PHYSICAL REVIEW B **67**, 165209 (2003)

Breakdown of the band-gap-common-cation rule: The origin of the small band gap of InN

Su-Huai Wei, Xiliang Nie, Iskander G. Batyrev, and S. B. Zhang
National Renewable Energy Laboratory, Golden, Colorado 80401

(Received 20 September 2002; published 25 April 2003)

It is well accepted that the band gap of a semiconductor compound increases as the atomic number decreases. However, recent measurements of the small band gap of InN ($E_g \sim 0.9$ eV) suggest that this rule may not hold for the common-cation In compounds. Using a band-structure method that includes band-gap correction, we systematically study the chemical trends of the band-gap variation in III-V semiconductors. The calculated InN band gap is 0.85 ± 0.1 eV, much smaller than previous experimental value of ~ 1.9 eV. The InN band-gap anomaly is explained in terms of atomic-orbital energies and the band-gap deformation potentials.

DOI: 10.1103/PhysRevB.67.165209 PACS number(s): 71.20.Nr, 71.15.Mb

Conventional wisdom holds that for common-anion (or cation) semiconductors, the direct band gap at Γ increases as the cation (or anion) atomic number decreases [the band-gap-common-anion (or cation) rule]. This observation is strongly supported by experimental data.[1] For example, the direct band gaps of the common-anion compounds InAs, GaAs, and AlAs increase from 0.42 to 1.52 to 3.13 eV. Similarly, the direct band gaps of the common-cation zinc-blende (ZB) compounds GaSb, GaAs, GaP, and GaN increase from 0.81 to 1.52 to 2.86 to 3.30 eV. This trend also would hold for common-cation In compounds if $E_g(\mathrm{InN}) = 1.9$ eV, as previously reported.[2] However, recent measurements[3–6] of InN band gap found that $E_g(\mathrm{InN}) \sim 0.9$ eV. This value is not only about 1 eV smaller than the previously reported value, but is also smaller than the band gap[1] of InP with $E_g(\mathrm{InP}) = 1.46$ eV, contrary to what one would expect from the band-gap-common-cation rule.

Among the group-III nitrides (AlN, GaN, InN), InN is not very well investigated, mostly because high-quality bulk InN samples are difficult to grow.[7] Early experimental studies[2,8,9] suggested that the band gap of wurtzite (WZ) InN was around 2 eV. Due to the poor sample quality, however, no band-edge photoluminescence spectra were reported. The band gaps were estimated from the absorption spectra, which can lead to overestimating the band gap if the sample quality is poor or if the sample is highly doped.[10] Despite that, the value of 1.89 eV for the InN band gap obtained by Tansley and Foley[2] is widely cited in the literature.[11,12] It is often used as an end-point value to interpret experimentally measured composition dependence of the band gap of InN alloys,[13,14] or is used to fit empirical pseudopotentials for modeling InN alloy optical properties.[15,16] However, recent progress of growth techniques using molecular-beam epitaxy has led to improved InN samples, which show intense photoluminescence as well as a clear absorption edge.[3–6] These new measurements have challenged the previous widely accepted band gap value[2] and suggest that the actual fundamental band gap of InN is much smaller, between 0.7 and 1.1 eV.

Theoretical study of the band gap of InN is also quite confusing. It is well-known that the local-density approximation[17] (LDA), which is widely used in modern band-structure calculation, severely underestimates the band gap.[18,19] For example, the LDA-calculated band gap[20] of GaAs (~ 0.2 eV) is much smaller than the experimental value[1] of 1.52 eV. For InN in the WZ structure, the LDA-calculated band gap is about -0.3 eV.[21–24] This value is clearly much smaller than the true band gap of InN. Various methods have been used to correct the LDA band-gap error. Using the self-interaction and relaxation correction, Vogel *et al.*[24] find that the corrected band gaps are 1.3 eV for the ZB InN and 1.6 eV for the WZ InN. Using the screened-exchange approach, van Schilfgaarde and co-workers[22] find that the band gaps are 0.7 and 0.8 eV for ZB and WZ InN, respectively. However, a recent quasiparticle calculation by Kotani and van Schilfgaarde[23] using the GW approximation found that the band gap of ZB InN is only 0.01 eV. Model GW calculation of Johnson and Ashcroft gives a band gap of 1.79 eV for InN.[25] We have performed also the GW calculation using the ABINIT code.[26] We find that the band gap for WZ InN is 0.5 eV. Although there are strong indications from recent calculations that the true band gap of InN should be much smaller than the previously reported experimental value of 1.9 eV, the uncertainty of these calculations is still quite large.

