A FIRST-PRINCIPLES STUDY ON THE ENHANCEMENT OF BERYLLIUM DOPING IN GALLIUM NITRIDE

by

XIAO WANG

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN PHYSICS

THE UNIVERSITY OF TEXAS AT ARLINGTON

August 2005

ACKNOWLEDGEMENTS

I would like to first sincerely thank Dr. Qiming Zhang for both suggesting and supervising this research. He taught me a lot about logical and systematic study in scientific research. His intuition, patience and encouragement make possible my completeness in this thesis. I would also like to extend my gratitude to the members of my thesis defense committee: Dr. Asok K. Ray and Dr. J. Ping Liu. I am very grateful to the faculty members of the Physics Department from whom I have learned a lot in their lectures: Dr. John L. Fry, Dr. Asok K. Ray and Dr. Qiming Zhang. I would also like to present my sincere thank to Dr. James Horwitz, Dr. Roy S. Rubins and Dr. Qiming Zhang for their kind help in my class teaching. During years I have spent here, physics office staff gave me a lot of help. Mr. Doug Coyne and Mr. Jay Atman have offered many valuable suggestions in my lab teaching. And Ms. Margaret Jackymack gave me consistent help in my two years' study and living in UTA.

I would especially like to thank the members of my research group for their discussion, suggestion and friendship. I would also like to thank my family members and friends for their support and encouragement when I was depressed. At last to my roommates Haining and Qingchun: I am so thankful to be taken care of by you, and I really had a good time with you!

July 17, 2005

ABSTRACT

A FIRST-PRINCIPLES STUDY ON
THE ENHANCEMENT OF BERYLLIUM DOPING
IN GALLIUM NITRIDE

Publication No. _____

Xiao Wang, M. S.

The University of Texas at Arlington, 2005

Supervising Professor: Qiming Zhang

The excellent physical and electrical properties of Gallium Nitride (GaN) have made it a good candidate in light-emitting diodes and UV detecting semiconductor materials. However, GaN's p-type doping has long been a difficulty. Although beryllium (Be) substitutials arise as shallow acceptors in GaN, the concentration of Be substitutials, and hence the population of holes, is not high enough. In the present work, formation energies of Be point defects and complex defects are calculated and compared by first-principles density functional theory (DFT) method. We find self-compensation is easily formed when Be substitutials and Be interstitials co-exist in GaN, which is responsible for the low solubility of Be p-type doping. We have examined the idea using oxygen as a co-doping element to overcome the self-

compensation. The formation of oxygen involved complex in GaN has been studied energetically. The charge states have also been considered. The results are compared with the formation of Be-only complexes.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Materials fall into three categories: conductors, insulators and semiconductors. A useful way to visualize the difference between conductors, insulators and semiconductors is to plot the allowed energy for electrons in the materials. Instead of having discrete energy levels for electrons in the case of free atoms, the allowed energy states form energy bands. Crucial to the conduction process is whether or not there are electrons in the conduction band. In insulators, electrons fully fill in the valence band and are separated by a large gap from the conduction band. In conductors, such as metals, the valence band overlaps with the conductor band, and both bands are partially filled. A semiconductor is basically an insulator except that there is a small enough gap between the valence and conduction bands that thermal or other excitations can bridge the gaps. With such a small gap, the presence of a small percentage of a shallow-level doping material can increase conductivity dramatically.

An important physical quantity in the band theory is the Fermi level, the top of the available electron energy level at low temperatures. The position of the Fermi level with the relation to the conduction band is a crucial factor in determining electrical properties.

The properties of materials are often controlled by defects and impurities. This is particularly true in the case of semiconductors, where the incorporation of impurities

in different concentrations determines the electrical properties. There are two types of impurity doping in semiconductors. One is called n-type doping if the addition of an impurity element results in a large number of free electrons available for conduction. A semiconductor can be also made p-type by doping it with a different element so that there are a large number of positive charge carriers, called holes, available for conduction. The holes actually correspond to vacancies or deficiencies of electrons in the bonds holding the atoms in the crystal lattice. Imagine that a p-type block of semiconductor can be placed in perfect contact with an n-type block. Free electrons from n-type region will diffuse across the junction to the p-type side where they will combine with some of the many holes in the p-type material. Similarly, holes will diffuse across the junction in the opposite direction and recombine. In practice, a p-n junction is formed within a single crystal rather than simply joining two pieces together. Electrical contacts on either side of the crystal enable a connection to an external circuit. The resulting device is called a junction diode. In a transistor, a basic element in integrated circuits, two p-n junctions exist. So it is crucial for a semiconductor material to have both p-type and n-type dopants.

