Mg acceptor energy levels in $Al_xIn_yGa_{1-x-y}N$ quaternary alloys: An approach to overcome the *p*-type doping bottleneck in nitrides

Fei Wang, Jingbo Li,* Shu-Shen Li, and Jian-Bai Xia
State Key Laboratory for Superlattices and Microstructures, Institute of Semiconductors, Chinese Academy of Sciences, P.O. Box 912, Beijing 100083, People's Republic of China

Su-Huai Wei
National Renewable Energy Laboratory, Golden, Colorado 80401, USA

(Received 25 November 2007; published 10 March 2008)

The Mg_{Ga} acceptor energy levels in GaN and random $Al_8In_4Ga_{20}N_{32}$ quaternary alloys are calculated using the first-principles band-structure method. We show that due to wave function localization, the Mg_{Ga} acceptor energy level in the alloy is significantly lower than that of GaN, although the two materials have nearly identical band gaps. Our study demonstrates that forming $Al_xIn_yGa_{1-x-y}N$ quaternary alloys can be a useful approach to lower acceptor ionization energy in the nitrides and thus provides an approach to overcome the *p*-type doping difficulty in the nitride system.

DOI: 10.1103/PhysRevB.77.113202 PACS number(s): 61.72.Bb

GaN and related nitride semiconductor alloys are currently among the most important materials for short-wavelength optoelectronic device applications, such as blue and ultraviolet light-emitting diodes and laser diodes, and for high-temperature and/or high-power electronic device applications.[1] This is because these nitride materials have many unique physical properties such as a large direct band gap and a high thermal conductivity.[2] Despite the tremendous progress in these fields in the last decade, further improvements in the performance of GaN-based devices are contingent on solving the doping asymmetry problem in the nitride system.[3,4] As in most wide-band-gap nitride and oxide materials, GaN can be easily doped *n* type but is difficult to dope *p* type. This is because nitrogen is strongly electronegative with a deep 2*p* atomic orbital; thus, the valence band maximum (VBM) of the nitrides, which contains mostly N 2*p* orbital, is low in energy. This leads to a relatively deep acceptor energy level and is liable for charge compensation.[5] Among all the dopants, magnesium is one of the best *p*-type dopants in GaN,[6] because its size is close to Ga, so when it substitutes on the Ga site, it has a small formation energy and causes only a relatively small perturbation on the VBM. Moreover, Mg has no occupied *d* orbital, so the Mg_{Ga} acceptor level is relatively shallow.[7] However, even in this case, the measured Mg_{Ga} (0/−) acceptor level in GaN is still too high, around ~0.20 eV.[2,3] Therefore, various approaches have been sought to lower the acceptor levels in GaN without significantly altering the band structure of the host material.

Recently, the quaternary nitride alloys $Al_xIn_yGa_{1-x-y}N$ have attracted much attention[8–10] because they could be good candidates to achieve easier *p*-type doping.[11] Unlike in the AlGaN, GaInN, and AlInN ternary alloys, the band gap, band edge energy level, and lattice constant of the quaternary alloy can be individually tuned. Therefore, one could find a quaternary $Al_xIn_yGa_{1-x-y}N$ alloy with the same band gap as in pure GaN but with a higher VBM, so it could be doped more easily *p*-type by lowering the acceptor ionization energy level and lowering the charge compensation.[12]

In this work, to test this idea, we have calculated the acceptor levels (Mg_{Ga}) in GaN and $Al_xIn_yGa_{1-x-y}N$ quaternary alloys using the first-principles band-structure method. The alloy in this study is represented using the special quasirandom structures approach,[13,14] in which the cations are distributed in a way that its atomic correlation functions are closest to that of the random alloy. We find that the calculated Mg_{Ga} acceptor transition energy level in random $Al_xIn_yGa_{1-x-y}N$ quaternary alloys with the same band gap as GaN is about 20–60 meV lower than that in GaN, depending on the location of the Mg atom in the alloy. Therefore, *p*-type doping in $Al_xIn_yGa_{1-x-y}N$ quaternary alloys indeed should be easier than that in GaN with the same band gap, consistent with our expectations.