In this paper, using an LDA-based semiempirical method,[20,27] we calculate the band gap of InN and *systematically* study the chemical trends of the band-gap variation in all III-V semiconductors. We find that the band gap of WZ InN is 0.85 ± 0.1 eV, in very good agreement with recent experimental measurements,[3–6] but much smaller than the previously accepted value[2] of 1.89 eV. We show that the reason that InN has a smaller band gap than InP is due to the combined effects of the much lower N $2s$ orbital energy and the much smaller band-gap deformation potential for the ionic InN. Furthermore, we show that the breakdown of the band-gap-common-cation rule is not unusual in ionic semiconductors with small band-gap deformation potentials.

The LDA band structure calculations in this study are performed using the fully relativistic (including spin-orbit coupling), general potential, linearized augmented plane-wave method.[28,20] Highly converged k-points sampling for the Brillouin-zone integration and cutoff energy for the basis function are used. The Ga $3d$ and In $4d$ states are treated as

TABLE I. Fitted parameters $\overline{V}$, V_0, and r_0 for group III and group V atoms. ES denotes empty sphere. $\overline{V}$ is nonzero only for ES. For nitrides, $R_{MT}(\mathrm{ES})=1.68$ a.u. For all other compounds, $R_{MT}(\mathrm{ES})=2.05$ a.u.

Atom	$\overline{V}$ (Ry)	V_0 (Ry)	r_0 (a.u.)
N, P, As, Sb	0.00	80	0.025
Al	0.00	360	0.025
Ga	0.00	280	0.025
In	0.00	200	0.025
ES	0.36	100	0.025

valence electrons. The band structures are calculated at experimental[1] lattice constants. To correct the LDA band gap error, we use a self-consistent approach with atom-dependent LDA corrections (LDA+C). Specifically, we add to the LDA calculations δ-function-like external potentials[20,29] inside the muffin-tin (MT) spheres centered at each atomic site α,

$$V_{ext}^{\alpha}(r)=\overline{V^{\alpha}}+V_0^{\alpha}\left(\frac{r_0^{\alpha}}{r}\right)e^{-(r/r_0^{\alpha})^2}, \tag{1}$$

and perform the calculation self-consistently. The functional form of the correction potential is based on the observation that the LDA band-gap error is orbital dependent. To correct the band-gap error one needs to have a potential that is more repulsive for the *s* orbital than to the *p* orbital. Since *p* orbital has zero charge density at the nuclear site while *s* orbital has finite density at the nuclear site, a δ-like function centered at the nuclear site can increase the band gap. The parameters $\overline{V^{\alpha}}$, V_0^{α}, and r_0^{α} in Eq. (1) are first fitted to the available experimental energy levels[1] and to the quasiparticle energies calculated by Zhu and Louie[18] at high-symmetry *k*-points for AlP, GaP, and InP.[20] In order to improve the fit, empty spheres centered at tetrahedral sites[29] are also used. The MT radii for the empty sphere are 2.05 a.u. The fitting parameters are given in Table I. We chose to have $\overline{V^{\alpha}}=0$ at all the atom sites so the potential in Eq. (1) becomes δ-function-like. This indicates that the calculated band gaps are not sensitive to the muffin-tin radii centered at the atomic sites. The same parameters given in Table I are then used to predict the band gaps of arsenides, antimonides, and nitrides. For the nitrides, however, we have to use smaller muffin-tin radii to avoid having overlapping MT spheres.[28] In this case, we use $R_{MT}=1.68$ a.u. for the empty spheres. This value is chosen to obtain the correct band gap of GaN. The same parameters given in Table I are used to predict the band gaps of AlN and InN. To find the band gap for the wurtzite structure, we add the LDA-calculated band-gap differences between the WZ and ZB compounds to the calculated LDA+C band gap for the ZB compound. It is well known that although the LDA underestimates the band gap, the LDA-calculated WZ-ZB band-gap differences are quite reliable.[21] The overall band-gap uncertainty associated with this fitting procedure is estimated to be 0.1 eV.