Semiconductors can be divided into two different types: elemental semiconductors and compound semiconductors. Elemental semiconductors are mainly from group IV, such as silicon (Si), germanium (Ge), carbon (C), *etc*. Si and Ge are the most frequently used elemental semiconductors. They have been widely used for various integrated circuits in everyday life. In addition to the pure elemental semiconductors, many alloys and compounds are semiconductors. These may be formed

from group III and group V elements (III-V semiconductor), such as GaN, GaAs, and from group II and group VI (II-VI semiconductor), such as ZnSe. The particular advantage of compounds is that they provide the device engineer with a wide range of energy gaps and mobilities, so that materials are available with properties which match exactly specific requirements.

As mentioned above, GaN is one of the III-V compound semiconductors. Structurally, GaN has two polytypes: wurtzite 2H polytype, where the molecules follow the hcp ABABAB… stacking sequence, and the zinc-blende 3C polytype, where the molecules follow the fcc ABCABC…stacking sequence. GaN is a direct and wide band gap (E_g =3.4eV) semiconductor. In the past decade the excellent physical and electrical properties of GaN have made it a good candidate in high-temperature, high-power and high-frequency applications. GaN-based semiconductors have emerged as a very important materials system in light emission in the green, blue, and UV regions of the spectrum, which were previously not accessible with solid-state light emitters. Some high-quality heterostructures are also applied in heterojunction field-effect transistors (HFETs), heterojunction bipolar transistors and light emitting diodes.[1]

GaN-based LEDs (light-emitting diodes) are semiconductor device that emit visible light when electricity is passed through them. The GaN LEDs typically comprise two thin layers of gallium nitride grown on sapphire or silicon carbide (SiC) substrates. The emitted light is monochromatic and the main benefits of this technology include: low power requirement, high efficiency and long life. They are mainly used in

manufacturing of displays, signs and traffic signals, automotive lightning, domestic lightning and medical sensors.

For UV detection and imaging, GaN photocathods are stable of period over 1 year and are robust enough to be re-activated many times. It's reported that it is a simple, quick, sensing and promising method to measure oil concentration in water in accordance with different UV light absorption between oil and water using GaN UV detector.[2]

It is mentioned earlier that the fabrication of p-type and n-type doped layers underlies the design of electronic and optoelectronic devices. GaN's n-type doping has never been a difficulty, and electron concentrations exceeding 10^{19} cm^{-3} can easily be achieved. P-type doing, however, has long been a difficulty. In 1989 Amano *et al.*[3] observed that Mg-doped GaN grown by MOCVD (metal-organic chemical vapor deposition) was highly resistive after growth, but could be activated by low-energy electron beam irradiation. Nakamura *et al.* subsequently showed that the Mg activation can also be achieved by thermal annealing at 700℃ under N_2 ambient.[4] Nakamura further observed that the process was reversible, with p-type GaN reverting to semi-insulating when annealed in a NH_3 ambient. Since then, hole concentrations on the order of 10^{18} cm^{-3} have been achieved and used in devices. However, the limited conductivity of p-type doped layers constitutes an impediment for progress in device applications.

Beryllium (Be) is one of the more promising acceptor impurities for GaN, exhibiting higher solubility and lower ionization energy than Mg, the most commonly

used p-type dopant.[5,6] The ionization energy for Be reported is 170 meV. However, the solubility of Be p-type doping is even lower than Mg in GaN, which is only 10^{17} cm^{-3} under N-rich condition .[7] The cause of the low Be p-type doping is most likely the self-compensation. It will be investigated in this thesis. Meanwhile, oxygen arises as a stable donor in GaN. A recent experiment studying Be and O coimplantation showed the improvement of p-type doping characteristics. It motivates us to study theoretically the co-dope oxygen together with beryllium to enhance the solubility of Be p-type dopants.

