich materials, respectively. Similar results were found by Jansen and Sankey.[7] This defect structure is much richer than that used in many previous analyses of native defects in II-VI semiconductors.[5] The difficulty in producing *p*-type ZnSe cannot be explained by deviations from stoichiometry because any deviation that compensates *p*-type doping would compensate *n*-type doping equally well.

The deviations from stoichiometry that we are discussing are too small to measure experimentally, which precludes a *direct* confirmation of our predictions. There is, however, indirect evidence to verify one of our predictions, namely, that the zinc vacancy is the dominant native defect in *n*-type Se-rich ZnSe. As-grown bulk ZnSe samples are highly compensated, and must be annealed in a Zn-rich atmosphere to be made well conducting. One known cause of this compensation is large numbers of "self-activated" (acceptor) centers, which are donor-V_{Zn} pairs.[22] This shows that zinc vacancies are a prominent defect in as-grown *n*-type ZnSe. Furthermore, analysis of the Zn-Se phase diagram suggests that ZnSe grown under equilibrium conditions from a melt is Se rich. Thus, our results for Se-rich *n*-type ZnSe provide a natural explanation of the occurrence of self-activated centers in ZnSe.

Having settled the native-defect compensation issue quantitatively, we now reexamine the notion that native-defect compensation increases with the width of the band gap. Let us restate the standard argument for this trend: For *p*-type material, imagine a prototypal compensating native donor defect that, when neutral, introduces one electron into a state in the gap; the formation energy for this defect, E^0, is assumed not to depend on the width of the band gap. The energy gained by transferring the electron from the level in the gap (E_L) to the Fermi level (E_F) should, in contrast, increase with the width of the gap; thus the net energy needed to form compensating defects, $E^0-(E_L-E_F)$, should decrease as the band gap increases. The flaw in this argument is that it assumes that the level in the gap (E_L) and E^0 are independent of one another. Actually, the level in the gap is *defined* by $E_L=E^0-E^+$, where E^++E_F is the (Fermi-level-dependent) energy of formation of the posi-

tive charge-state defect. Using this definition, we find that the net energy required to create a compensating defect is $E^0-(E_L-E_F)=E^++E_F$, independent of the energy of formation of the neutral defect. We see that native-defect compensation will increase with the width of the band gap if and only if E^++E_F *decreases* with increasing band gap. The existence of such a trend has not been convincingly established.

Having eliminated native defects as a generic source of compensation in wide-band-gap materials, it is fruitful to identify problems associated with specific dopants. We are studying the technologically important case of Li_{Zn}, a promising acceptor in ZnSe.[23] In a separate publication, we will report on the properties of Li impurities in ZnSe, including possible defect reactions.

In conclusion, we have shown that native defects alone cannot be responsible for difficulties in doping the wide-band-gap semiconductors ZnSe and diamond. Native-defect concentrations in MBE-grown stoichiometric ZnSe are too low to compensate. Deviations from stoichiometry in ZnSe *do* produce large numbers of native defects which, however, compensate *n*-type as well as *p*-type material.

We are very grateful to P. Blöchl for many fruitful suggestions, and for making his unpublished work available to us. We are indebted to D. Vanderbilt for his iterative diagonalization program. We acknowledge helpful conversations with R. Bhargava, J. M. DePuydt, T. Marshall, J. Tersoff, and G. D. Watkins. D.B.L. acknowledges support from an IBM Graduate Fellowship. This work was supported in part by NSF Grant No. ECS-89-21159 and ONR Contract No. N00014-84-0396.

---

[a]Present address: Solar Energy Research Institute, Golden, CO 80401.

[1]Y. S. Park and B. K. Shin, in *Electroluminescence,* edited by J. I. Pankove, Topics in Applied Physics Vol. 17 (Springer, Berlin, 1977), p. 133.

[2]R. Bhargava, J. Cryst. Growth **59**, 15 (1982).

[3]G. F. Neumark, Phys. Rev. Lett. **62**, 1800 (1989).

[4]G. Mandel, Phys. Rev. **134**, A1073 (1964).

[5]A. K. Ray and F. A. Kröger, J. Electrochem. Soc. **125**, 1348 (1978).

[6]Y. Marfaing, Prog. Crystal Growth Charact. **4**, 317 (1981).

[7]R. W. Jansen and O. F. Sankey, Phys. Rev. B **39**, 3192 (1989).

[8]C. G. Van de Walle, P. J. Denteneer, Y. Bar-Yam, and S. T. Pantelides, Phys. Rev. B **39**, 10791 (1989), and references therein.

[9]R. Car, P. J. Kelly, A. Oshiyama, and S. T. Pantelides, Phys. Rev. Lett. **52**, 1814 (1984).

[10]Y. Bar-Yam and J. D. Joannopoulos, J. Electron. Mater. **14A**, 261 (1985).

[11]G. A. Baraff and M. Schlüter, Phys. Rev. Lett. **55**, 1327 (1985).

[12]S. H. Wei and A. Zunger, Phys. Rev. B **37**, 8958 (1988).

[13]S. Louie, M. Ho, and M. Cohen, Phys. Rev. B **19**, 1774 (1979).

[14]The eigenvalue problem was solved using the iterative diagonalization scheme of R. Natarajan and D. Vanderbilt, J. Comput. Phys. **82**, 218 (1989).

[15]For supercells, 16-, 32-, and 64-atom cells were checked. In most cases, 32-atom cells were adequate. The plane-wave cutoff was typically 9 Ry.

[16]The *d* states were treated as part of the core.

[17]P. Blöchl (private communication).

[18]Defects in ZnSe move readily even at $T=400$ K, so that kinetic barriers do not prevent the attainment of thermal equilibrium. [G. Watkins, in Proceedings of the International Conference on Science and Technology of Defect Control in Semiconductors, Yokohama, 1989 (to be published)].

[19]Occupied defect levels in the gap were shifted upward by the LDA band-gap error. The true defect concentrations in *n*-type ZnSe may actually be higher than those shown in Fig. 1.

[20]J. Bernholc, A. Antonelli, T. M. Sole, Y. Bar-Yam, and S. T. Pantelides, Phys. Rev. Lett. **61**, 2689 (1988).

[21]We refer only to deviations from stoichiometry accommodated by native-point defects. Deviations from stoichiometry due to higher-dimensional defects, precipitates, and substitutional dopants are not included.

[22]R. K. Watts, *Point Defects in Crystals* (Wiley, New York, 1977), p. 252.

[23]J. Depuydt, M. Haase, H. Cheng, and J. Potts, Appl. Phys. Lett. **55**, 1103 (1989).