Bowing of the defect formation energy in semiconductor alloys

Jie Ma and Su-Huai Wei

Giant bowing of the defect formation energy in semiconductor alloys

Jie Ma and Su-Huai Wei*

National Renewable Energy Laboratory, Golden, Colorado 80401, USA

Using first-principles method and special quasirandom structure (SQS) approach, we have studied the formation energies of two prototype defects in alloys: Ge_{As} in $Al_xGa_{1-x}As$ and Cu_{Cd} in CdS_xTe_{1-x}. We find that giant bowing effects for the defect formation energy can exist in semiconductor alloys. The bowing effect originates from the concentrated distribution of defects at low energy sites caused by the defect wavefunction localization and the size-mismatch-induced strain effect. Because the bowing effect can drastically reduce the defect formation energy — even in dilute semiconductor alloys — it can have wide applications, such as alloy-enhanced defect solubility in semiconductors.

Alloying different semiconductors is a popular approach to broaden the range of available material properties for specific applications. Considerable efforts have been made to understand the compositional dependence of properties of isovalent $A_xB_{1-x}C$ semiconductor alloys. Many alloy properties $P(x)$ can be described as a linear average of the corresponding quantities in two pure constituents, plus a quadratic term [1, 2]: $P(x) = [xP(AC) + (1-x)P(BC)] - bx(1-x)$, where b is the so-called bowing coefficient. Generally, for isovalent alloys with small chemical and size mismatch, the bowing coefficients are small constant numbers; for alloys with large chemical and size mismatch, the bowing coefficients could be large and composition dependent [3–5].

Most previous studies on alloying effects were focused on macroscopic properties, such as the band gap and mixing enthalpy of substitutional alloys. The compositional dependence of defect properties in disordered alloys, however, has not been well studied, despite the fact that the electrical and optical properties of semiconductors and their alloys are strongly determined by defects. It is quite common that defect properties of alloys were discussed based on data of the pure constituents [6–8], assuming simply a linear compositional dependence based on the virtual-crystal approximation (VCA) [9]. The bowing effects of the defect formation energy have never been systematically discussed. However, because the defect solubility has an exponential dependence on the formation energy, it is important to understand the compositional dependence of the defect formation energy (solubility) and the physical origin for possible bowings.

Generally, defect properties are determined primarily by local motifs because the wavefunction of the defect state is usually localized around the defect site. In a tetrahedral $A_xB_{1-x}C$ alloy, there are many distinct C sites with different local motifs. For example, the first-neighbor motif around a C atom could be A_4, A_3B_1, A_2B_2, A_1B_3, or B_4. Each of these motifs could contribute differently to a given property. The old-fashioned VCA or coherent-potential approximation (CPA) [10] assumes that the C atom has a single averaged environment, and therefore cannot simulate the contributions from different motifs. More advanced atomistic models, such as the special quasirandom structure (SQS) approach [11, 12], must be used.

Using first-principles method and SQS approach, we have calculated the defect formation energies in two prototype systems: Ge substitution on As site (Ge_{As}) in $Al_xGa_{1-x}As$, which is nearly lattice matched, and Cu substitution on Cd site (Cu_{Cd}) in CdS_xTe_{1-x}, which has highly size mismatched mixing elements, with the composition $x = 0$, 0.25, 0.5, 0.75, and 1. We find that the formation energies of the defects strongly depend on the first-neighbor motif. To calculate the solubility, it is convenient to define an effective formation energy. For Ge_{As} in $Al_xGa_{1-x}As$ alloys, the effective formation energy has a nearly linear compositional dependence at the high temperature limit (i.e., no bowing); however, at the low temperature limit, the effective formation energy can even change discontinuously at $x \sim 1$ (i.e., infinitive bowing), due to the concentrated distribution of defects. For Cu_{Cd} in CdS_xTe_{1-x} alloys, there exists a strain-induced bowing even at the high temperature limit. Due to the large bowing, the formation energy (solubility) of Cu_{Cd} actually decreases (increases) compared to that in CdTe, as a few S atoms are incorporated, even though the formation energy of Cu_{Cd} is higher in pure CdS than in pure CdTe.