First-principles calculations based on density-function-theory reaching high level of accuracy in treating increasingly larger systems at the microscopic level will be used. It obtains detailed information about atomic structure (including atomic relaxations), wave functions, charge densities, formation energies, *etc*. All of these data can be used to elucidate the properties of Be-related defects including Be point defects and Be complex defects. The formation energy of defects will play a key role in the doping study. The energetics of charge states will be also addressed. The comparison of formation energies of those defects will explain the low solubility of Be p-type doping in GaN. Then formation energy of defect complexes arisen from oxygen co-doping will be calculated to compare with Be-only complexes.

The rest of the thesis is organized as follows. In chapter 2, theoretical background will be introduced. Chapter 3 will be devoted to computational details and supercell of GaN. Chapter 4 and 5 will discuss Be point defects and Be complexes. Then oxygen co-doping is involved in chapter 6, and it is found not work well as we speculated. At last conclusion will be presented in chapter 7.

CHAPTER 2

THEORY

2.1 Density Function Theory and Local Density Approximation

Density-function theory (DFT)[8] calculations based on peudopotentials, a plane-wave basis set and a supercell geometry， are now one of the major methods for performing first-principles studies of defects in semiconductors. DFT in the local density approximation (LDA)[8] allows a description of the many-body electronic ground state in terms of single-particle equations and an effective potential. The effective potential consists of the ionic potential due to the atomic cores, the Hartree potential describing the electrostatic electron-electron interaction, and the exchange-correlation potential that takes into account the many-body effects. This approach has proven to describe with high accuracy such quantities as atomic geometry, electronic structures, formation energies, etc. For an isolated N-electron atomic or molecular system under the Born-Oppenheimer nonrelativistic approximation, this is given by

$$\hat{H}\Psi = E\Psi \qquad (2.1.1)$$

where E is the electronic energy, $\Psi = \Psi(r_1, r_2, ...r_n)$ is the electronic wave function and $\hat{H}$ is the Hamiltonian operator in atomic unit

$$\hat{H} = \sum_{i=1}^{N}\left(-\frac{1}{2}\nabla_i^2\right) + \sum_{i=1}^{N} v(r_i) + \sum_{i\neq j}^{N}\frac{1}{r_{ij}} \qquad (2.1.2)$$

where $v(r_i) = -\sum_{I=1}^{M} \frac{Z_I}{|r_i - R_I|}$ (2.1.3)

is the "external" potential due to M nuclei of charges Z_I's.

Thus, for an N-electron system, N and $v(r)$ determine all properties for the ground state.

In place of N and $v(r)$, the first Hohenberg-Kohn theorem legitimizes the use of electron density $\rho(r)$ as the basic variable.[9] It states that, t*he external potential* $v(r)$ *is determined, with a trivial additive constant, by the electron density* $\rho(r)$. Since ρ determines the number of electrons by $\int \rho(r)dr = N$, it follows that $\rho(r)$ also determines the ground-state wave function Ψ and all other electronic properties of the system, for example the kinetic energy $T[\rho]$, the potential energy $V[\rho]$ and the total energy $E[\rho]$. Then we have, for the total energy

$$E[\rho] = T[\rho] + V_{ne}[\rho] + V_{ee}[\rho] = \int \rho(r)v(r)dr + F_{HK}[\rho] \qquad (2.1.4)$$

where $F_{HK}[\rho] = T[\rho] + V_{ee}[\rho]$ is a *universal functional* of $\rho(r)$ in a sense that $F_{HK}[\rho]$ is defined independently of the external potential $v(r)$. Note that $v(r)$ is not restricted to Coulomb potentials.

The second Hohenberg-Kohn theorem provides the energy variational principle.[9] It states that for a trial density $\tilde{\rho}(r)$, such that $\tilde{\rho}(r) \geq 0$ and $\int \tilde{\rho}(r)dr = N$,

$$E_0 \leq E[\tilde{\rho}] \qquad (2.1.5)$$

where $E[\tilde{\rho}]$ is the energy functional of (2.1.4). This means the ground-state electron density is the density that minimizes $E[\rho]$.

