Review Article

Progress in ZnO Acceptor Doping: What Is the Best Strategy?

Judith G. Reynolds¹ and C. Lewis Reynolds²

¹ Department of Physics, North Carolina State University, Raleigh, NC 27695, USA
² Department of Materials Science and Engineering, North Carolina State University, Raleigh, NC 27695, USA

Correspondence should be addressed to C. Lewis Reynolds; lew_reynolds@ncsu.edu

Received 13 November 2013; Revised 23 April 2014; Accepted 23 April 2014; Published 22 May 2014

Academic Editor: Jianrong Qiu

Copyright © 2014 J. G. Reynolds and C. L. Reynolds. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper reviews the recent progress in acceptor doping of ZnO that has been achieved with a focus toward the optimum strategy. There are three main approaches for generating p-type ZnO: substitutional group IA elements on a zinc site, codoping of donors and acceptors, and substitution of group VA elements on an oxygen site. The relevant issues are whether there is sufficient incorporation of the appropriate dopant impurity species, does it reside on the appropriate lattice site, and lastly whether the acceptor ionization energy is sufficiently small to enable significant p-type conduction at room temperature. The potential of nitrogen doping and formation of the appropriate acceptor complexes is highlighted although theoretical calculations predict that nitrogen on an oxygen site is a deep acceptor. We show that an understanding of the growth and annealing steps to achieve the relevant acceptor defect complexes is crucial to meet requirements.

1. Introduction

Although zinc oxide has been investigated for many years, its potential for photonic and electronic applications has led to significant resurgence in interest during the last decade. Its use as a widely diverse functional material is enhanced by the fact that it can be grown in bulk, thin film, and nanostructures, examples of the latter being nanowires, nanobelts, and other morphologies that are dependent upon growth conditions. Although bulk substrates are readily available, ZnO thin films can also be grown heteroepitaxially across the misfit scale via the paradigm of domain matching epitaxy [1, 2]. This is relevant for fabrication of next generation devices in which multiple functionalities are integrated on a given substrate. Potential applications of ZnO are blue/UV LEDs and lasers, photodetectors, transparent thin film transistors, transparent conducting oxides, gas sensors, and nanostructured piezoelectric nanogenerators, and substantial markets have been predicted. Many of the potential device applications, however, require both donor and acceptor doping, and growth of reproducible and stable p-type ZnO has been difficult to achieve. This doping asymmetry problem in which n-type doping is easily achieved while p-type is quite problematic is well known [3], and thus widespread development of ZnO-based devices has been inhibited. For this special issue on structural, electronic, and optical properties of functional metal oxides, we do not intend to provide an exhaustive review of ZnO materials and devices as many review articles [3–9] have already been published but instead to focus on recent progress in acceptor doping in ZnO and to provide our assessment of the strategies pursued thus far.

Zinc oxide is a wide band gap semiconductor that typically crystallizes in the hexagonal wurtzite structure with a 3.37 eV band gap at 300 K that is comparable to GaN, both of which are suitable for UV and blue photonic devices. Zinc oxide also has several other characteristics that make it an attractive semiconducting material. As a consequence of its less than ideal $c/a = \sqrt{(8/3)} = 1.633$ ratio in the ZnO wurtzite structure and tetrahedral coordination, it exhibits spontaneous polarization along the c-axis, the strength of which is related to the deviation from ideality. Strains in thin film heterostructures can also give rise to an additional piezoelectric polarization. On the basis of the Pauling electronegativity scale, the Zn–O bond has a significant, 60%, ionic nature instead of being considered a covalent compound. Furthermore, substitution of Mg or Cd
ions on the Zn cation sublattice enables band gap tuning above and below the nominal 3.37 eV band gap for growth of heterostructures. Advantages that ZnO has in comparison to GaN are availability of a native substrate, relative ease of wet chemical etching for device fabrication, its large exciton binding energy of approximately 60 meV [10] and bixciton binding energies on the order of the 25 meV thermal energy at room temperature [11]. These latter characteristics make it attractive for low threshold and large differential quantum efficiency photonic devices in the UV and blue portion of the electromagnetic spectrum. While current microelectronic, nanoelectronic, and optoelectronic devices are based on the flow of electric charge, integration of true multifunctionality on a chip will only be achieved when the spin of the electron can also be taken into consideration. Spin of the electron can be manipulated by an applied magnetic field and has relatively long relaxation times, which implies that devices can be smaller, use less energy, and provide logic operations. Zinc oxide has been predicted to exhibit ferromagnetic behavior at room temperature via a hole-mediated exchange mechanism [12], and thus stable p-type behavior is a requirement for successful demonstration of ZnO spintronics in electronic and photonic applications. While the above points to the significant promise of ZnO-based devices, two major issues must be resolved: high defect densities and doping asymmetry. The former is readily addressed by fabrication of ZnO-based structures that have been grown homoepitaxially. As stated above, however, the doping asymmetry problem is the dominant issue that limits ZnO-based materials and devices.

There are three fundamentally different approaches for growth of p-type ZnO: substitution of Group IA impurities, for example, Li, Na, or K on a Zn site, substitution of Group VA impurities such as N, P, As, and Sb on the O site, and codoping with donors and acceptors. It is clear that the successful approach will be that which enables substantial p-type conduction at room temperature; that is, the acceptor ionization energy must be relatively shallow and compensation by background impurities and/or defects must be minimized. Here, we attempt to summarize the recent accomplishments that have been achieved via each of these methods and to assess the trends in order to postulate what we believe to be the optimum strategy for both thin film heterostructures and nanostructures. Below, we will first address the origin of background donors, which is highly dependent upon whether the material is grown in a Zn-rich or an O-rich environment, the role of H as a donor in ZnO, and the issue of the acceptor ionization energy of substitutional N on the O sublattice. We follow this with a summary of recent reports on the p-type behavior for the various potential doping schemes and then conclude with our outlook.

2. Origin of Background Donors in ZnO

In order to be able to achieve the high (>10^{17} cm^{-3}) acceptor concentration necessary for useful p-type conductivity, the solid solubility of the acceptor impurity must be relatively high while the self-compensating defects and acceptor ionization energies must be low. Nominally undoped ZnO is n-type with carrier concentrations ranging from 10^{14} to mid-10^{16} cm^{-3}. Traditionally, the underlying cause of the unintentional n-type conductivity has been assigned to the formation of zinc interstitials (Zn_i) and oxygen vacancies (V_O) [13]. Comprehensive calculations [13, 14] of the formation energies of dilute native point defects in ZnO via first principles density functional theory (DFT) within the local density approximation (LDA) [15, 16] have thus been performed. The pseudopotential plane-wave method [17] employed uses constant volume supercells and a pair potential approach. At equilibrium, the concentration, c, of a defect in the crystal lattice as a function of its formation energy, E_f', is as follows:

\[ c = N_{\text{sites}} \exp \left( \frac{-E_f'}{k_B T} \right), \]

where \( N_{\text{sites}} \) is the concentration of sites in the crystal where the defect can occur and \( k_B \) and \( T \) are the Boltzmann constant and temperature, respectively. Hence, a low \( E_f' \) implies a high equilibrium concentration of the defect and a high \( E_f' \) means that defect formation is less likely to occur. Additionally, the formation energy \( E_f' \) of a point defect in a charge state \( q \) is given by [13]

\[ E_f'(q) = E_{\text{tot}}(q) - n_{\text{Zn}} \mu_{\text{Zn}} - n_{\text{O}} \mu_{\text{O}} - qE_F, \]

where \( E_{\text{tot}}(q) \) is the total energy of a system containing \( n_{\text{Zn}} \) and \( n_{\text{O}} \) zinc and oxygen atoms, \( \mu_{\text{Zn}} \) and \( \mu_{\text{O}} \) are the chemical potentials for zinc and oxygen, respectively; and \( E_F \) is the Fermi energy. The chemical potentials depend upon the growth conditions. For high zinc partial pressure, \( \mu_{\text{Zn}} = \mu_{\text{Zn(bulk)}} \), while for high oxygen partial pressure, \( \mu_{\text{O}} < \mu_{\text{O}_2} \). For intermediate II-VI ratios, \( \mu_{\text{Zn}} < \mu_{\text{Zn(bulk)}} \) and \( \mu_{\text{O}} < \mu_{\text{O}_2} \). Since Zn and O are in equilibrium within ZnO, their chemical potentials are not independent necessitating \( \mu_{\text{Zn}} + \mu_{\text{O}} < \mu_{\text{ZnO}} \).