The first-principles band-structure and total-energy calculations in this study are performed using the density functional theory in the local density approximation (LDA)[15] as implemented in the plane wave VASP code.[16] The electron and core interactions are included using the frozen-core projected augmented wave approach.[17,18] The Ga 3*d* and In 4*d* electrons are explicitly treated as valence electrons. An energy cutoff of 400 eV is used in all cases. For the Brillouin zone integration, we used the *k* points that are equivalent[19] to the $4\times4\times4$ Monkhorst-Pack special *k*-point meshes[20] in the zinc-blende (ZB) Brillouin zone. In all calculations, all the atoms are allowed to relax until the quantum mechanical forces acting on them become less than 0.01 eV/Å. For computation convenience, our calculations are performed for cubic ZB nitride alloys. Because the band structures of the ZB compounds and the wurtzite (WZ) compounds are very similar near the band edge at Γ, we expect that our results for the ZB alloy are also applicable for the ground state WZ alloys.[21–23]

The defect calculation is performed using the supercell approach. The ionization energy of an acceptor ($q<0$) with respect to the VBM in the impurity limit is calculated following the procedure described in Ref. 5,

 ©2008 The American Physical Society

TABLE I. Calculated formation energy $\Delta H_f(\alpha,0)$ at $\mu_i=0$ and the transition energy level $\varepsilon(0/-)$ for AlInGaN:Mg and GaN:Mg. The nNN denotes that there are n In atoms in the fcc nearest neighbor sites centered around the defect Mg atom in the $Al_8In_4Ga_{20}N_{32}$ quaternary alloys. The formation energy of the $q=-1$ charged Mg_{Ga} defect can be obtained using Eq. (2).

	$\Delta H_f(\alpha,0)$	$\varepsilon(0/-)$
4NN	1.282	0.184
3NN	1.240	0.163
2NN	1.234	0.153
1NN	1.206	0.147
0NN	1.201	0.139
GaN	1.272	0.198

$$\varepsilon(0/q) = \{E(\alpha,q) - [E(\alpha,0) - q\varepsilon_D^k(0)]\}/(-q) + [\varepsilon_D^\Gamma(0) - \varepsilon_{\mathrm{VBM}}^\Gamma(\mathrm{host})], \quad (1)$$

where $E(\alpha,q)$ and $E(\alpha,0)$ are the total energies of the supercell at charge state q and neutral, respectively, for defect α, $\varepsilon_D^k(0)$ and $\varepsilon_D^\Gamma(0)$ are the defect levels at the special k points (averaged) and at the Γ point, respectively, and $\varepsilon_{\mathrm{VBM}}^\Gamma(\mathrm{host})$ is the VBM energy of the host supercell at the Γ point and is aligned using core electron energy levels away from the defect. The first term on the right-hand side of Eq. (1) determines the U energy parameter (including both the Coulomb contribution and the atomic relaxation contribution) of the charged defects calculated at the special k points, which is the extra cost of energy after moving $(-q)$ charge from the VBM of the host to the neutral defect level. The second term gives the single-electron defect level at the Γ point.

The formation energy of a charged defect is given by

$$\Delta H_f(\alpha,q) = \Delta H_f(\alpha,0) - q\varepsilon(0/q) + qE_F, \quad (2)$$

where E_F is the Fermi level with respect to the VBM and $\Delta H_f(\alpha,0)$ is the formation energy of the charge-neutral defect and is a function of the atomic chemical potential.[5] The chemical potentials for $N(\mu_N)$, $Mg(\mu_{Mg})$, and $Ga(\mu_{Ga})$ are referenced to gaseous N_2, fcc Mg, and α-Ga. Our calculated value for $\Delta H_f(\mathrm{GaN})$ is −1.58 eV.