Assuming differentiability of $E[\rho]$, the variational principle (2.1.5) plus the normalization constraint requires that the ground-state density satisfies the stationary principle

$$\delta\left\{E[\rho]-\mu\left[\int \rho(r)dr - N\right]\right\}=0 \qquad (2.1.6)$$

which gives the Euler-Lagrange equation

$$\mu=\frac{\delta E[\rho]}{\delta\rho(r)}=v(r)+\frac{\delta F_{HK}[\rho]}{\delta\rho(r)}. \qquad (2.1.7)$$

The quantity μ is the chemical potential.

Eq. (2.1.7) is the basic working equation of *density-functional theory*.[10] If we know the exact $F_{HK}[\rho]$, eq. (2.1.6) would be an exact equation for the ground-state electron density. Once given an explicit form (approximate or accurate) of $F_{HK}[\rho]$, we can apply this method to any system. However, accurate calculational implementations of the density-functional theory are far from easy to achieve, because of the unfortunate (but challenging) fact that the functional $F_{HK}[\rho]$ is hard to come by in explicit forms.

In a trade of simplicity for accuracy, Kohn and Sham (1965) invented an ingenious indirect approach to calculate the kinetic-energy functional $T[\rho]$, the Kohn-Sham (KS) equation.[9] They thereby turned density-functional theory into a practical tool for rigorous calculations.

In analogy with the Hohenberg-Kohn's definition of the universal functional $F_{HK}[\rho]$, Kohn and Sham invoked a corresponding noninteracting reference system, with the Hamiltonian

$$\hat{H}_S = \sum_i^N \left(-\frac{1}{2}\nabla_i^2 \right) + \sum_i^N v_S(r_i) \qquad (2.1.8)$$

in which there are no electron-electron repulsion terms. Instead, the noninteracting electrons are moving in an effective external potential $v_{eff}(r)$, and the ground-state electron density is exactly ρ. In this system there will be an exact determinantal ground-state N-electron wave function

$$\Psi_s = \frac{1}{\sqrt{N!}} \det[\psi_1 \psi_2 .....\psi_N] \qquad (2.1.9)$$

where the $\psi_i's$ are the N lowest eigenstates of the one-electron Hamiltonian $\hat{h}_s$:

$$\hat{h}_S \psi_i = \left[-\frac{1}{2}\nabla^2 + v_s(r) \right] \psi_i = \varepsilon_i \psi_i \qquad (2.1.10)$$

and setting

$$\rho(r) = \sum_{i=1}^N |\psi_i(r)|^2 \qquad (2.1.11)$$

The kinetic energy is $T_s[\rho]$, given by

$$T_s[\rho] = \sum_i^N \left\langle \psi_i \left| -\frac{1}{2}\nabla^2 \right| \psi_i \right\rangle \qquad (2.1.12)$$

The quantity $T_s[\rho]$ in eq. (2.1.12), though uniquely defined for any density, is still not the exact kinetic energy functional $T[\rho]$ specified in eq. (2.1.4). The sparkling idea of

Kohn and Sham is to set up a problem of interest in such a way that $T_s[\rho]$ is its kinetic-energy component, exactly. Then the universal functional $F_{HK}[\rho]$ can be rewritten as:

$$\begin{aligned} F_{HK}[\rho] &= T[\rho] + V_{ee}[\rho] \\ &= T_s[\rho] + J[\rho] + (T[\rho] - T_s[\rho] + V_{ee}[\rho] - J[\rho]) \\ &= T_s[\rho] + J[\rho] + E_{xc}[\rho] \end{aligned} \tag{2.1.13}$$

where $J[\rho]$ is the classical Coulomb interaction energy,

$$J[\rho] = \frac{1}{2}\int \frac{\rho(r)\rho(r')}{|r-r'|} dr dr' \tag{2.1.14}$$

and the defined quantity $E_{xc}[\rho] \equiv T[\rho] - T_s[\rho] + V_{ee}[\rho] - J[\rho]$ is called the *exchange-correlation energy,* which contains the difference between T and T_s, presumably fairly small, and the nonclassical part of $V_{ee}[\rho]$.