All native point defect formation energies in ZnO have been calculated [13]. Figure 1 summarizes the \( E_f' \)'s as a function of the Fermi energy \( E_F \) [13] for zinc-rich conditions including those for octahedral Zn_i, Zn antisites, ZnO, and zinc vacancies \( V_{\text{Zn}} \), in various charge states. For oxygen-rich conditions, \( E_f' \)'s as a function of \( E_F \) for octahedral O_i and O antisites, O_{Zn}, and V_{Zn} in various charge states are presented in Figure 2.

As the incorporation of donors or acceptors shifts the Fermi energy, spontaneous formation of these charged defects acts to compensate the prevailing conductivity in ZnO. The formation energy of Zn_i across most of the Fermi level range is high making them an unlikely self-compensating species in n-type material. However, in p-type material, the formation energies for charged species, \( Zn_{i+1} \) and \( Zn_{i+2} \) are more favorable, but the strain energy introduced by the self-interstitial limits their concentration. The absence of the charged \( V_{O}^+ \) and \( V_{O}^{2+} \) species in the energy diagram in Figure 2 reflects their high formation energy near the conduction band minimum (CBM). Since the formation energy of \( V_{O}^+ \) is always higher than either \( V_{O}^{2+} \)
Advances in Condensed Matter Physics

Figure 1: Calculated defect formation energy for selected defects as a function of the Fermi level and for $\mu_{Zn} = \mu_{Zn}^{\text{bulk}}$ (high zinc partial pressure). Only the lowest formation energy values are shown. The zero of the Fermi level is set to the top of the valence band. Reprinted with permission from [13]. Copyright 2000 by the American Physical Society.

Figure 2: Calculated defect formation energy for selected defects as a function of the Fermi level and for $\mu_{Zn} = \mu_{Zn}^{\text{bulk}} + \Delta E_{ZnO}^{\text{tot}}$ (low zinc partial pressure). Only the lowest formation energy values are shown. The zero of the Fermi level is set to the top of the valence band. Reprinted with permission from [13]. Copyright 2000 by the American Physical Society.

or $V_{O}$, the positive charge state is never thermodynamically stable [14]. However, by virtue of their low $E^f$'s, $V_{O}^{2+}$ or $V_{O}$ may be self-compensating species across the Fermi energy range.

Since hydrogen is unintentionally incorporated during MOVPE growth processes, it is also a plausible contributor to n-type conductivity in ZnO films. The formation energy of the hydrogen interstitial in charge state $q$ is defined by [14]

$$E^f (H^q) = E^{\text{tot}} (H^q) - E^{\text{tot}} (\text{bulk}) - \mu_H + qE_F,$$

where $E^f (H^q)$ is the total energy derived from a supercell calculation for the hydrogen interstitial, $E^{\text{tot}} (\text{bulk})$ is the total energy for a supercell containing only bulk ZnO, $\mu_H$ is the hydrogen chemical potential, and $E_F$ is the Fermi level which is set to 0 at the top of the valence band [14]. Defect species $H^+$ and $H^+$ are never stable in ZnO because they are at a much higher energy level in the band structure than $H^+ [14]$. The calculated formation energy for $H^+$ in ZnO is shown in Figure 3. The lowest energy position for $H$ is at the bond-center (BC) site with the antibonding (AB) position slightly higher in energy [14]. When $H$ is at the $AB_{Zn}$ site, the corresponding state is localized near the $H$ atom. When $H$ is at the BC and AB sites, the occupied state becomes an extended state because it is at a higher energy level [14].

The formation energy of interstitial $H$ shows that the solubility of $H^+$ is higher under p-type conditions than under n-type conditions. While that may seem deleterious to accomplishing p-type conductivity in ZnO, hydrogen does have the beneficial effect of increasing acceptor solubility and suppressing compensation by native defects in GaN [18]. Also, since $H^+$ prefers positions where the charge density is high and it remains close to the donated electron [18], it may easily form complexes with double acceptor species resulting in single acceptor complexes. Additionally, $H^+$ can be removed from the ZnO lattice during high temperature post-growth anneals.

3. Strategies for p-Type Doping

As mentioned above, there are three primary different approaches that have been investigated to generate appreciable hole conduction in ZnO at room temperature: substitution of Group IA impurities on the Zn sublattice, codoping of donors and acceptors, and substitution of Group VA impurities on the O sublattice. And multiple growth approaches that have been investigated to generate appreciable hole conduction in ZnO at room temperature: substitution of Group IA impurities on the Zn sublattice, codoping of donors and acceptors, and substitution of Group VA impurities on the O sublattice. Multiple growth approaches that have been investigated to generate appreciable hole conduction in ZnO at room temperature: substitution of Group IA impurities on the Zn sublattice, codoping of donors and acceptors, and substitution of Group VA impurities on the O sublattice. Multiple growth approaches that have been investigated to generate appreciable hole conduction in ZnO at room temperature: substitution of Group IA impurities on the Zn sublattice, codoping of donors and acceptors, and substitution of Group VA impurities on the O sublattice.
Table 1: Summary of room temperature Hall data for p-type ZnO.

<table>
<thead>
<tr>
<th>Growth technique</th>
<th>Structure</th>
<th>Structure a</th>
<th>Acceptor doping scheme</th>
<th>Hole concentration (10^{18} \text{ cm}^{-3})</th>
<th>Mobility (\text{cm}^2 \text{V}^{-1} \text{s}^{-1})</th>
<th>Resistivity (\Omega \text{-cm})</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVD</td>
<td>U</td>
<td>NH3 and excess Zn</td>
<td>0.015</td>
<td>12</td>
<td>34</td>
<td>Minegishi et al. [19]</td>
<td></td>
</tr>
<tr>
<td>MBE</td>
<td>SC</td>
<td>RF plasma N2</td>
<td>0.09</td>
<td>2</td>
<td>40</td>
<td>Look et al. [20]</td>
<td></td>
</tr>
<tr>
<td>RF diode sputtering</td>
<td>SC</td>
<td>Ga, N codoping</td>
<td>0.09</td>
<td>6</td>
<td>11.77</td>
<td>Singh et al. [21]</td>
<td></td>
</tr>
<tr>
<td>MBE (ECRO source)</td>
<td>U</td>
<td>Sh-doped + 800°C anneal</td>
<td>1.7</td>
<td>20</td>
<td>0.2</td>
<td>Xiu et al. [22]</td>
<td></td>
</tr>
<tr>
<td>DC reactive magnetron sputtering</td>
<td>PC b</td>
<td>Li-doped</td>
<td>0.14</td>
<td>2.65</td>
<td>16.4</td>
<td>Zeng et al. [23]</td>
<td></td>
</tr>
<tr>
<td>RF magnetron sputtering</td>
<td>SC</td>
<td>Ga and N codoping (N2 sputtering gas)</td>
<td>0.39</td>
<td>0.4</td>
<td>38</td>
<td>Kumar et al. [24]</td>
<td></td>
</tr>
<tr>
<td>Plasma assisted MOVPE</td>
<td>U</td>
<td>NO plasma as O and N source</td>
<td>2.29</td>
<td>1.59</td>
<td>1.72</td>
<td>Zeng et al. [25]</td>
<td></td>
</tr>
<tr>
<td>Sol-gel</td>
<td>PC</td>
<td>Al and N codoping + O anneal</td>
<td>0.08–0.2</td>
<td>1.6–1.73</td>
<td>19–45</td>
<td>Dutta et al. [26]</td>
<td></td>
</tr>
<tr>
<td>RF magnetron sputtering</td>
<td>SC</td>
<td>O and As dual implantation</td>
<td>6.8–19.8</td>
<td>19–32.9</td>
<td>0.05–0.55</td>
<td>Kim et al. [27]</td>
<td></td>
</tr>
<tr>
<td>MOVPE</td>
<td>U</td>
<td>P-doped</td>
<td>0.21</td>
<td>1.2</td>
<td>—</td>
<td>Du et al. [28]</td>
<td></td>
</tr>
<tr>
<td>PLD</td>
<td>PC</td>
<td>Na-doped + 254 nm illumination</td>
<td>0.21</td>
<td>7.9</td>
<td>3.8</td>
<td>Lin [29]</td>
<td></td>
</tr>
<tr>
<td>PLD</td>
<td>SC</td>
<td>N+ implantation + dynamic annealing</td>
<td>0.024–0.52</td>
<td>0.7–3.71</td>
<td>18–71</td>
<td>Myers et al. [30]</td>
<td></td>
</tr>
<tr>
<td>Plasma assisted MBE</td>
<td>U</td>
<td>Sh-doped + 800°C anneal</td>
<td>0.06</td>
<td>—</td>
<td>—</td>
<td>Huang et al. [31]</td>
<td></td>
</tr>
<tr>
<td>MOVPE</td>
<td>SC</td>
<td>As-doped via outdiffusion</td>
<td>0.35</td>
<td>0.79</td>
<td>—</td>
<td>Shi et al. [32]</td>
<td></td>
</tr>
<tr>
<td>Magnetron sputtering</td>
<td>U</td>
<td>P and N codoping + 800°C anneal</td>
<td>1.16</td>
<td>1.35</td>
<td>3.98</td>
<td>Sui et al. [33]</td>
<td></td>
</tr>
<tr>
<td>Plasma assisted MBE</td>
<td>SC</td>
<td>Na doping</td>
<td>0.002</td>
<td>0.4</td>
<td>8575</td>
<td>Ding et al. [34]</td>
<td></td>
</tr>
<tr>
<td>Spin coating</td>
<td>PC</td>
<td>N, Al codoping + double anneal in NH3 and N2</td>
<td>0.61</td>
<td>198.8</td>
<td>0.05</td>
<td>Kalyanaranan et al. [35]</td>
<td></td>
</tr>
<tr>
<td>MOVPE</td>
<td>SC</td>
<td>3% NO in N2</td>
<td>3.6</td>
<td>0.4</td>
<td>4.4</td>
<td>Reynolds et al. [36]</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)U: undefined; SC: single crystal; PC: polycrystalline. All are thin films.
\(^b\)(002) preferential orientation.