Our calculations were based on density functional theory (DFT) within the generalized gradient approximation (GGA) [13] as implemented in VASP code [14]. The projector augmented wave (PAW) pseudopotentials [15] were employed and the wavefunctions were expanded in planewave basis-set with an energy cutoff of 300 eV. To simulate random alloys, we constructed SQSs ($x = 0.25$, 0.5, and 0.75) in a 512-atom cubic cell. In constructing the SQSs, the atomic correlation functions of pairs up to the fourth neighbor, the first neighbor triangle and tetrahedral were required to be the same as those in perfectly random alloys. The Brillouin zone integration was sampled with Γ-point only. As expected, the calculated alloy lattice constant follows the Vegards rule [16]. In

a perfectly random alloy $A_xB_{1-x}C$, the probability of finding a C site with the first-neighbor motif A_nB_{4-n} is $p_n(x) = C_4^n x^n (1-x)^{4-n}$. In the SQSs we constructed, the probabilities are exactly the same as those in random alloys.

The formation energy of a defect α at the site s in the alloy with the composition x is defined as [17]

$$\Delta H(\alpha, s, x) = E(\alpha, s, x) - E(x) + \sum_i n_i \mu_i, \quad (1)$$

where $E(\alpha, s, x)$ is the energy of the alloy containing a defect α at the site s and $E(x)$ is the energy of the alloy in the same supercell without defects. μ_i is the chemical potential of the ith constituent, and n_i is the number of atoms transferred from the supercell to the reservoirs in forming the defect. Here, we employ their elemental solid values for chemical potentials.

The solubility of the defect α is the summation of the probabilities of finding α at every site s [18]. In alloys, the defect formation energy $\Delta H(\alpha, s, x)$ depends on s and x. To calculate the defect solubility, it is convenient to define an effective formation energy $\Delta H_{eff}(\alpha, x, T)$, which fulfills

$$\exp[-\Delta H_{eff}(\alpha, x, T)/k_B T] = \frac{1}{N} \sum_s \exp[-\Delta H(\alpha, s, x)/k_B T],$$
$$(2)$$

where T is the doping temperature, k_B is the Boltzmann constant, and N is the number of all defect sites in the alloy. According to this definition, the solubility of the defect α (depending on x and T) can be calculated in the traditional way [18]

$$[\alpha]_{x,T} = N \exp[-\Delta H_{eff}(\alpha, x, T)/k_B T]. \quad (3)$$

It is interesting to see that at the high temperature limit ($T \to \infty$), the Boltzmann factor $\Delta H(\alpha, s, x)/k_B T \to 0$, and the effective formation energy is

$$\Delta H_{eff}(\alpha, x, \infty) = \frac{1}{N} \sum_s \Delta H(\alpha, s, x), \quad (4)$$

which is the arithmetic average of the formation energies at all sites. This is expected because as $T \to \infty$, all the sites have equal opportunity to be occupied.

At the low temperature limit ($T \to 0$), $\Delta H(\alpha, s, x)/k_B T \to \infty$. Therefore, under equilibrium conditions, the defect can only occupy the sites with the lowest formation energy, i.e.,

$$\Delta H_{eff}(\alpha, x, 0) = \Delta H(\alpha, s_0, x), \quad (5)$$

where s_0 is the site at which the defect α has the lowest formation energy. This is a very interesting observation, suggesting the defect formation energy at low T is controlled by local properties, not by averages over the whole alloy.
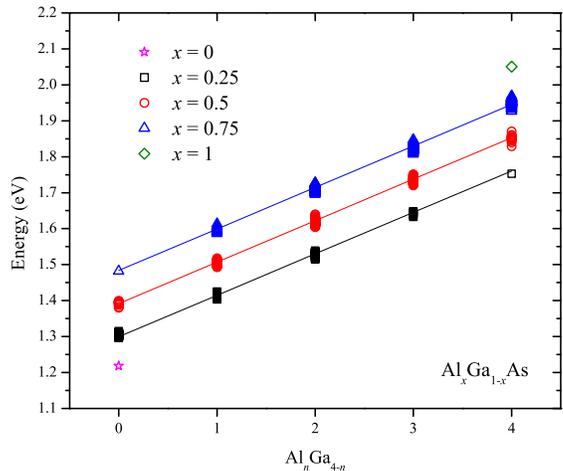