The predicted direct band gaps at the Γ point for the III-V semiconductors are shown in Table II. These values are compared with available experimental data[1] and quasiparticle calculations.[18,19] We find that for nearly *all* the III-V semiconductors, the differences between the predicted and the experimental band gaps are less than 0.1 eV. For InN, however, our predicted value of 0.85 eV is much smaller than the previous experimental value[2] of 1.9 eV, but it is in very good agreement with recent experimental measurements.[3–6]

Our calculations above show convincingly that the band gap of InN is around 0.85 eV. However, this value is about 0.6 eV smaller than that of InP, thus contradicting the conventional wisdom that the band gaps of common-cation (anion) compounds increase as the anion (cation) atomic number decreases. Table II shows that the common-anion rule indeed holds for all the common-anion system. For common-cation system, this rule holds for Al compounds and Ga compounds, but fails for In compounds. It is interesting to notice that although LDA underestimates the band gaps, similar chemical trends already exist in the LDA-calculated band gaps. To understand the general trends of the band gap variation in the common-cation system, we study the chemical and size contributions to the band gap. For the chemical contribution, we calculate the band gaps of Al, Ga, and In compounds at the fixed lattice constants of AlP, GaP, and InP, respectively. The results are shown in Table III. LDA corrections are included. We find that at the *fixed* phosphide volume, the band gaps of the common-cation system decrease from MSb to MP to MAs to MN ($M=$Al, Ga, and In), following the same trend of the anion atomic valence *s* orbital energies shown in Table IV. This trend can be understood as follows. The conduction-band minimum (CBM) at the Γ point is an anion *s* plus cation *s* state. The anion contribution increases as the compound becomes more ionic.[30] Since the variation of the anion *p* orbital energy, which determines the position of the valence-band maximum (VBM), is much less than the *s* orbitals (Table IV), the band gaps of the common-cation compounds at fixed volume generally follow the same trend of the valence *s* orbital energies of the anion. Therefore, since Sb $5s$ has the highest orbital energy, the band gap of MSb is also the largest at fixed volume. N $2s$ orbital energy is the lowest, 5.3 eV lower than the Sb $5s$ orbital energy, so its band gap is also the lowest. The $4s$ orbital energy of As is lower than the $3s$ orbital energy of P due to the incomplete screening of the $3d$ orbitals in As, so the band gap of MAs at fixed volume is also lower than that of MP.

Since the order of the band gaps calculated at the fixed volume is generally opposite to what is observed at the equilibrium lattice constants, except between InN and InP, the chemical contribution alone cannot explain the experimentally observed trend in the band gaps at equilibrium lattice constants. Next, we investigate the size or volume deformation contribution to the band gap. The calculated volume deformation potentials[27] with the LDA correction for III-V zinc-blende semiconductors are listed in Table II. We see that all the compounds have negative volume deformation potentials, i.e., when the volume decreases, the band gap increases. Therefore, it is clear that the common-cation rule and the common-anion rule for the band gap are mainly due to the large deformation potential of the III-V compounds.