that mobility values indicate that the thin films are quite compensated. A summary of the recent data in chronological order is given in Table 1. The fundamental goal is significant p-type conduction at room temperature, which implies that one must achieve shallow acceptor ionization energies with minimal compensation by unintentional impurities and/or defects.

3.1. Group IA Li and Na on a Zinc Site. While Park et al. [37] have predicted Li and Na to exhibit relatively shallow, \(~0.1–0.17\text{ eV}\), acceptor levels in ZnO, a more recent hybrid density functional investigation [38] has shown the ionization energy to be \(~0.6–1.1\text{ eV}\) above the valence band maximum, that is, becomes significantly deeper. The latter attributed this discrepancy to the well-known band gap error in DFT and its influence on the energy of the \(\text{Li}_{\text{Zn}}\) acceptor level. As the band gap widens, the acceptor level shifts upward away from the VBM due to becoming more localized. However, there have been limited reports of successful acceptor behavior that is based on Group IA doping. Using DC reactive magnetron sputtering, Zeng et al. [23] have demonstrated hole concentrations of \(~10^{17} \text{ cm}^{-3}\) at a growth
temperature of 550°C in Li-doped thin films; however, the acceptor concentration decreases an order of magnitude as the growth temperature decreases or increases 50°C on either side of the optimum. They also suggest that Li\(_{2}\)Zn may be relatively unstable because of a higher Madelung energy. As we will see later, a lowered Madelung energy is suggested as being fundamental to understanding why codoping should result in p-type behavior. For Na-doped ZnO, p-type ZnO has been reported for thin films grown by PLD [29] and plasma assisted MBE [34]. In the former, it was suggested that absorbed O species at grain boundaries are deleterious to p-type behavior, and thus, it was necessary to subject films to 254 nm UV illumination to remove these species to obtain a p-type film. Nevertheless, they were able to achieve a mobility of 7.9 cm\(^2\) V\(^{-1}\) s\(^{-1}\) at a hole concentration of 2.1 \(\times\) 10\(^{17}\) cm\(^{-3}\), and the acceptor ionization energy was estimated to be 110 meV from temperature-dependent Hall data. For the plasma assisted MBE ZnO, nonpolar a-plane films were grown on r-plane sapphire. The relatively low hole concentration of \(\sim 2 \times 10^{15}\) cm\(^{-3}\) was attributed to compensation by oxygen vacancies. Such weak p-type behavior is not anticipated to yield substantial hole conduction at room temperature. In general, the low mobilities observed in these investigations indicate that the films are most likely compensated. Group IA elements are relatively reactive, and thus we expect that anomalous p-n junction behavior may be an issue as observed for Group IIA dopants in traditional III-V semiconductors.

Although codoping of donors and acceptors has been predicted to lead to improved p-type conduction via the presence of shallow acceptor levels in comparison to the isolated nitrogen on an oxygen site, experimentally, relatively low hole concentrations, \(<\sim\) mid-10\(^{17}\) cm\(^{-3}\), with low mobilities are observed (see Table I) [21, 24, 26, 35]. Singh et al. [21] demonstrated p-type ZnO thin films for (Ga, N) codoped films grown by RF diode sputtering but found that the observation of p-type behavior depended upon the oxygen partial pressure to total pressure ratio. For films grown with less than 50% O partial pressure, n-type conduction was observed, but films grown with 50% and above exhibited p-type behavior that appeared to saturate at 60% O partial pressure. P-type conduction was attributed to suppression of \(V_0\) and \(Zn_i\) and the associated compensation with codoping. In addition, films exhibiting p-type behavior revealed both (002) and (100) reflections in X-ray diffraction (XRD) patterns. Most likely, this is related to N incorporation in the films as we will discuss in the next section for substitutional N on an O site. They claim a reduction in band gap to 3.27 eV for p-type films; however, this emission peak is most likely related to an acceptor-related transition [21]. Using RF magnetron sputtering, Kumar et al. [24] have reported p-type behavior for (Ga, N) codoped films grown on both sapphire and Si in which \(N_2O\) was used as the sputtering gas. For films grown below 450°C, conduction was n-type. At a growth temperature of 550°C, they achieved a resistivity of 38 \(\Omega\)-cm with a hole concentration of 3.9 \(\times\) 10\(^{17}\) cm\(^{-3}\). Somewhat puzzling though is the 1.3 \(\times\) 10\(^{19}\) cm\(^{-3}\) hole concentration reported for films grown on Si. In spite of the significant difference in misfit and thermal expansion coefficients between ZnO and sapphire and between ZnO and Si (see, e.g., [6]) and resultant defect densities, it is most likely that the high hole concentration is not representative of the film but is indicative of an issue with the Hall measurements on the p-type substrate. Although these authors claim that only (002) and (004) diffraction peaks are observed in the XRD spectra for films grown on both substrates, close examination of the spectrum on Si reveals a (101) ZnO reflection. As we will discuss in the next subsection, we believe that this reflection is associated with N incorporation in the film.

Other codoped films have utilized (Al, N) codoping and have been deposited via sol-gel [26] and spin coat [35] processes. Dutta et al. [26] reported that films that were only N-doped exhibited mixed n and p conduction while the Al and N codoped films were stable p-type with hole concentrations of (0.8–2) \(\times\) 10\(^{17}\) cm\(^{-3}\) that were dependent upon the Zn/N/Al ratios. All films were annealed in an oxygen ambient. While their undoped and solely N-doped films were polycrystalline, their codoped (Al, N) ones exhibited a strong (002) diffraction peak with much weaker (100) and (101) peaks. We believe that presence of the latter weak reflection is related to N incorporation. Grain sizes were in the range 25–75 nm. Although the (Al, N) codoped ZnO films grown by spin coating were polycrystalline with grain sizes \(<25\) nm, they exhibited reasonably high hole concentrations, 6 \(\times\) 10\(^{17}\) cm\(^{-3}\), and extraordinarily high mobility, 198.8 cm\(^2\) V\(^{-1}\) s\(^{-1}\) [35]. This is the highest mobility that we
have been reported for holes in ZnO and seems unreasonable considering the polycrystalline nature of the films. A 550°C anneal in NH₃ served as the N-dopant source, and films were subjected to another post-growth anneal at 700°C to remove H. This last step is crucial as it has been suggested that H is the dominant donor in ZnO [42, 43]. Comparison of Ga and N versus Al and N codoped films enables one to make an interesting observation that seems counter to first-principles DFT calculations [40]. For (Ga, N) codoped films, the reported [21, 24] hole concentrations are in the range of 0.9–3.9 × 10¹⁷ cm⁻³, whereas those [26, 35] codoped with Al and N are 0.8–6 × 10¹⁷ cm⁻³; that is, there appears to be no significant difference between codoping with Ga or Al and N with regard to N dopant solubility and/or acceptor ionization energy. As mentioned in the beginning of this subsection, DFT calculations implied that Ga and N codoping should increase N solubility compared to Al and N codoping because the interaction between (Al-2N) and a neighboring N on an oxygen site is weak in comparison to that between two (Ga-N) complexes and the binding between (Ga-2N) and a nearest neighbor N [40]. A more recent investigation [44] on p-type conduction for N and Te codoped ZnO films has reinforced the proposed role [39] of codoping for generating p-type behavior; that is, N/Te codoping lowers the Madelung energy for enhanced N incorporation. However, it was necessary to subject the codoped films to a 700°C/30 min anneal in an O₂ environment in order to reduce donor-related defects and enable n- to p-type conversion in the films. However, the luminescence data are somewhat puzzling with regard to the relative importance of N₂ flow and anneal related to the Hall data (see Figure 6 in Park et al. [44] in relation to carrier concentration data in Table 1). An added benefit of the thermal anneal step is that it improves crystal quality that had been degraded by codoping as evidenced by the increase in the FWHM of the (0002) X-ray reflection with increasing N₂ flow. Their highest reported hole concentration of 1.6 × 10¹⁸ cm⁻³ after the anneal occurred for the highest N2 flow and also exhibited a mobility of 1.8 cm² V⁻¹ s⁻¹, the latter implying that the films are heavily compensated.