The Euler equation now turns to

$$\mu = v_{eff}(r) + \frac{\delta T_s[\rho]}{\delta \rho(r)} \tag{2.1.15}$$

where the KS *effective potential* is defined by

$$\begin{aligned} v_{eff}(r) &= v(r) + \frac{\delta J[\rho]}{\delta \rho(r)} + \frac{\delta E_{xc}[\rho]}{\delta \rho(r)} \\ &= v(r) + \int \frac{\rho(r')}{|r-r'|} dr' + v_{xc}(r) \end{aligned} \tag{2.1.16}$$

with the *exchange-correlation potential*

$$v_{xc}(r) = \frac{\delta E_{xc}[\rho]}{\delta \rho(r)}. \tag{2.1.17}$$

The Kohn-Sham treatment runs as follows. Eq. (2.1.15) with the constraint $\int \rho(r)dr = N$ is precisely the same equation as one obtains from conventional density-functional theory when one applies it to a system of noninteracting electrons moving in the external potential $v_s(r) = v_{eff}(r)$. Therefore, for a given $v_{eff}(r)$, one obtains the $\rho(r)$ that satisfies eq. (2.1.15) simply by solving the N one-electron equations

$$\left[-\frac{1}{2}\nabla^2 + v_{eff}(r)\right]\psi_i = \varepsilon_i\psi_i \qquad (2.1.18)$$

and setting (2.1.11). Here, v_{eff} depends on $\rho(r)$ through eq. (2.1.17); hence eqs. (2.1.16), (2.1.18) and (2.1.11) must be solved self-consistently. One begins with a guessed $\rho(r)$ such as a superposition of atomic charge densities, constructs $v_{eff}(r)$ from eq. (2.1.16), and then finds a new $\rho(r)$ from eqs. (2.1.18) and (2.1.11). The total energy can be computed directly from eq. (2.1.4) with eq. (2.1.13).

Eqs. (2.1.16) through (2.1.18) plus (2.1.11) are the celebrated Kohn-Sham equations. Figure 2.1 shows the self-consistent calculation procedure.

In practice, the exact functional $v_{xc}[\rho(r)]$ is unknown, which makes it difficult to solve Kohn-Sham equations. Approximations are needed to be considered at this time. The simple, yet remarkably successful one is *local density approximation* (LDA).[9] The local density approximation states that, for regions of a material where the charge density is slowly varying, the exchange correlation energy at that point can be considered the

same as that for a locally uniform electron gas of the same charge density $\rho(r)$. In this case

$$E_{xc}^{LDA} = \int \rho(r)\varepsilon_{xc}(\rho(r))dr \qquad (2.1.19)$$

where the function $\varepsilon_{xc}(\rho(r))$ is the exchange and correlation energy per electron of a uniform electron gas of density ρ and depends locally on the density at the position r (see Figure 2.2).

The spin polarized variation (local spin density approximation, or LSDA) replaces the spin averaged energy in the above equation with the energy density of a polarized homogenous electron gas.

Although this approximation is extremely simple, it is surprisingly accurate, and forms the core of the most modern DFT codes. It even works reasonably well in systems where the charge density is rapidly varying. However it tends to underpredict atomic ground state energies and ionization energies, while overpredicting binding energies. It is also known to overly favor high spin state structures. For these reasons there have been attempts to move beyond the LDA, notably through the addition of gradient corrections to incorporate longer range gradient effects.