FIG. 1: The formation energies of Ge$_{As}$ at every site in Al$_x$Ga$_{1-x}$As alloys ($x = 0.25$, 0.5, and 0.75) as a function of the number of Al atoms in the first neighbor around the defect. For comparison, the defect formation energies in GaAs and AlAs ($x = 0$, 1) are also plotted. The solid lines are the linear fitting [Eq.(6)].
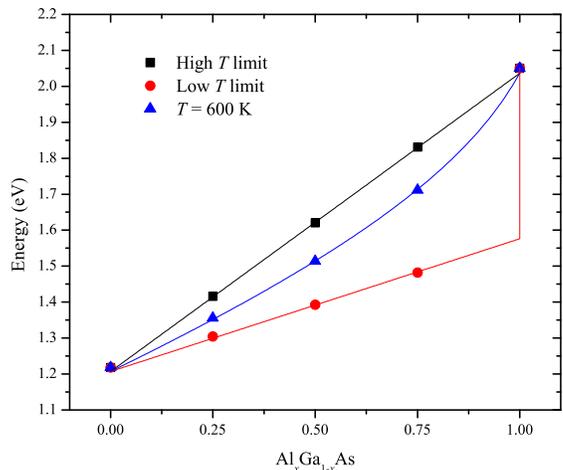

FIG. 2: The effective formation energy of Ge$_{As}$ in Al$_x$Ga$_{1-x}$As alloys ($x = 0$, 0.25, 0.5, 0.75, and 1) at the high temperature limit, the low temperature limit, and a finite temperature $T = 600$ K. The solid lines are calculated with the linear fitting [Eq.(6)].

We first consider Ge$_{As}$ in Al$_x$Ga$_{1-x}$As alloys. Al$_x$Ga$_{1-x}$As alloys have many applications in optical and electronic devices, and Ge is one of the most widely used p-type dopants [19]. AlAs and GaAs are nearly lattice matched and Ge is also close to As in size, so there is little strain effect in this system. The formation energies of Ge$_{As}$ at all the 256 defect sites in each alloy ($x = 0.25$, 0.5, and 0.75), and those in GaAs and AlAs, are displayed in Fig. 1. In this figure, the defect formation energies in alloys are plotted as a function of the number of Al atoms in the first neighbor around the defect. First, we notice that although the formation energies of Ge$_{As}$ in the al-

loys spread in a large energy range, Ge_{As} with the same first-neighbor motif (i.e., the same n in Fig. 1) in each alloy actually have very similar formation energies. This is because the wavefunction of the defect state is localized around the defect site, and therefore the defect formation energy is determined mostly by the first-neighbor motif. Next, we observe that the formation energies in each alloy increase linearly as the number of Al atoms in the first neighbor around the defect (n), which is consistent with the trend that the formation energy of Ge_{As} is 0.83 eV higher in AlAs than that in GaAs. This is because the localized defect state has strong interactions with the first-neighbor atoms. For defects surrounded by Ga atoms in the first neighbor, the Ga d orbitals push up the energy of the localized defect state through the p-d coupling [20]. It costs less energy to create holes at high-energy levels, so the formation energy is low. We also find that as the composition x increases from 0 to 1, the formation energies increase. This is also explained by the electronic effect. As more Al atoms are incorporated into the system, the valence band maximum (VBM) of the alloy decreases in energy, because the VBM of AlAs is lower in energy than that of GaAs [20]. Therefore, it costs more energy to create holes and the formation energy increases. In fact, to a good approximation, the Ge_{As} formation energy in $Al_xGa_{1-x}As$ can be expressed (in eV) as