3.3. Group VA Elements on an Oxygen Site. Although the Group VA elements (N, P, As, and Sb) have predicted acceptor ionization energies greater than those of the Group IA elements on the zinc site, the Group VAs have been the most extensively investigated. Of the Group VA elements, nitrogen seems to be the most suitable because of its electronic structure, and its ionic radius, 0.168nm, is closer to that of oxygen, 0.138 nm, than the other elements for which the ionic radii differences are >50% [3]. One might anticipate that such a larger ionic radius compared to that of oxygen would introduce considerable strain into the wurtzite lattice. More importantly, the ionization energies are predicted to be >-1 eV for P, As, and Sb [45]. However, an acceptor ionization energy of 197–227 meV has been reported for Sb-doped ZnO grown by MBE [22]. Substitutional nitrogen on an oxygen site alone is particularly interesting. Ionization energies of 0.33 [40], -0.4 [37, 46], and 1.3 eV [47] have been predicted, none of which would be compatible with appreciable hole conduction at room temperature. Nevertheless, on the basis of Hall measurements, p-type conduction has been reported for nitrogen and the other Group VA dopant elements. Reported ionization energies for N₂O are in the range of 100–200 meV (see Table 2 for a summary of the ionization energies). In view of these results, it seems clear that complexes involving nitrogen are mainly responsible for p-type behavior as we will discuss below. We begin first though with a brief summary of p-type conduction for P-, As-, and Sb-doped ZnO. Recall that a summary of relatively recent room temperature Hall data are given in Table 1.

P-type behavior for P-doped ZnO has been reported for films grown by MOVPE [28] and magnetron sputtering [33]. In the latter reference, Sui et al. demonstrated that P and N codoping provided an order of magnitude increase to the hole concentration compared to P-doping alone. They attributed this to the addition of nitrogen to a neutral P₂N₃O₂ complex to form a P₂N₃O₂ acceptor complex, which gives rise to formation of an impurity band above the valence band maximum similar to the discussion above for donor-acceptor codoping. Their optimum hole concentration of 1.16 × 10¹⁸ cm⁻³ was achieved after a post-growth anneal at 800°C for 30 min at a pressure of 10⁻³ Pa. The low mobility of 1.35 cm² V⁻¹ s⁻¹ implies that the films are heavily compensated. Du et al. [28] confirmed their p-type Hall data with C-V measurements and claimed their films to be stable for four months. Of particular relevance in the latter investigation is that they were able to demonstrate lasing under electrical pumping for a P-doped p-ZnO/n-GaN heterostructure; the FWHM of the electroluminescence spectrum narrowed to ≤1 nm at 9 mA and above. However, other resistivity and luminescence data have suggested that phosphorous doping for p-type behavior in ZnO is more problematic. Von Wencskernst et al. [50] have grown P-doped ZnO heteroepitaxially and homoepitaxially and found quite varying results. For example, as-grown heteroepitaxial films were semi-insulating to n-type, whereas scanning capacitance microscopy of homoepitaxial films indicated regions of mixed carrier type; both results are consistent with the predictions of Park et al. [37], based on the amphoteric behavior of P in ZnO.

As-doped films have been achieved via MOVPE [32] and (O, As) dual implantation [27] into a film that was grown by magnetron sputtering. In the former, Shi et al. [32]...
formed p-ZnO by As outdiffusion from a neighboring GaAs layer. They reported a hole concentration of $3.56 \times 10^{12} \text{ cm}^{-3}$, which was attributed to formation of an $\text{As}_{2}\text{Zn}_{2} \rightarrow 2\text{V}_{\text{Zn}}$ complex, and also demonstrated electroluminescence from a p-n junction light emitting diode type structure based on As-doped ZnO. For the implanted films [27], mixed conduction was observed for As implantation alone, whereas dual implantation followed by an 800°C anneal in N$_2$ resulted in hole concentrations $\sim 10^{19} \text{ cm}^{-3}$ with relatively high hole mobilities that depended upon the fluence of As and O. The observation of hole conduction in As-doped ZnO however is contrary to the predictions of Park et al. [37], who based their analysis on the amphoteric nature of these dopants. In spite of the large ionic radii size difference between Sb and O, hole conduction has been observed in Sb-doped ZnO [31, 45]. For the Group VA elements P, As, and Sb, Xiu et al. have reported the highest hole concentration, $1.7 \times 10^{18} \text{ cm}^{-3}$, with a corresponding mobility of $20 \text{ cm}^{2} \text{ V}^{-1} \text{ s}^{-1}$ for Sb-doped ZnO [22]. Their films were grown by MBE with elemental Zn and Sb sources and an oxygen plasma electron-cyclotron resonance source for the oxygen. Growth was followed by an 800°C in situ anneal; the hole concentration increased with the Sb effusion cell temperature. Huang et al. [31] used plasma assisted MBE to grow their films and also required an 800°C in situ anneal to activate their Sb. Earlier though, Friedrich et al. [51] reported a deterioration of the sample surface with an increase in the Sb concentration due to the increase in lattice stress. They claim that there is a tendency for Sb-O precipitate formation, which suppresses formation of $\text{Sb}_{2}\text{Zn}_{2} \rightarrow 2\text{V}_{\text{Zn}}$ complexes for p-type behavior. More recently, it was demonstrated that the conduction type in Sb-doped ZnO depended upon the concentration of Sb in films grown by plasma enhanced MBE [52]. Liu et al. demonstrated that films with Sb concentrations in the doping regime (varied by changing the Sb flux) exhibited n-type behavior for electron concentrations in the $10^{18}$-mid-$10^{19} \text{ cm}^{-3}$ range while films with $\sim 1$ at % Sb revealed compensating acceptor-like defects with a resultant decrease in electron concentration. It was suggested that Sb incorporates on the Zn site for relatively low Sb concentrations but the observed sudden increase in the c-lattice parameter suggested that Sb begins to reside on the O site at higher concentrations, $\sim 1$ at % [52]. Structural degradation of the films was also observed to occur with increasing Sb as also reported by Friedrich et al. [51] above. Yankovich et al. [53] have reported that stable acceptor conduction is observed in Sb-doped ZnO nanowires in which Sb decorates basal plane inversion domain boundaries. This electron acceptor behavior was related to the phenomenon of Sb and O codoping and shown to be consistent with density functional theory calculations. This suggests however that acceptor behavior via Sb-doping is not necessarily simple formation of an isolated acceptor level but involves defect complexes within the material, which we believe to be the relevant mechanism for appreciable room temperature hole conduction.