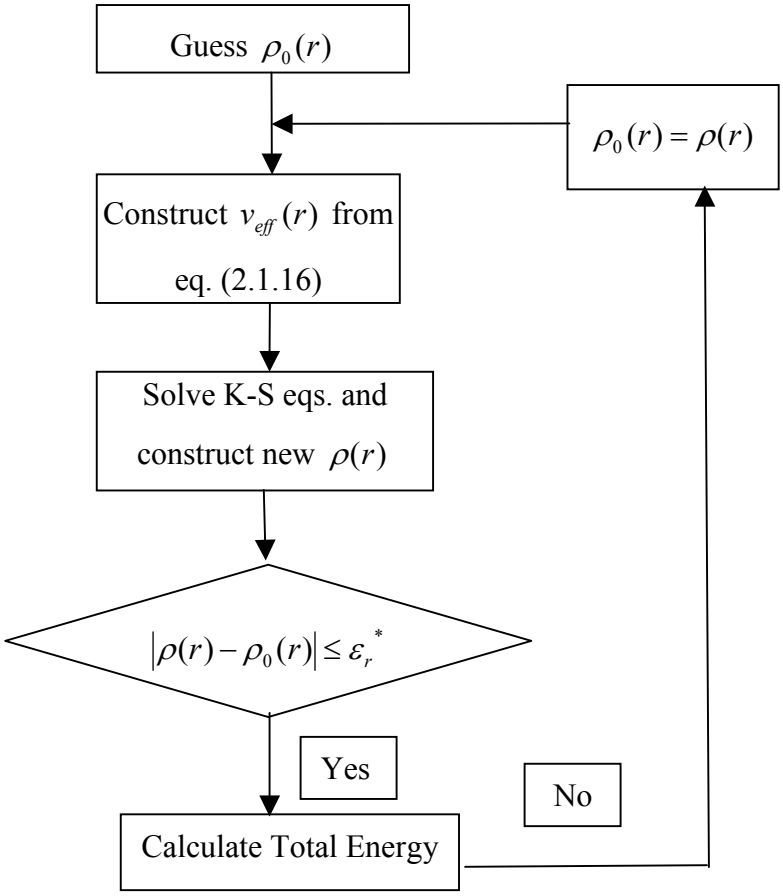

Figure 2.1 Self-consistent Calculation Flow Chart.

* ε_r is the required accuracy and can be determined from experience.

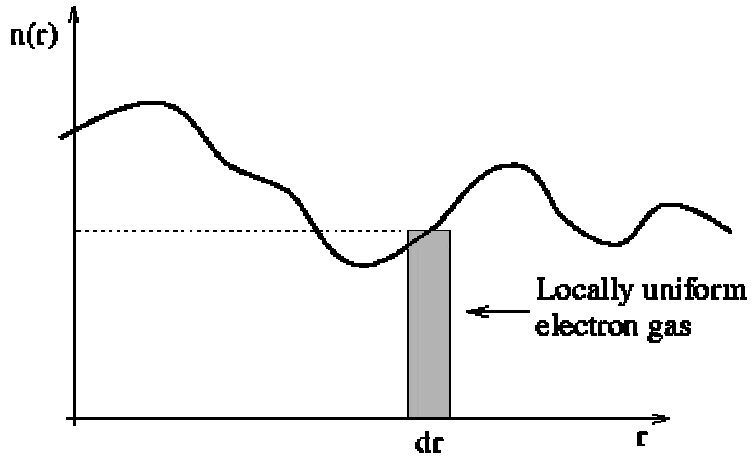

Figure 2.2 Schematic diagram showing the principle of the local density approximation, namely that for a given radial slab, dr, the local charge density can be considered to be n(r), the density of an equivalent uniform homogeneous electron gas.

2.2 Generalized Gradient Approximation

The local density approximation, while the only moderate accuracy, has proven extremely reliable over three decades of use. However, LDA is not accurate enough for most chemical applications, which require the determination of energy differences with considerable precision. New, gradient-corrected functionals, of the form

$$E_{xc}^{GGA} = \int f(\rho(r), \nabla\rho(r))dr \qquad (2.2.1)$$

have reduced LDA atomization energy errors by a factor of 5.[11] Unfortunately a plethora of functions *f* are in use, each yielding different energies for the same system. These functionals may be divided into two broad classes: "locally based" fucntionals, whose construction starts from the uniform electron gas, [12-16] and semiempirical functionals,[17,18] which contain one or more parameters fitted to a particular finite system (or class of system). Such functionals are called generalized gradient approximation (GGA). The culmination of many years of theoretical work has produced the most modern locally based functional, Perdew-Wang 1991(PW91),[15] which contains no empirical input, while Becke exchange[17] and Lee-Yang-Parr (LYP) correlation[18] (BLYP) are perhaps the most popular semiempirical formulas.