TABLE II. Calculated band gaps and deformation potentials at Γ for ZB III-V compounds and WZ nitrides at the experimental (expt.) lattice constants (Ref. 1) using the LDA and LDA plus correction (LDA+C) methods. The E_g^{LDA+C} values with $*$ are fitted values, whereas all the others are predicted values. Our calculated results are compared with available experimental data (Ref. 1) and quasiparticle (QP) values (Refs. 18 and 19).

	a_{expt} (Å)	E_g^{LDA} (eV)	E_g^{LDA+C} (eV)	E_g^{QP} (eV)	E_g^{expt} (eV)	$-a_g^{LDA+C}$ (eV)
AlSb	6.133	1.24	2.28	2.23	2.32	8.9
GaSb	6.096	-0.38	0.81	0.62	0.81	8.0
InSb	6.479	-0.70	0.15	0.08	0.24	6.4
AlAs	5.660	1.75	3.05	2.88	3.13	8.9
GaAs	5.653	0.09	1.43	1.22	1.52	8.2
InAs	6.058	-0.64	0.36	0.31	0.42	5.7
AlP	5.467	3.06	4.42*	4.38		9.5
GaP	5.451	1.50	2.86*	2.85	2.86	8.8
InP	5.869	0.37	1.40*	1.44	1.46	5.9
AlN	4.360	4.28	6.00	6.0		10.2
GaN	4.500	1.72	3.34*	3.1	3.3	7.4
InN	4.980	-0.48	0.70			3.7
AlN(WZ)	$a=3.112$ $c=4.982$ $u=0.3819$	4.23	5.95	5.8	6.1	10.4
GaN(WZ)	$a=3.189$ $c=5.185$ $u=0.3768$	1.87	3.49	3.5	3.5	7.8
InN(WZ)	$a=3.544$ $c=5.718$ $u=0.3790$	-0.34	0.85			4.2

For example, at GaP lattice constant, the band gap of GaSb is 0.81 eV *larger* than that of GaP. However, GaSb is about 34% larger in volume than GaP. So, with an average deformation potential of -8.4 eV ($a_g=-8.0$ eV near its equilibrium volume), the band gap of GaSb at its equilibrium lattice constant is about 2.05 eV *smaller* than the band gap of GaP at its equilibrium lattice constant. The same situation applies to AlN and GaN: Even though AlN and GaN have much smaller band gaps than AlP and GaP at the lattice constants of AlP and GaP, respectively, their band gaps are larger than the phosphides at their own equilibrium lattice constants (Table III). This is because the volume of AlN is 68% smaller than that of AlP and GaN is 58% smaller than GaP, and AlN and GaN have large band-gap deformation potentials [$a_g(\text{AlN})=-10.2$ eV and $a_g(\text{GaN})=-7.4$ eV]. However, for InN, although its volume is about 49% smaller than InP, its band-gap deformation potential is small, $a_g(\text{InN})=-3.7$ eV. Because of this small $|a_g|$, the contribution due to the size or deformation potential is not sufficient to reverse the band-gap order due to the contribution of the chemical effect. This explains why InN has smaller band gap than InP.

From the analysis above, we see that the breakdown of the common-cation rule for the band gap in In compounds is due to the small $|a_g|$. We find that[27] the small $|a_g|$ for InN is due to the combined effects of (i) a large difference between the cation In 5s and anion N 2s orbital energies, (ii) a large repulsion between the N 2p and the high-lying In 4d orbitals, and (iii) a large In-N bond length (relative to AlN and GaN). Since a similar situation also exists in II-VI semiconductors,[27] one would expect that the breakdown of the common-cation rule should also apply to the II-VI sys-

TABLE III. Calculated (LDA+C) direct band gaps (in eV) at Γ of zinc-blende Al, Ga, and In compounds at their equilibrium (eq) lattice constants and at their respective phosphides lattice constants.

	$a=a_{\text{AlP}}$	a_{eq}		$a=a_{\text{GaP}}$	a_{eq}		$a=a_{\text{InP}}$	a_{eq}
AlN	0.45	6.00	GaN	-0.61	3.34	InN	-1.27	0.70
AlP	4.42	4.42	GaP	2.86	2.86	InP	1.40	1.40
AlAs	4.04	3.05	GaAs	2.36	1.43	InAs	0.92	0.36
AlSb	5.69	2.28	GaSb	3.67	0.81	InSb	2.15	0.15

TABLE IV. Atomic s and p orbital energy levels (in eV) for group III and group V elements.