Table I presents the calculated Mg_{Ga} acceptor formation energies and (0/−) transition energy levels for GaN and the random $Al_8In_4Ga_{20}N_{32}$ quaternary alloys. We chose this alloy because our previous study[11] has shown that when Al and In have a ratio $x/y\sim 1.8$, the alloy band gap is close to that of GaN. Our direct LDA calculation shows that the band gap of the random $Al_8In_4Ga_{20}N_{32}$ quaternary alloys is only 0.01 eV smaller than that of GaN. For the $Al_8In_4Ga_{20}N_{32}$ alloy, we have calculated Mg_{Ga} at different sites. The nNN in Table I denotes that there are n In atoms in the fcc nearest neighbor (NN) sites centered around the Mg atom. Figure 1 shows the crystal structure of the random AlInGaN quaternary alloys doped with Mg_{Ga} in the 0NN site; i.e., there are no In atoms in the fcc nearest neighbor sites centered around the Mg atom. Our results show the following trend.

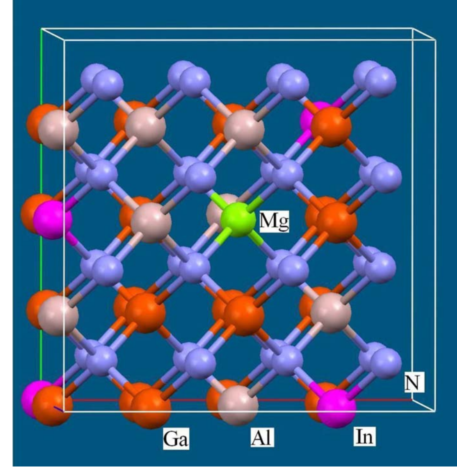

FIG. 1. (Color online) Crystal structure of the random AlInGaN quaternary alloys doped with Mg_{Ga}. (The gray, pink, red, blue, and green balls are Al, In, Ga, N, and Mg atoms, respectively.)

(i) The formation energy of the neutral Mg_{Ga} defect in the quaternary alloys decreases slightly when the neighboring sites contain decreasing number of In atoms. This is because the strain associated with the Mg_{Ga} substitution is slightly smaller in the (Ga,Al)-rich region than in the In-rich region. Moreover, as an acceptor defect, Mg_{Ga} also prefers to be at a low charge-density region to reduce the electron-electron Coulomb repulsion (see below). The formation energy of Mg_{Ga} in the GaN is slightly larger than that in the alloy. This is because the volume of GaN is about 0.8% smaller than that of the $Al_8In_4Ga_{20}N_{32}$ alloy, thus Mg_{Ga} in pure GaN experiences a slightly larger strain.

(ii) For Mg_{Ga} in GaN, the calculated Mg_{Ga} (0/−) acceptor level is at 0.198 eV. This agrees well with previous first-principles calculations[6] and experimental observations.[2] For the $Al_8In_4Ga_{20}N_{32}$ alloy, we find that the position of the Mg_{Ga} (0/−) acceptor level depends on the position of the Mg defect. It decreases from 0.184 eV when Mg has four In fcc NN atoms to 0.139 eV when there are no In atoms in the fcc NN. This indicates that for the $Al_8In_4Ga_{20}N_{32}$ alloy, which has nearly the same band gap as GaN, the acceptor ionization energy is about 20–60 meV lower than that in GaN. Combined with (i) above, which shows that Mg prefers to be at a 0NN position, we see that forming the random alloy is expected to enhance the p-type dopability significantly. The origin of the above-observed trends can be understood as follows: Due to the large coupling between the N $2p$ orbital and the In $4d$ orbital and the larger In-N bond length, the VBM of InN is 1.11 and 2.39 eV higher than that of GaN and AlN, respectively. Similarly, due to the large In-N bond length, the conduction band minimum (CBM) of InN is also much lower than that of GaN and AlN. Because of the unique band alignment, when the AlInGaN quaternary alloy is formed, the VBM as well as CBM wave function are more localized on the InN site. Figure 2 plots our calculated charge density of the alloy VBM and CBM states in different planes containing the Ga-N-In bond and Ga-N-Al bond. We can see from the plots that the VBM state [Figs. 2(a) and 2(b)] shows strong antibonding p-d character in the Ga-N