From the preceding section, it should be clear that the reliability of LDA depends on more than just the particular choice of local energy per particle, $\varepsilon_{xc}(\rho(r))$. Any GGA that hopes to emulate this reliability should retain the good features of LDA. The simplest form, already suggested by Kohn and Sham, is to expand the energy in a gradient

expansion and include the $\left|\nabla(\rho(r))\right|^2$ terms. However, this gradient expansion approximation (GEA) often does worse than LDA because the associated hole is not the hole of any physical system, and so violates some exact conditions.

Perdew devised a simple procedure to overcome this difficulty, by removing all obviously unphysical contributions to the GEA hole. This defines a numerical GGA which retains the good features of LDA, while improving the description of the average hole (and therefore the energy) by using the gradient. PW91 is an analytic parameterization of this numerical functional, which incorporates several further exact conditions.

Clearly, functionals that have not been constructed in this way should not be expected to be as reliable. However, at the exchange (or λ =0) level, the Becke 88 functional is quantitatively very similar to numerical GGA, and so should be reasonably reliable. Such agreement can be ascribed to the universal nature of the functional, to Becke's use of the correct $\varepsilon_{xc}(\rho(r))$, and to the relatively simple nature of exchange. However, LYP underestimates the correction energy of the uniform electron gas and so cannot be close to the results of numerical GGA.

While retaining many of the best features of LDA, PW91 still incorporates some inhomogeneous effects. The problems include: (1) The derivation is long, and depends on a mass of detail; (2) The analytic function *f*, fitted to the numerical results of the real-space cutoff, is complicated and nontransparent; (3) *f* is overparameterized; (4) The

parameters are not seamlessly joined[19], leading to spurious wiggles in the exchange-correlation potential $\delta E_{XC}/\delta\rho(r)$ for small[20] and large[21] dimensionless density gradient, which can bedevil the construction of GGA-based electron-ion pseudopotentials[22-24]; (5) Although the numerical GGA correlation energy functional behaves properly[25] under Levy's uniform scaling to the high-density limit[26], its analytic parameterization (PW91) does not[27]; (6) Because PW91 reduces to the second-order gradient expansion for density variations that are either slowly varying or small, it describes the linear response of the density of a uniform electron gas less satisfactory than LDA[28].

Anyway, in most of the cases for atoms, molecules, solids and surface, it has been found that Perdew-Wang generalized gradient approximation (PW91) works well.

2.3 Pseudopotential Theory

2.3.1 The choice of a Basis set-Plane waves

The Kohn-Sham orbitals ψ_i may be represented in terms of any complete basis set. Many choices are possible including atomic orbitals, Gaussians, Linearized Augmented Plane Wave Method (LAPW) and plane waves, the basis set we use in practice. The use of a plane wave (PW) basis set offers a number of advantages, including the simplicity of the basis functions, which makes no preconceptions regarding the form of solution, the absence of basis set superposition error, and the ability to efficiently calculate the forces of atoms.

In principle, the representation of an arbitrary orbital in terms of a PW basis set would require a continuous, and hence infinite, basis set. However, the imposition of periodic boundary conditions allows the use of Bloch's Theorem whereby the ψ_i may be written

$$\psi_{i,k}(\vec{r}) = \sum_G c_{i,k}(\vec{G}) e^{i(\vec{k}+\vec{G})\cdot\vec{r}} \qquad (2.3.1)$$

where the sum is over reciprocal lattice vectors $\vec{G}$, $\vec{k}$ is a wave vector which lies within the first Brillion Zone, and $c_{i,k}(\vec{G})$ is the expansion coefficient. Thus the basis set for a given $\vec{k}$ will be discrete, although in principle it will still be infinite. In practice, due to the usual lack of high-frequency PW in a wavefunction, the set of plane waves is restricted to a sphere in reciprocal space most conveniently represented in terms of a cut-off energy, E_{cut}, such that for all values of G used in the expansion

$$\frac{\hbar^2 \left|\vec{k}+\vec{G}\right|^2}{2m_e} \leq E_{cut} \qquad (2.3.2)$$

Thus, the convergence of the calculation with respect to basis set may be ensured by variation of a single parameter, E_{cut}. This is the significant advantage over many other basis set choices, with which calculated properties often show extreme sensitivity to small change in basis set and no systematic scheme for convergence is available.

The principle disadvantage of the use of a PW basis set is the number of basis functions required to accurately represent Kohn-Sham orbitals. This problem may be reduced by the use of *pseudopotentials* as described in the next section.