$$\Delta H(Ge_{As}, n, x) = 1.207 + 0.115n + 0.369x. \quad (6)$$

The effective formation energies $\Delta H_{eff}(Ge_{As}, x, T)$ in $Al_xGa_{1-x}As$ alloys ($0 \leq x \leq 1$) at the high temperature limit, the low temperature limit, and a finite temperature $T = 600$ K are displayed in Fig. 2. We notice that at the high temperature limit ($T \to \infty$), the effective formation energy of Ge_{As} increases almost linearly with the alloy composition x (the defect solubility decreases exponentially with x), which indicates that there is almost no bowing effect at the high temperature limit. This is because as $T \to \infty$, all the sites can be occupied equally, and therefore the effective formation energy reflects the averaged property of the whole alloy. As the temperature decreases, the defect distribution becomes more concentrated at sites with lower formation energies, so the effective formation energy bows downwards. At the low temperature limit ($T \to 0$), the bowing effect is the largest. As discussed before, when $T \to 0$, only sites with the lowest formation energies can be occupied, so the defects are all concentrated at sites surrounded by four Ga atoms. The effective formation energy is determined by these sites. At $x = 1$, there is no site surrounded by Ga atoms. However, as Ga is incorporated, sites surrounded by four Ga atoms suddenly appears, so the effective formation energy shows a discontinuous change at $x \sim 1$ as $T \to 0$.

It is interesting to see that the bowing is asymmetric in $Al_xGa_{1-x}As$, larger on the Al-rich side ($x \sim 1$), but smaller on the Ga-rich side ($x \sim 0$). As discussed above, the bowing is induced by the different formation energies of defects at various sites, which is caused by the defect wavefunction localization. Because Ga creates a level above the VBM of AlAs, Ge_{As} surrounded by Ga atoms have strong wavefuction localization. As Ga atoms are incorporated into AlAs, the strong wavefuction localization of defects surrounded by Ga atoms causes the large bowing at $x \sim 1$. On the other hand, Al does not create localized levels in GaAs, so the localization of the defect wavefunction is weak, and the bowing is also weak at $x \sim 0$.
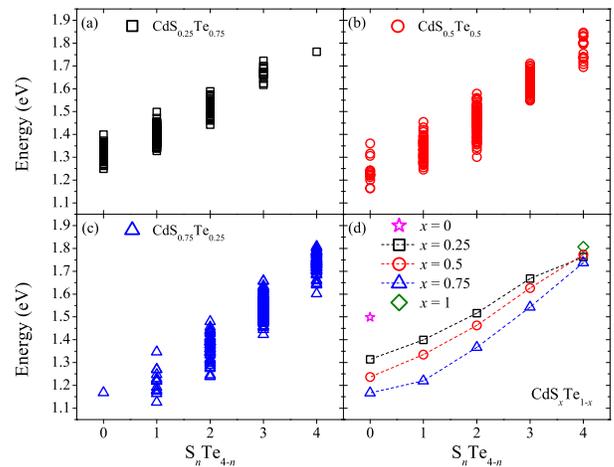

FIG. 3: The formation energies of Cu_{Cd} at every site in CdS_xTe_{1-x} alloys ($x = 0.25, 0.5,$ and 0.75) as a function of the number of S atoms in the first neighbor around the defect. The arithmetic averages of the formation energies of defects with the same first-neighbor motif are plotted in (d). For comparison, the defect formation energies in CdTe and CdS ($x = 0, 1$) are also plotted. The dashed lines are guides for eyes.

Next, we consider Cu_{Cd} in CdS_xTe_{1-x} alloys. CdS_xTe_{1-x} alloys commonly exist in CdTe/CdS solar cells [21], and Cu is the most important p-type dopant [22]. The mixing elements S and Te exhibit large size mismatch and the size of Cu is much smaller than that of Cd, so the strain is expected to play an important role. For simplicity, we only consider zincblende CdS_xTe_{1-x} alloys. The formation energies of Cu_{Cd} at all the 256 sites in each alloy ($x = 0.25, 0.5,$ and 0.75) as a function of the number of S atoms in the first neighbor are displayed in Fig. 3 (a-c). Similar to the Ge_{As} case, the formation energies of Cu_{Cd} in CdS_xTe_{1-x} alloys are strongly dependent on the first-neighbor motif. However, due to the large size mismatch and local atomic relaxations, the formation energies of defects within a given first-neighbor motif are more scattered. The arithmetic averages of the formation energies of defects with the same first-neighbor motif in each alloy are displayed in Fig. 3 (d), along with the formation energies of Cu_{Cd}
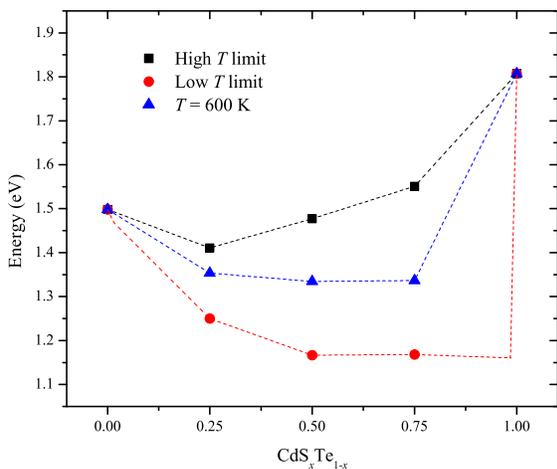
FIG. 4: The effective formation energy of Cu_{Cd} in CdS_xTe_{1-x} alloys ($x = 0$, 0.25, 0.5, 0.75, and 1) at the high temperature limit, the low temperature limit, and a finite temperature $T = 600$ K. The dashed lines are guides for eyes.