There have been considerable calculations on substitu- tional Group VA impurities, mostly nitrogen, on the oxygen site. Although the magnitude of the ionization energies varies from 0.33 to 1.3 eV, all are consistent with a deep acceptor for a single N$_O$. This strongly suggests that the experimental observation of p-type conduction in N-doped ZnO is predominantly related to complexes involving nitrogen. As mentioned above, Park et al. [37] performed first principles DFT calculations using the pseudopotential method within the LDA [15]. Since their calculations included a nonlinear partial core correction (NPC) [54], they gave a 2x correction of the bandgap for ZnO compared to those that only included the Zn-3d states as valence electrons. Their underestimate of the bandgap, 1.56 eV, is a characteristic error in the LDA approach and their 25% underestimate of the calculated heat of formation of ZnO follows. However, their calculations do reveal the trend for the increase in defect energy levels when Group VA elements are substituted on the O site compared to Group IA defects substituted on the Zn site. Since Group IA elements lack an active d orbital, they have reduced $p-d$ coupling which lowers their defect energies. The higher defect energy levels for Group VA elements are a result of increasing impurity $p$-orbital energy from N $\rightarrow$ P $\rightarrow$ As [37]. The key point revealed by their calculations was the importance of AX centers, which are deep defect complexes that compensate for acceptors. The positive energy required to form positively charged AX centers from the substitutional acceptors on a Zn site indicates that Group IA elements are only metastable in ZnO [37]. The negative energy for Group VA elements, P and As, means they are more stable as AX centers than substitutional states on the O site [37]. Their calculations thus suggested that only N is a viable substitutional acceptor in ZnO. Another consequence of their calculations is the formation of complexes between Zn$_n$ and N$_{nO}$ as a function of growth environment. For example, the formation energy $E_f$ of (Zn$_n$ + N$_{nO}$) under O-rich conditions is 6.58 eV but can be as low as 0.8 eV at the Zn-rich limit [37].

Lee et al. [46] also used first principles pseudopotential calculations within the LDA to describe the localized nature of the Zn 3d and O 2p wave functions to explain compensation mechanisms in N-doped ZnO. The conclusion from Lee’s calculations mirrors others [13, 14] in that $\text{V}_{\text{Zn}}$ is the most stable defect under O-rich conditions from midgap to the CBM. While $\text{V}_{\text{Zn}}$ acts as an acceptor, his focus is on the compensating species such as $\text{V}_{\text{O}}$, Zn$_n$, Zn$_{nO}$, and the shallow double donor, (N$_2$)$_{O}^{-2+}$, as $E_f$ changes from the CBM to the VBM. The low formation energies for complexes of those defects with N$_O$ suggest that N-doping is inefficient under O-rich conditions because the total atomic concentration, [N], of incorporated N impurities is below $8 \times 10^{12} \text{ cm}^{-3}$ which leads to very low hole densities [46]. In the Zn-rich limit however, since the $E_f$ of N$_O$ decreases significantly, both hole and N concentrations increase, leading to [N] $\sim 10^{15}$-2 $\times 10^{17} \text{ cm}^{-3}$ with resultant hole carrier densities that saturate at $\sim 2 \times 10^{15} \text{ cm}^{-3}$. For growth processes using a N$_2$ source in an electron cyclotron resonance, ECR, and plasma, N solubility is expected to increase under O-rich conditions. However, (N$_2$)$_{O}^{-2+}$ molecules are the dominant compensating species. So even though [N] increases with this active source, the hole carrier density decreases because of compensation by (N$_2$)$_{O}^{-2+}$ molecules and unintentional H impurities from H$_2$O within the growth system. Lee et al. [46] concludes that for low [N],
N_2O acceptors are compensated by V_0, while for high [N], N_2O acceptors are compensated by N_2O-ZnO complexes.

Troubled by acceptor levels for N in ZnO generated by DFT calculations within the LDA or generalized gradient approximation (GGA) [37, 46], Lyons et al. [47] performed first principles calculations using hybrid functionals. These calculations lead to a band gap of ZnO of 3.4 eV. They found that N_2O can be stable in either the neutral or −1 charge states, with the acceptor level occurring at 1.3 eV above the VBM [47], as mentioned above. The discrepancy between their calculations and those of Park et al. [37] and Lee et al. [46] in which N is predicted to be a more shallow acceptor were attributed to the downward shift of the VBM on an absolute energy scale. They verified their calculations for N_2O in ZnO by repeating them for N_sc in ZnSe. Their calculations were consistent with experimental findings that the ionization energy for N_sc is 100 meV [47]. They assigned the difference in ionization energies of the N acceptors to the band structure of the host as follows. In ZnO, the VBM is derived from the O 2p orbitals. In ZnSe, the VBM is derived from the Se 4p orbitals. Since N 2p orbitals are ~3 eV lower than Se 4p orbitals in ZnSe, but ~3 eV higher than O 2p orbitals in ZnO, it follows that the transfer of an electron from the VBM to the N acceptor is more energetically favorable in ZnSe, but nearly impossible in ZnO [47]. Experimental confirmation of N as a deep acceptor in bulk ZnO grown in an NH_3 ambient was reported by Tarun et al. [55] based on two key observations: (1) an IR absorption peak at 3148 cm\(^{-1}\) attributed to a N-H complex and (2) a broad PL emission centered at 1.7 eV, the intensity of which increased with the activated N concentration. Activation of N as a deep acceptor was accomplished by annealing in a 0.5 atm O_2 ambient, which dissociated N-H pairs to form the isolated N. Presence of the broad 1.7 eV emission was claimed to be in agreement with the deep acceptor model [47] for an isolated substitutional N in ZnO proposed by Lyons, et al. Nevertheless, numerous publications have reported N-related relatively shallow acceptor levels (see Table 2) that are consistent with p-type conduction. As discussed below, these shallow acceptor levels are related to defect complexes involving N_2O and not an isolated N on the oxygen site.

Minegishi et al. [19] was the first to realize p-type conduction at room temperature in ZnO films by incorporating N during chemical vapor deposition (CVD) of ZnO on (0001) sapphire substrates. Secondary ion mass spectroscopy (SIMS) was used to confirm the presence, but not absolute concentration, of N in those films given the simultaneous addition of NH_3 to the H carrier gas and a 10 mol% mixture of metallic Zn and ZnO powder. Their N-doped films exhibited resistivities that ranged from 34 to 175 \(\Omega\)-cm with Hall mobilities of 12–30 cm\(^2\) V\(^{-1}\) s\(^{-1}\) for substrate temperatures of 650–750°C, respectively. However, there was a very narrow temperature window in which the excess Zn was able to catalyze N to combine with H in the film to form ZnNH\(_2\), resulting in the activation of the N acceptor [19]. For the film that did invert to p-type conduction, the acceptor ionization energy was estimated to be 100 meV which is the lowest reported to date (see Table 2) and seems low for the N incorporation suggested by a hole carrier density of only 1.5 \(\times\) 10\(^{16}\) cm\(^{-3}\). If one assumes that their estimate of the ionization energy is correct, then one can only conclude that their nitrogen incorporation is quite low.

Probably the results that have stimulated the most renewed interest in ZnO were those by Look et al. [20]. They were the first to show reproducible N-doped p-type ZnO grown by MBE on Li-doped semi-insulating ZnO substrates by adding N_2 to the O_2 gas flow in the RF plasma source. SIMS measurements revealed the nitrogen concentration, [N], at the surface of 8 \(\times\) 10\(^{18}\) cm\(^{-3}\) with a corresponding hole concentration of 9 \(\times\) 10\(^{16}\) cm\(^{-3}\), which implied a 1% activation of the N acceptor. Van der Pauw Hall measurements also gave average values of \(\rho = 40 \Omega\)-cm and \(\mu_p = 2 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}\). Their low temperature (4 K) PL reported the acceptor-bound exciton (A\(^{0}\)X) associated with N_2O at 3.315 eV, and the acceptor ionization energy was estimated to be 0.17–0.20 eV [20]. Their hole concentrations were consistent with the estimated ionization energy and were in agreement with those predicted by the nondegenerate, single-donor/single-acceptor model [56]. SIMS, Hall, and PL measurements were also consistent with those for p-type GaN and other p-type II-VI compounds.