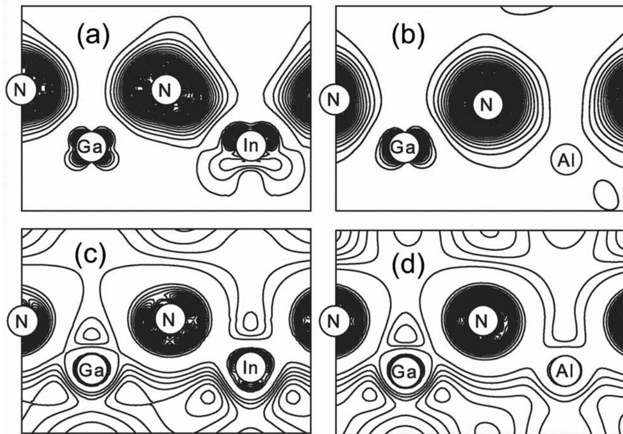

FIG. 2. Charge-density plot of [(a) and (b)] VBM and [(c) and (d)] CBM states in the planes of Ga-N-In bonds and Ga-N-Al bonds in the random AlInGaN quaternary alloys.

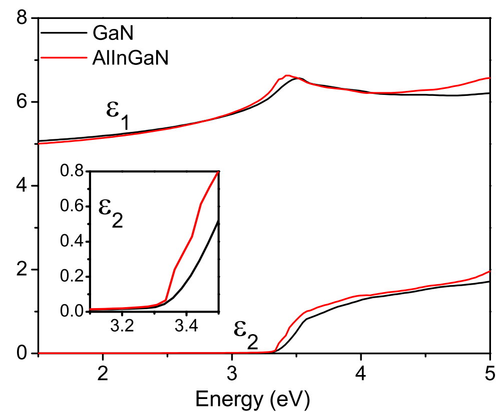

FIG. 3. (Color online) The real (ε_1) and imaginary (ε_2) components of the complex dielectric function for GaN and $Al_8In_4Ga_{20}N_{32}$. The inset shows the ε_2 near the band gap.

and In-N bonds and is more localized on the In-N site. Because Al has no active *d* orbital, the VBM contains only a negligible Al state. Similarly, the CBM states [Figs. 2(c) and 2(d)] also show some localization on the In-N site. The degree of the localization in the CBM is smaller than that in the VBM because the CBM band has smaller effective mass. We have shown that due to this wave function localization, the VBM of the $Al_8In_4Ga_{20}N_{32}$ alloy is higher than that of GaN.[11] Because of this, the Mg_{Ga} defect level is closer to the VBM, so the acceptor ionization energy is smaller. Also, because the potential for the VBM-derived defect level is high near the InN site, when the Mg atom moves away from the In site, the defect level decreases relative to the VBM, so the acceptor ionization energy also decreases.

It is important to point out that these wave function localization and reduction in the acceptor ionization energy are realized naturally when the distribution of cation atoms in the quaternary alloy is random. Therefore, it is not necessary to have In clustering as a precondition of the wave function localization, although In clustering is expected to enhance the effect of wave function localization.

Another consequence of the band edge wave function localization is the increased overlap of the VBM and CBM wave functions (Fig. 2) and, thus, enhanced optical transition intensity. To confirm this, we have calculated the real (ε_1) and imaginary (ε_2) components of the complex dielectric function for pure GaN and the $Al_8In_4Ga_{20}N_{32}$ alloy using the optical package[24] in the WIEN2K code.[25] The results are shown in Fig. 3. To compare directly with experiment, a rigid shift is made in the plot to correct the LDA band-gap error. We find that the calculated results are consistent with experimental results for the cubic GaN.[26] We also see that, near the band edge, the ε_2 for the quaternary alloys is larger than that for GaN, consistent with our expectation and experimental observations.[10]

In summary, we have calculated the Mg_{Ga} acceptor energy level in $Al_8In_4Ga_{20}N_{32}$ quaternary alloys and GaN using the first-principles band-structure method. The calculated Mg_{Ga} acceptor energy level is significantly lower (about 20–60 meV) than that of GaN, although the two materials have nearly identical band gaps, indicating that forming $Al_xIn_yGa_{1-x-y}N$ quaternary alloys can be a useful approach to lower acceptor ionization energy in the nitrides, thus improving their *p*-type dopability. We also show that because the wave function is more localized in the InN site, the optical transition near the band edge is enhanced in the alloy.