Atom	ϵ_s	ϵ_p
Al	-7.91	-2.86
Ga	-9.25	-2.81
In	-8.56	-2.78
N	-18.49	-7.32
P	-14.09	-5.68
As	-14.70	-5.34
Sb	-13.16	-5.08

tems. Indeed, experimental data show that the ZnO band gap of 3.4 eV is smaller than the ZnS band gap of 3.8 eV. Our calculations show similar trends between CdO and CdS and between HgO and HgS in the ZB phase. In the past, this band-gap anomaly between ZnO and ZnS was explained in terms of the stronger p-d repulsion[30] in ZnO than in ZnS, because the Zn $3d$ to O $2p$ orbital energy difference is smaller. The larger p-d repulsion pushes up the VBM of ZnO more than that of ZnS, therefore, reducing the band gap of ZnO. However, the calculated VBM offset $\Delta E_V = 1.0$ eV between ZnO and ZnS[31] indicates that the p-d repulsion mechanism, which affects only the VBM, is not sufficient to explain this band-gap anomaly. As explained above, to understand the breakdown of the band-gap-common-cation rule for the II-VI compounds, we also have to take into account the contribution at CBM, i.e., the fact that O $2s$ orbital is much lower in energy than the S $3s$ orbital, and that oxides have very small band-gap deformation potentials.[27]

In summary, using an LDA band-structure method with band-gap correction we have systematically studied the chemical trends of band-gap variation in III-V semiconductors. We find that InN has a band gap of 0.85 ± 0.1 eV,[32] in good agreement with recent experimental measurements. We show that the previously accepted band-gap-common-cation rule does not hold for ionic InN and the II-VI oxides.

S.-H.W. would like to thank Professor H. X. Jiang, Professor Y. F. Chen, and Professor G. X. Gao for valuable discussions. I.G.B. would like to thank X. Gonze, V. Olevano, and G. M. Rignanese for help in using the ABINIT code. This work was supported by the U.S. Department of Energy, Contract No. DE-AC36-99GO10337.

---

[1] *Intrinsic Properties of Group IV Elements and III-V, II-VI, and I-VII Compounds*, edited by O. Madelung, M. Schulz, and H. Weiss, Landolt-Börnstein, New Series, Group III, Vol. 22, Pt. a (Springer, Berlin, 1987); The band gap for AlN is quoted from J. Li, K.B. Nam, M.L. Nakami, J.Y. Lin, and H.X. Jiang, Appl. Phys. Lett. **81**, 3365 (2002).

[2] T.L. Tansley and C.P. Foley, J. Appl. Phys. **59**, 3241 (1986).

[3] T. Inushima, V.V. Mamutin, V.A. Vekshin, S.V. Ivanoc, T. Sakon, M. Motokawa, and S. Ohoya, J. Cryst. Growth **227/228**, 481 (2001).

[4] V. Yu *et al.*, Phys. Status Solidi B **230**, R4 (2002); **229**, R1 (2002).

[5] J. Wu, W. Walukiewicz, K.M. Yu, J.W. Ager III, E.E. Haller, H. Lu, J. Schaff, Y. Saito, and Y. Nanishi, Appl. Phys. Lett. **80**, 3967 (2002).

[6] J. Wu, W. Walukiewicz, K.M. Yu, J.W. Ager III, E.E. Haller, H. Lu, and J. Schaff, Appl. Phys. Lett. **80**, 4741 (2002).

[7] T. L. Tansley, *Properties of Group III Nitrides*, edited by J.H. Edgar (INSPEC, London, 1999), p. 39.

[8] K. Osamura, K. Nakajima, Y. Murakami, P.H. Shingu, and A. Ohtsuki, Solid State Commun. **11**, 617 (1972).