Zeng et al. [25] grew N-doped ZnO thin films on a-plane (II–20) sapphire substrates by plasma-assisted low-pressure (5 Pa) MOVPE. An NO plasma was used as the source for both the oxygen source and the N dopant source. Growth temperatures ranged from 250 to 500°C in 50°C increments. All films were grown at 450°C and below exhibited p-type behavior. Room temperature (RT) van der Pauw Hall characteristics were optimal for a growth temperature of 400°C. Resistivity was lowest (1.72 \(\Omega\)-cm) as was hole mobility (1.59 cm\(^2\) V\(^{-1}\) s\(^{-1}\)). Hole concentrations were also maximized at 2.29 \(\times\) 10\(^{18}\) cm\(^{-3}\). Scanning electron microscopy (SEM) of surfaces of films grown at 300°C, 400°C, and 500°C supported Zeng et al.'s [25] assertion that less N is incorporated at higher temperatures as did the relative intensities from RT photoluminescence. The free electron-to-neutral-acceptor (e, A\(^0\)) transition was most pronounced for their 400°C film. From the (e, A\(^0\)) peak position, the acceptor ionization energy was estimated to be 180 meV. The very impressive aspect of this study was the wide range of temperature and the repetitive consistency with which p-type conductivity was achieved.

As mentioned in the codoped section above, Dutta et al. [26] used a sol gel process to investigate both N-doped and (Al, N) codoped ZnO. The sol-gel film was spin coated, post-baked, and then heated at 550°C in an oxygen ambient for 30 min. The decomposition pathway of the ammonium acetate gave NO and NO\(_2\) to act as the N source. X-ray diffraction (XRD) patterns of their ZnO:N showed (002) as well as (101) plane reflections. However, their N-doped only films showed unstable behavior. Hall measurements fluctuated between p- and n-type with concentrations and mobilities in the range of (−)6.53 \(\times\) 10\(^{13}\) cm\(^{-3}\) to (+)5.95 \(\times\) 10\(^{14}\) cm\(^{-3}\) and 66.5 cm\(^2\) V\(^{-1}\) s\(^{-1}\) to 6.4 cm\(^2\) V\(^{-1}\) s\(^{-1}\). The anomalous Hall behavior is consistent with their low XRD c-axis value of 0.5218 nm suggesting <0.1 atomic % N in the film [57].

Myers et al. [30] used PLD and ion implantation of N\(^+\) ions at three different fluences 3 \(\times\) 10\(^{14}\) cm\(^{-2}\), 6 \(\times\) 10\(^{14}\) cm\(^{-2}\),
and $1.2 \times 10^{15} \text{ cm}^{-2}$ and dynamic annealing at 300, 380, and 460°C for each fluence. Simulations predicted maximum implanted concentration of $8 \times 10^{19} \text{ cm}^{-3}$ at a depth of ~120 nm for the highest implant fluence of $1.2 \times 10^{15} \text{ cm}^{-2}$. All samples implanted at 300 and 380°C were n-type except for the lowest fluence at 380°C. However, samples for all fluences implanted at 460°C exhibited p-type conductivity and their resistivity decreased from $71 \Omega \cdot \text{cm} \rightarrow 50 \Omega \cdot \text{cm} \rightarrow 18 \Omega \cdot \text{cm}$ as fluence increased. As resistivity decreased, hole carrier concentration increased from $2.4 \times 10^{16} \text{ cm}^{-3} \rightarrow 1.9 \times 10^{17} \text{ cm}^{-3} \rightarrow 2.4 \times 10^{17} \text{ cm}^{-3}$. Hall mobilities ranged from 0.7 to 3.7 cm$^2$ V$^{-1}$ s$^{-1}$. Transmission electron microscopy (TEM) of films that exhibited n-type conductivity confirmed that low temperature implantation created defect clusters and damage instead of substitutional incorporation of dopants. Conversely, stacking faults were a characteristic of all the p-type samples. Based on their simulations, at an implant temperature of 460°C, they were able to activate 0.3% of the implanted N$^+$ ions.

Ion implantation provides a means of incorporating controllable amounts of dopant into the ZnO matrix to help elucidate the relevant mechanisms for p-type conduction. Recently, Stehr et al. [49] used this technique to maximize the formation of certain intrinsic defects by coimplanting N$^+$ with O$^+$ or Zn$^+$ into bulk ZnO crystals at 300 K. Following implantation, crystals were annealed in either N$_2$ or O$_2$ at 800°C for 2 min. Raman spectra of their crystals contained local vibration modes at 277 cm$^{-1}$, 511 cm$^{-1}$, and 581 cm$^{-1}$ which have been previously associated with N [58, 59]. The PL spectra of their codoped ion implanted samples revealed three major points. First, the $I_{44}$ line normally present for undoped ZnO disappeared after doping implantation with either N or O or N and Zn. They attributed this to consumption of H by the formation of complexes with N in the implanted ZnO. Secondly, PL lines at 3.3128 and 3.2405 eV appeared after implantation. These lines have been assigned to the recombination of excitons bound to a neutral acceptor ($A^0X$) [20, 60] and also to free electron to acceptor (FA) transitions [61]. The presence of these PL signatures is consistent with the observations of Look et al. [20]; however, the interpretation of the origin, single N$_0$ versus an acceptor complex, remains quite controversial. Stehr et al. [49] suggested that this defect is most likely a complex that contains a N atom because the acceptor energies estimated by Stehr et al. [49] and Look et al. [20] are too small to be substitutional nitrogen on the oxygen site. Thirdly, the intensity of the A$^0X$ emission for ZnO samples implanted with N and Zn was highest for those annealed in O$_2$ and lowest for those annealed in N$_2$. This annealing behavior suggested that while defect formation was favored under Zn-rich conditions, it was also suppressed by oxygen deficiency [49]. For those samples coimplanted with N and O, a PL band at 3.23 eV accompanied by an LO-assisted transition was assigned as the donor to acceptor pair (DAP) transitions related to a nitrogen acceptor. Assuming a shallow (52 meV) donor is participating in the DAP transitions, the ionization energy (160 meV) suggests this N-related acceptor is also a complex defect. This binding energy is similar to the energy level of a N$_{2\text{Zn}}$-2V$_{\text{Zn}}$ complex seen in Zn-deficient conditions [49]. Their optically detected magnetic resonance (ODMR) measurements of samples with coimplantation of N and O gave insight about another defect. When studied in the visible spectrum, the ODMR signal for samples annealed in N$_2$ was broad and its intensity was diminished relative to those annealed in O$_2$ or the reference ZnO crystal. The angular dependence of that broad signal suggested the defect was related to the deep donor, N$_{2\text{Zn}}$, which has favorable formation energy under oxygen-rich conditions [62]. However, when studied under near-IR emission, that same deep donor signal was detected but was paired with an additional defect, a deep acceptor N$_0$. These same defects were not present for samples coimplanted with N and Zn. These ODMR signals present the first evidence for the N antisite that Liu et al. [63] suggested and is fundamental to the work done by Reynolds et al. [36].

Sui et al. [33] codoped ZnO films with Group VA elements P and N by magnetron sputtering and post-growth annealing to generate p-type material. Their source material was a mixture of argon and nitrogen used to (RF) sputter ZnO and 2 wt.% P$_2$O$_5$ powders onto quartz substrates at 500°C. Post-growth annealing was performed in a tube furnace for 30 min at 800°C. Their RT resistivity was found to be 3.98 Ω-cm with a Hall mobility of 1.35 cm$^2$ V$^{-1}$ s$^{-1}$ and a hole concentration of 1.16 × 10$^{18}$ cm$^{-3}$. The electrical properties for ZnO: (P, N) were greatly improved, 10x and 100x, over those samples doped with P or N alone, respectively. The ZnO:P and ZnO:N resistivities were an order of magnitude lower and their carrier concentrations were 2.3 × 10$^{17}$ cm$^{-3}$ and 1.18 × 10$^{16}$ cm$^{-3}$, respectively. The higher mobility (10.8 cm$^2$ V$^{-1}$ s$^{-1}$) for the ZnO:N sample is consistent with the lower carrier concentration. All samples showed p-type conductivity. Sui et al. [33] also presented $I$-$V$ characteristics for their homojunction of undoped n-type ZnO and p-type ZnO, which used Ni/Au for the p-contacts and In for the n-contacts. The rectifying behavior is clearly present albeit the current is low ($-2 \mu A < I < 2 \mu A$) for a drive voltage from $-15 V < V < 15 V$. Using SIMS, they compared the N incorporation in a film that was codoped with N and P with one doped with N alone. While no absolute concentrations were given, the relative incorporation of N is an order of magnitude higher in the (P, N) codoped sample. Their P$_2$O$_5$ XPS spectra gave a binding energy of between 133.3 and 133.8 meV for the ZnO: (P, N) sample indicating that P did not substitute for O but rather occupied a Zn site. The N$_{1s}$ XPS spectrum showed two peaks, one at 397.6 eV and another at 402.5 eV. The lower peak indicates that the N atom substitutes for the O atom to form the N$_0$ acceptor in the ZnO: (P, N) sample. The higher, less intense peak is attributed to (N$_2$)$_{13}$ which is considered a double donor. Their PL data over the temperature range from 83 to 300 K showed the FA transition at 3.310 eV, a DAP transition at 3.241 eV, and its longitudinal optical (LO) phonon replica separated by 72 meV at 3.168 eV. Their FA peak exhibits the characteristic redshift while the DAP emission continually shows a blueshift as shallow donors are ionized with increasing temperature as the ionized free electrons in the conduction band prefer to recombine with acceptors to form FA [64]. The formation
mechanism for the p-type ZnO: (P, N) is that P occupies the Zn site, $P_{Zn}$, and N occupies the O site ($N_{O}$) forming a neutral passive $P_{Zn}$-$3N_{O}$ complex which may form an additional fully occupied impurity band above the VBM. When additional N is introduced into the system with the $P_{Zn}$-$3N_{O}$ defect complex, the N and $P_{Zn}$-$3N_{O}$ combine to form more energetically favored $P_{Zn}$-$4N_{O}$ complex acceptors. As a consequence, the electrons transit from the impurity band which lowers the ionization energy, and the p-conduction arises from the $P_{Zn}$-$4N_{O}$ acceptor complex.