in zincblende CdTe and CdS. We find that in each alloy, the formation energy of Cu_{Cd} increases as the number of S atoms in its first neighbor, which is consistent with the trend that the formation energy of Cu_{Cd} is 0.31 eV higher in CdS than that in CdTe. This is because the Te 5p orbital has higher energy than the S 3p orbital. For defects surrounded by more Te atoms in the first neighbor, the localized defect wavefunctions contain more Te 5p components, and therefore have higher energies. It costs less energy to create holes at high-energy levels, so the formation energy is low. Different from the Ge_{As} case, the Cu_{Cd} formation energies generally decrease as x increases from 0.25 to 0.75. The formation energy decrease is more significant for defects surrounded by more Te atoms in the first neighbor (small n). This is because in addition to the electronic effect discussed in the Ge_{As} case, the strain effect also plays an important role. The size of Cu is small, so after forming Cu_{Cd}, the system carries a compressive strain. Due to the large lattice mismatch between CdS and CdTe, as S atoms are incorporated into the system (x increases), the lattice constant of the alloy decreases. Therefore, the strain induced by the defect is released, so the formation energy decreases [23]. In each alloy, sites surrounded by more S atoms have shorter bond lengths, so the strain effect is small; sites surrounded by more Te atoms have longer bond lengths, so the strain effect is large and the formation energies change more significantly. For alloys with $x \sim 1$, the strain effect is the least significant due to the smallest bong lengths, so the electronic effect is dominant. Because CdS has the lowest-energy VBM, it costs the largest energy to create holes and therefore the defect formation energy in CdS is higher than those in the three alloys.

The effective formation energies $\Delta H_{eff}(Cu_{Cd}, x, T)$ in CdS_xTe_{1-x} alloys ($x = 0$, 0.25, 0.5, 0.75, and 1) at the high temperature limit, the low temperature limit, and a finite temperature $T = 600$ K are displayed in Fig. 4. Similar to the Ge_{As} case, the bowing increases as the temperature decreases, because the distribution of the defect becomes more concentrated at low-energy sites. In contrast to the Ge_{As} case, at the high temperature limit, the effective formation energy shows a large positive bowing and the slop is even negative on the Te-rich side ($x \sim 0$). This is due to the strain effect that does not exist in the previous case. As discussed before, the Cu_{Cd} formation energy decreases as x increases, due to the release of the strain energy. The effective formation energy at the high temperature limit is the average of formation energies at all sites, so it bows downwards. When the composition x is small, most Cd atoms are in the Te_4, S_1Te_3, or S_2Te_2 first-neighbor motifs. Because the defect formation energies at these sites are lower than that in CdTe, the effective formation energy is lower than that in CdTe (i.e., the solubility of Cu_{Cd} is larger). However, when x is large, most Cd atoms are in the S_4 or S_3Te_1 first-neighbor motifs. Because the defect formation energies at these sites are higher, the effective formation energy will eventually increase as x increases (i.e., the solubility decreases) at high temperatures. Therefore, due to the large bowing, the defect solubility is enhanced in small x alloys, compared to those in the two pure constituents.

Our analysis indicates that for Cu_{Cd} in CdS_xTe_{1-x} alloys, both the electronic effect and strain effect play roles in the defect formation energy. The wavefunction localization effect is dominant for S-rich alloys ($x \sim 1$) because Te create a localized level in CdS, but the strain effect is dominant for Te-rich alloys ($x \sim 0$) because S can reduced the large strain in CdTe caused by Cu_{Cd}.