J.L. gratefully acknowledges financial support from the "One-Hundred Talents Plan" of the Chinese Academy of Sciences. This work was supported by the National Natural Science Foundation of China under Grants Nos. 60325416, 60521001, and 90301007. The work at NREL was funded by the (U.S.) Department of Energy under Contract No. DE-AC36-99GO10337.

*jbli@semi.ac.cn

[1] *Nitride Semiconductor and Devices*, edited by H. Morkoc (Springer, New York, 1999).
[2] I. Vurgaftman and J. R. Meyer, J. Appl. Phys. **94**, 3675 (2003).
[3] C. G. Van de Walle and J. Neugebauer, J. Appl. Phys. **95**, 3851 (2004).
[4] Y. Yan, J. Li, S.-H. Wei, and M. M. Al-Jassim, Phys. Rev. Lett. **98**, 135506 (2007).
[5] S.-H. Wei, Comput. Mater. Sci. **30**, 337 (2004).
[6] J. Neugebauer and C. G. Van de Walle, Appl. Phys. Lett. **68**, 1829 (1996).
[7] Be can have even lower acceptor transition energy due to its high electronegativity. However, due to its small size, it can also go easily to the interstitial sites, making it a double donor. See, C. G. Van de Walle, S. Limpijumnong, and J. Neugebauer, Phys. Rev. B **63**, 245205 (2001).

[8] M. Asif Khan, J. W. Yang, G. Simin, R. Gaska, M. S. Shur, H. Zur Loye, G. Tamulaitis, A. Zukauskas, D. J. Smith, D. Chandrasekhar, and R. Bicknell-Tassius, Appl. Phys. Lett. **76**, 1161 (2000).
[9] M. Marques, L. K. Teles, L. M. R. Scolfaro, J. R. Leite, J. Furthmuller, and F. Bechstedt, Appl. Phys. Lett. **83**, 890 (2003).
[10] C. H. Chen, L. Y. Huang, Y. F. Chen, H. X. Jiang, and J. Y. Lin, Appl. Phys. Lett. **80**, 1397 (2002).
[11] F. Wang, S. S. Li, J. B. Xia, H. X. Jiang, J. Y. Lin, J. Li, and S.-H. Wei, Appl. Phys. Lett. **91**, 061125 (2007).
[12] S. B. Zhang, S.-H. Wei, and A. Zunger, J. Appl. Phys. **83**, 3192 (1998).
[13] A. Zunger, S.-H. Wei, L. G. Ferreira, and J. E. Bernard, Phys. Rev. Lett. **65**, 353 (1990).
[14] S.-H. Wei, L. G. Ferreira, J. E. Bernard, and A. Zunger, Phys. Rev. B **42**, 9622 (1990).
[15] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[16] G. Kresse and J. Furthmuller, Comput. Mater. Sci. **6**, 15 (1996); Phys. Rev. B **54**, 11169 (1996).
[17] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
[18] P. E. Blochl, Phys. Rev. B **50**, 17953 (1994).
[19] L. E. Urbach, K. L. Percival, J. M. Hicks, E. W. Plummer, and H. L. Dai, Phys. Rev. B **45**, 3769 (1992).
[20] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1972).
[21] C.-Y. Yeh, S.-H. Wei, and A. Zunger, Phys. Rev. B **50**, 2715 (1994).
[22] S.-H. Wei and S. B. Zhang, Phys. Rev. B **62**, 6944 (2000).
[23] G. M. Dalpian and S.-H. Wei, Phys. Rev. Lett. **93**, 216401 (2004).
[24] R. Abt, C. Ambrosch-Draxl, and P. Knoll, Physica B **194**, 1451 (1994).
[25] K. Schwarz, P. Blaha, and G. K. H. Madsen, Comput. Phys. Commun. **147**, 71 (2002).
[26] M. Munoz, Y. S. Huang, F. H. Pollak, and H. Yang, J. Appl. Phys. **93**, 2549 (2003).