The fact that low experimental acceptor ionization energies for Group VA on an oxygen site have been reported suggests that the origin of p-conductivity in ZnO is most likely not associated with a single VA substitution on an O site but must be explained by a more complex and multistep approach. Catlow et al. [65] showed that the fundamental reactions necessary to generate defect species $V_{Zn}^{-2}, V_{O}^{+2}, O_{2}^{-}, Zn_{2}^{+2}$, and hole formation require energy while reactions for electron generation and ZnO(s) formation release energy. So, from a thermodynamic perspective, except under conditions of very high chemical potential, defect compensation is more favorable than hole compensation under equilibrium conditions [65]. However, experimental data suggests that by controlling the growth conditions, defect solubility may be coerced such that donor defect species provide metastable routes to acceptor complexes capable of providing mobile holes in ZnO. Liu et al. [63] suggested that since the absorption energy for $N_{Zn}$ is 0.08 eV for the hexagonal close packed structure of the Zn-polar surface, it may be possible for N to absorb to bulk Zn sites, resulting in $N_{Zn}$ bonded to 3 neighboring O atoms on the zinc-polar oxygen surface. This suggestion is consistent with the formation energies for $V_{Zn}^{-2}$ under O-rich conditions as shown in Figure 2. But it necessitates the amphoteric behavior of N in ZnO under O-rich conditions leading to the formation of $N_{Zn}$ antisites.

Reynolds et al. [36] suggested that if the MOVPE growth conditions were alternated from O-rich to Zn-rich, it would be energetically less favorable for O to bond to the $N_{Zn}$ than to form an $V_{O}^{0}$ since its $E_{f}$ is $\sim 0.1$ eV across the entire Fermi level range. The result would be the formation of $N_{Zn}V_{O}$ which according to Liu et al. [63] are metastable double donors. By repeating the O-rich then Zn-rich growth environment in a cyclic manner, ZnO films containing high concentrations of $N_{Zn}V_{O}$ can be achieved. During the cool down after growth, the metastable nature of these donors requires an input of energy via an in situ anneal that must satisfy two criteria. The first is that the ambient gas must create sufficient overpressure to minimize nitrogen from leaving the film. Secondly, the thermal energy must be sufficient to promote the N on the Zn site to hop to the adjacent vacant O site, the result of which is to convert the double donor complex, $N_{Zn}V_{O}$, to a double acceptor complex, $V_{Zn}$-$N_{O}$. The first condition can be met with an 450°C in situ anneal in N$_2$O since the change in the Gibbs free energy, $\Delta G$, favors formation of NO and N rather than $N_{2}$. Since the thermal energy will also break bonds between H and native defects, H$^{-}$ will diffuse through the lattice to compensate other defects. Reynolds et al. [36] used Raman spectroscopy to reveal a local vibration mode consistent with a $V_{Zn}$-$N_{O}$-H$^{-}$ single acceptor complex at a vibration frequency of 3078 cm$^{-1}$ after the in situ anneal (black) as well as after an ex situ 800°C O$_2$ anneal.
for more than one year. They attributed this instability for ZnO on c-plane sapphire to disappearance of nitrogen on the oxygen site due to compressive stresses associated with lattice misfit. Low temperature (11.6 K) PL studies of a N-doped film exhibiting RT p-type conductivity, shown in Figure 7, determined that the dominant peak is the neutral donor bound exciton (D\(^0\)X) transition at 3.361 eV \[68\] and the FA transition at 3.314 eV \[69\] that undergoes a continuous redshift with increasing temperature. The acceptor ionization energy assigned (using the method described by Wang and Giles \[48\]) to the \(\text{V}_{\text{Zn}}\)-\(\text{N}_\text{O}^{-}\)-H\(^+\) complex is 134 meV, which is sufficiently low to allow appreciable room temperature hole conduction as we reported. The structural quality and N incorporation in our films can be assessed by X-ray diffraction scans. Figure 8 compares patterns for a N-doped film to a nominally undoped one for ZnO grown on sapphire. Both the sapphire (0006) and (0002) and (0004) reflections for ZnO are observed. Furthermore, the N-doped sample also exhibits additional (10-11) and (20-22) reflections. In all samples investigated thus far, the FWHM of the (0002) and (0004) reflections for N-doped material are slightly wider, for example, 13.0 min versus 11.5 min for nominally undoped. Using the Scherrer formula for X-ray broadening, we estimate the average crystallite size to be 50.2 nm for nominally undoped material and 44.5 nm for N-doped films. On the basis of the d-spacing for the (0004) reflection (see Figure 3.3 in \[7\]), we conclude that our alternating O-rich to Zn-rich growth scheme enables us to incorporate \(\sim 0.3–0.5\) at % N in the films. Many have reported inferior crystalline quality of heavily N-doped films, which has led some to suggest that these films are actually polycrystalline and thus attribute p-type conduction in ZnO to be associated with grain boundaries \[70\]. These authors discuss segregation of impurities to grain boundaries and subsequent formation of an interfacial complex under O-rich conditions that behaves as an acceptor. Once again this appears to reinforce the concept that complex formation is required for p-type ZnO as opposed to simple impurity incorporation.

4. Summary, Conclusions, and Outlook

Zinc oxide is a fascinating material with numerous potential applications as we have discussed above. Wang et al. \[71\] has referred to ZnO as a unique material that has the “richest family of nanostructures among all materials, both in structure and in properties with novel applications in optoelectronics, sensors, transducers, and biomedical sciences.” These nanostructures include nanowires, nanobelts, nanorings, and nanocages, for example, and they have succeeded in fabricating a ZnO nanowire nanogenerator that is able to convert mechanical energy into electrical energy via piezoelectric to semiconductor coupling \[71\]. Yang et al. \[72\] have also described fabrication of N-doped ZnO nanowire arrays in which a DAP recombination has been observed. Although thin strain-free films of ZnO can be grown homoeptaxially, physical properties of the films can depend on uniformity of crystal quality and presence of defects over the ZnO substrate and details of the growth conditions. For example, homoepitaxial MOVPE growth of ZnO on (11-20) ZnO substrates revealed the existence of two different morphologies dependent upon the growth ambient \[73\]. At a growth temperature of 480°C, a needle microstructure was observed when grown in a N or N\(_2\)O + O\(_2\) environment, whereas a network structure occurred in a NO\(_2\) + O\(_2\) ambient. Both films coalesced however after 15 min at 800°C. More importantly, the nitrogen atomic concentration varied by two orders of magnitude for films grown in the
advances in condensed matter physics

After ex situ anneal
FA (3.303eV)
FA (3.317eV)
11.6K
Intensity (a.u.)