[9] K. Osamura, S. Naka, and Y. Murakami, J. Appl. Phys. **46**, 3432 (1975).

[10] T.S. Moss, Proc. Phys. Soc. London, Sect. B **67**, 775 (1954); E. Burstein, Phys. Rev. **93**, 632 (1954).

[11] *Properties of Group III Nitrides*, edited by J.H. Edgar (INSPEC, London, 1994).

[12] H. Morkoc, *Nitride Semiconductors and Devices* (Springer, New York, 1999).

[13] C. Wetzel *et al.*, Appl. Phys. Lett. **73**, 1994 (1998).

[14] S. Yamaguchi *et al.*, Appl. Phys. Lett. **76**, 876 (2000).

[15] L. Bellaiche, T. Mattila, L.W. Wang, S.-H. Wei, and A. Zunger, Appl. Phys. Lett. **74**, 1842 (1999).

[16] P.R.C. Kent and A. Zunger, Appl. Phys. Lett. **79**, 1977 (2001).

[17] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L.J. Sham, *ibid.* Phys. Rev. **140**, A1133 (1965).

[18] X. Zhu and S.G. Louie, Phys. Rev. B **43**, 14 142 (1991).

[19] A. Rubio, J.L. Corkill, M.L. Cohen, E.L. Shirley, and S.G. Louie, Phys. Rev. B **48**, 11 810 (1993).

[20] S.H. Wei and A. Zunger, Phys. Rev. B **57**, 8983 (1998).

[21] C.-Y. Yeh, S.-H. Wei, and A. Zunger, Phys. Rev. B **50**, 2715 (1994).

[22] M. van Schilfgaarde, A. Sher, and A.-B. Chen, J. Cryst. Growth **178**, 8 (1997); A. Sher, M. van Schilfgaarde, M.A. Berding, S. Krishnamurthy, and A.-B. Chen, MRS Internet J. Nitride Semicond. Res. **4S1**, G5.1 (1999).

[23] T. Kotani and M. van Schilfgaarde, Solid State Commun. **121**, 461 (2002).

[24] D. Vogel, P. Kruger, and J. Pollmann, Phys. Rev. B **55**, 12 836 (1997); C. Stampfl, C.G. Van de Walle, D. Vogel, P. Kruger, and J. Pollmann, *ibid.* **61**, R7846 (2000).

[25] K.A. Johnson and N.W. Ashcroft, Phys. Rev. B **58**, 15 548 (1998).

[26] Our GW calculation is performed using the open-source ABINIT3.3 code, which was developed by the Universite Catholique de Louvain, Corning Incorporated, and other contributors (URL: http://www.abinit.org). The pseudopotentials are generated using the FHI98PP code [M. Fuchs and M. Scheffler, Comput. Phys. Commun. **119**, 67 (1999)]. For GaAs, our calculated quasipartical gap is 1.4 eV. For the nitrides, we find that the GW gaps are 3.3, 3.5, 0.3, 0.5 eV, respectively, for ZB-GaN, WZ-GaN, ZB-InN, and WZ-InN. Detailed discussion of the GW results will be presented elsewhere.

[27] S.-H. Wei and A. Zunger, Phys. Rev. B **60**, 5404 (1999).

[28] S.-H. Wei and H. Krakauer, Phys. Rev. Lett. **55**, 1200 (1985); D.J. Singh, *Planewaves, Pseudopotentials, and the LAPW Method*, (Kluwer, Boston, 1994).

[29] N.E. Christensen, Phys. Rev. B **30**, 5753 (1984).

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

[31] S.-H. Wei and A. Zunger, Appl. Phys. Lett. **72**, 2011 (1998).

[32] After the submission of this paper we note that the quasipartical calculation of F. Bechstedt and J. Furthmuller, J. Cryst. Growth **246**, 315 (2002) also find that the band gap of InN is less than 1 eV, in agreement with our result.