In conclusion, we have investigated the formation energies of two prototype defects in alloys: Ge_{As} in $Al_xGa_{1-x}As$ and Cu_{Cd} in CdS_xTe_{1-x}. We find that the bowing has two origins: the concentrated distribution of defects at low energy sites at finite temperatures caused by the defect wavefunction localization and the size-mismatch-induced strain effect. The bowing increases as the temperature decreases. Since the electronic occupation and size of defects depend on charge states, the bowing of the defect formation energy should also be charge-dependent. Our theoretical model is consistent with experimental observations [24–26]. For examples, it is observed that Cu atoms accumulate in the mixed CdS_xTe_{1-x} alloy region when small amount of S is mixed into CdTe at the CdTe/CdS interfaces in CdTe based solar cell [25]. In p-type ZnSe, N solubility can also be significantly enhanced when small amount of ZnTe is mixed into the host [26].

* Electronic address: swei@nrel.gov

[1] *Physics of Group IV Elements and III-V Compounds*, edited by O. Madelung, M. Schultz, and H. Weiss, Landolt-Bornstein Numerical Data and Functional Relationships in Science and Technology, Vol. 17, Part A (Springer-Verlag, Berlin, 1982).

[2] S. Adachi, *Physical Properties of III-V Semiconductor Compounds* (Wiley, New York, 1992).

[3] S.-H. Wei and A. Zunger, Phys. Rev. Lett. **76**, 664 (1996).

[4] L. Bellaiche, S.-H. Wei, and A. Zunger, Phys. Rev. B **54**, 17568 (1996).

[5] J. Neugebauer and C. G. Van de Walle, Phys. Rev. B **51**, 10568 (1995).

[6] C. Stampfl and C. G. Van de Walle, Appl. Phys. Lett. **72**, 459 (1998).

[7] M. D. McCluskey, N. M. Johnson, C. G. Van de Walle, D. P. Bour, M. Kneissl, and W. Walukiewicz, Phys. Rev. Lett. **80**, 4008 (1998).

[8] C. Stampfl, J. Neugebauer, and C. G. Van de Walle, Materials Science and Engineering:B **59**, 253 (1999).

[9] L. Nordheim, Ann. Phys. (Leipzig) **9**, 607 (1931).

[10] P. Soven, Phys. Rev. **156**, 809 (1967).

[11] A. Zunger, S.-H. Wei, L. G. Ferreira, and J. E. Bernard, Phys. Rev. Lett. **65**, 353 (1990).

[12] S.-H. Wei, L. G. Ferreira, J. E. Bernard, and A. Zunger, Phys. Rev. B **42**, 9622 (1990).

[13] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[14] G. Kresse and J. Furthmuller, Phys. Rev. B **54**, 11169 (1996).

[15] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).

[16] L. Vegard, Z. Phys. **5**, 17 (1921).

[17] S.-H. Wei, Comp. Mater. Sci. **30**, 337 (2004).

[18] J. E. Northrup and S. B. Zhang, Phys. Rev. B **47**, 6791 (1993).

[19] Y. R. Yuan, K. Mohammed, and J. L. Merz, J. Appl. Phys. **57**, 2896 (1985), and references therein.

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

[21] D. Wang, Z. Hou, and Z. Bai, J. Mater. Res. **26**, 697 (2011), and references therein.

[22] J. Ma, S.-H. Wei, T. A. Gessert, and K. K. Chin, Phys. Rev. B **83**, 245207 (2011).

[23] J. Zhu, F. Liu, G. B. Stringfellow, and S.-H. Wei, Phys. Rev. Lett. **105**, 195503 (2010).

[24] R. Kube, H. Bracht, J. Lundsgaard Hansen, A. Nylandsted Larsen, E. E. Haller, S. Paul, and W. Lerch, J. Appl. Phys. **107**, 073520 (2010), and references therein.

[25] K. D. Dobson, I. Visoly-Fisher, G. Hodes, and D. Cahen, Solar Energy Mater. and Solar Cells **62**, 295 (2000).

[26] W. Lin, B. X. Yang, S. P. Guo, A. Elmoumni, F. Fernandez, and M. C. Tamargo, Appl. Phys. Lett. **75**, 2608 (1999).