3.15 3.20 ...

without appreciable p-type behavior at 300K. We thus conclude that p-type conductivity in N-doped ZnO involves different ambients (5 × 10^{17} \text{ cm}^{-3} for N_2O + O_2 and 9 × 10^{19} \text{ cm}^{-3} for NO_2 + O_2). Using the concept of domain matching epitaxy [1, 2], ZnO thin films are also grown heteroepitaxially across the misfit scale on nonnative substrates such as sapphire and Si with misfits up to 18%. It is the latter that is particularly exciting as this would enable integration of electronics and optoelectronics on a substrate for multiple functionality. As discussed above, the dominant drawback of ZnO as an optoelectronic material has been the lack of stable and reproducible p-type conduction at room temperature. To achieve this, one must have sufficient incorporation of the impurity species of interest with relatively shallow ionization energies without significant compensation by unintentional impurities and/or native defects. Satisfying both of these criteria is challenging.

As discussed above, there are three primary strategies that have been investigated to generate appreciable hole conductivity in ZnO at room temperature: substitution of Group IA impurities on the Zn sublattice, codoping of donors and acceptors, and substitution of Group VA impurities on the O sublattice. Recent room temperature Hall measurements are summarized in Table 1. On the basis of our analysis above regarding these data, we believe that we are able to assess the likelihood of success for functional devices utilizing each of these approaches, which enables us to postulate several conclusions that have significant impact on the promising outlook for zinc oxide.

With regard to Group IA impurities on a zinc site, the relatively low hole concentrations suggest that Li and Na are not optimum dopants in ZnO in spite of their relatively low ionization energies. The best that has been reported for Na-doped ZnO is 2.1 × 10^{17} \text{ cm}^{-3} after 254 nm UV illumination [29]. Such a low doping level is marginal at best for fabricating efficient lasers and/or LEDs. Similarly, codoping of donors and acceptors has resulted in only slightly higher hole concentrations in the low- to mid-10^{17} \text{ cm}^{-3} range [21, 24, 26, 35], and it was necessary to incorporate post-growth annealing in the (Al, N) codoped films. The latter especially suggests to us that the observed p-type conductivity may be related more to formation of complexes, as discussed above for N-doping of ZnO, during an anneal than codoping per se. In spite of DFT calculations [40] that imply Ga and N codoping should be more effective for p-type behavior than Al and N codoping, a comparison of the experimental data in Table 1 reveals no difference between the two. This certainly suggests that a complete understanding of donor/acceptor codoping has thus far not been achieved.

On the basis of analysis above of recent data, it appears in general that Group VA impurities on the oxygen site and more specifically complexes involving nitrogen on an oxygen site, N\textsubscript{O}, are the most likely to yield significant p-type conduction. While Sb-doped ZnO has resulted in an appreciable hole concentration of 1.7 × 10^{18} \text{ cm}^{-3} after an 800°C post-growth in situ anneal and the hole concentration increased with the Sb effusion cell temperature [45], Friedrich et al. [51] recently reported that increased Sb concentrations resulted in Sb-O phase separation while Liu et al. [52] found that lower doping level concentrations lead to donor behavior. This suggests then that a fundamental doping phase space exists for Sb-doped ZnO, which can be deleterious to electronic and photonic devices.

Our view then is that N-doped ZnO holds the most promise for stable and reproducible p-type behavior. However, as we have described, it cannot be a sole substitutional nitrogen on an oxygen site as the predicted ionization energies are greater than 0.3 eV, suggesting N\textsubscript{O} is a deep acceptor without appreciable p-type behavior at 300 K. We thus conclude that p-type conductivity in N-doped ZnO involves

![Figure 7: 11.6 K PL spectra of p-type ZnO before and after the ex situ 800°C O\textsubscript{2} anneal.](image)

![Figure 8: Comparison of X-ray spectra for nominally undoped (black) and N-doped (red) ZnO films grown using an alternating growth scheme.](image)
formation of complexes that are shallow acceptors with ionization energies that are consistent with those reported in Table 2. It was Liu et al. [63] who used density functional theory and first suggested evolution of a shallow acceptor from a $N_{Zn}$-$V_{O}$ double donor to a $N_{O}$-$V_{Zn}$ double acceptor. They based their analysis on the surface reaction pathways for N-doped ZnO. The initial step however was contrary to the statement by Park et al. [37] that nitrogen antisites, that is, $N_{Zn}$, do not exist in ZnO. But, the recent experimental results by Reynolds et al. [36] imply that formation of the initial double donor complex in an oxygen rich environment is crucial and thus support the pathway suggested by Liu et al. [63] with the caveat related to the significance of hydrogen. The significance of hydrogen in ZnO has been well documented by Van de Walle [42, 43]. Reynolds et al. [36] have recently shown that p-type conduction in ZnO thin films grown by MOVPE is associated with a three-step process in which one employs an alternating Zn-rich and O-rich growth environment to incorporate sufficient nitrogen as a metastable double donor, $N_{Zn}$-$V_{O}$, on the zinc site, an \textit{in situ} anneal that supplies sufficient activation energy for the nitrogen to hop to the adjacent oxygen site to form the shallow acceptor complex, $V_{Zn}$-$N_{O}$-$H^{+}$, and lastly, an \textit{ex situ} anneal to eliminate hydrogen from native defects without removing the acceptor complex. Basically then, one can view formation of p-type conduction during the \textit{ex situ} anneal as a competition between removing H-decorated native defects without significant reduction of the $V_{Zn}$-$N_{O}$-$H^{+}$ acceptor complex, having a resultant ionization energy of $134$ meV. Raman spectroscopy was used to follow the reaction pathways under different growth and annealing conditions that enabled the authors to identify [36] the specific complexes involved, which are consistent with the model [63] suggested by Liu, et al. However, there are competing models [74, 75] for hole conduction, which our existing data cannot definitively exclude. These models are based on $N_{O}$ pairs complexed with a hydrogen atom [74] and $N_{Zn}$ on a Zn site [75]. The latter in particular is intriguing considering that the deduced acceptor ionization energy of $165$ meV is reasonably close to that which we have estimated. Both Liu et al. [63] and Boonchun and Lambrecht [75] are in agreement that growth on the Zn-polar surface is critical for sufficient nitrogen incorporation. Additional EPR and Raman data may help to elucidate the nature of the relevant complexes. It seems clear to us that routine observations of p-type N-doped ZnO is based on a complete understanding of the fundamental mechanisms involved and general acceptance of the applicable model.

One might ask what the outstanding challenges that remain for p-type behavior in zinc oxide are. The dominant issue that remains is demonstration of the stability and reproducibility of p-type ZnO, which can only be accomplished by additional growth under the conditions described and subsequent Hall and Raman measurements over time. There must be a critical assessment of the Hall measurements and associated data. It is generally recognized that it is difficult to make contacts to ZnO, and, hence, one might question interpretation of Hall results. One of the issues that have been discussed with regard to interpretation of Hall measurements for polarity is sample uniformity and placement of contacts.

Ohgaki et al. [76] showed that false positive Hall coefficients can be observed on n-type ZnO crystals and attributed this behavior to sample inhomogeneity. A subsequent analysis by Bierwagen et al. [77] of various sample nonuniformities and contact placement supported the results in [76]. In particular, they showed that an incorrect polarity type can be deduced in the presence of a nonuniform carrier concentration when contacts are placed in the interior of the sample. Most importantly though, they demonstrated that the correct carrier type can be inferred if contacts are placed at sample corners even if inhomogeneities exist. Recently, Macaluso et al. [78] have also questioned assignment of carrier type in not-intentionally doped ZnO grown on an InP substrate based on conflicting results from Hall data and photocurrent and C-V measurements. While their Hall measurements suggested conversion from n-type to p-type material after a post-growth anneal, the latter two techniques indicated that the material remained n-type. Thus, they attributed the anomalous Hall results to a highly conducting p-type layer at the ZnO/InP interface formed during a post-growth $600^\circ$ C anneal. This may not be surprising if the defect structure in the ZnO film allowed P outdiffusion during the anneal, which would render the interfacial region to be p-type. Lastly, heteroepitaxial growth of ZnO on various substrates and orientations across the misfit scale may influence the incorporation of dopants as observed in the InP-based system [79].

In summary, recent results described herein strongly suggest that p-type conduction in zinc oxide is feasible based on nitrogen doping on the oxygen sublattice. An understanding of the reaction pathways and specific model to explain the acceptor level is the key to stable and reproducible p-type ZnO. This indeed implies a promising future for zinc oxide based thin film and nanostructured electronic and photonic devices.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

14 Advances in Condensed Matter Physics


C. L. Reynolds Jr. and J. G. Reynolds, "Preliminary magnetization vs applied field indicate that films are p-type," unpublished.


Submit your manuscripts at
http://www.hindawi.com