2.3.2 Pseudopotentials

If all the electrons in a system were explicitly included when performing a calculation and V_{ext} constructed from the full Coulombic potential of the nuclei, the computational cost would still be prohibitive using a plane wave basis set. The rapid oscillations of the wavefunctions near the nucleus, due to the very strong potential in the region and the orthogonality condition between different states, mean that a very large cut-off energy, and hence basis set, would be necessary. It is well known that most physical properties are dependent on the valence electrons to a much greater degree than that of the tightly bound core electrons. It is for the reason that the pseudopotential approximation is introduced. This approximation uses this fact to remove the core electrons and the strong nuclear potential and replace them with a weaker pseudopotential which acts on a set of pseudowavefunctions rather than the true valence wavefunctions. Thus, the core electron states may be assumed to be fixed and a pseudopotential may be constructed for each atomic species which takes into account the effects of the nucleus and core electrons. The pseudowavefunctions corresponding to this modified potential do not exhibit the rapid oscillations of the true wavefunctions, dramatically reducing the number of plane waves needed for their representation. The calculations then need only explicitly consider the valence electrons, offering a further saving in effort.

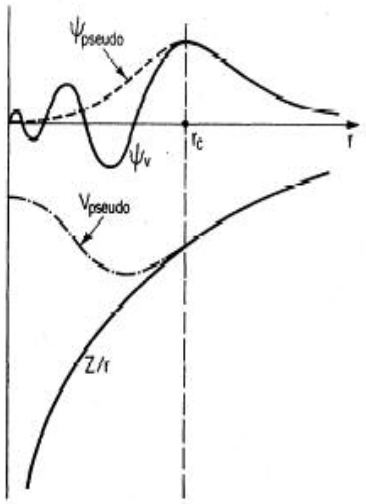

Figure 2.3 Schematic illustration of all-electron (solid line) and pseudoelectron (dashed lines) potentials and their corresponding wave functions.

A pseudopotential is constructed such that it matches the true potential outside a given radius, designated the core radius. Similarly, each pseudowavefunction must match the true wavefunction beyond this distance. In addition, the charge densities obtained outside the core region must be identical to the true density. Thus, the integral of the squared amplitudes of the real and pseudowavefunctions over the core region must be identical. This condition is known as norm-conservation.

The most general form for a non local pseudopotential is

$$V_{NL} = \sum_{lm} |lm\rangle V_l \langle lm| \qquad (2.3.3)$$

where $|lm\rangle$ are the spherical harmonics and V_l is the pseudopotential for angular momentum l. Acting on the electronic wave function with this operator decomposes the wave function into spherical harmonics, each of which is then multiplied by the relevant pseudopotential V_l.

Pseudopotentials are usually constructed using an *ab initio* procedure. The 'true' wavefunctions are calculated for an isolated atom using an all-electron DFT approach. The resulting valence wavefunctions are then modified in the core region to remove the oscillations while obeying the norm-conservation constraint. The Schrodinger equation is then inverted to find the pesudopotenail which will reproduce the pseudowavefunctions. This procedure produces a pseudopotenial which may be transferred between widely varying systems. This contrasts with semi-empirical potentials which are constructed to describe a particular atomic environment and may not be simply transferred to different environment.

2.3.3 Vanderbilt ultra soft pseudopotentials

The utility of that approach by norm-conserving pseudopotentials to systems containing highly localized valence orbitals (e.g., for first-row and transition-metal atoms) has been limited, because of the difficulty of representing the pseudo-wave-functions in a plane-wave basis. Moreover, the norm-conserving condition requires that the total pseudocharge inside the core match that of the all-electron (AE) wave function. Thus for many important cases, e.g., O 2p or Ni 3d orbitals, it has been proven impossible to construct a pseudo-wave-function which is much smoother than the AE one.

Vanderbilt (1990) suggested a rather more radical approach to modifying pseudopotentials for use in plane-wave calculations[29], in which a fully non-local pseudopotential is generated directly, as shown in Figure 2.3. It has the following

desirable properties: (1) It takes the form of a sum of a few separable terms. (2) It becomes local and vanishes outside the core. (3) The scattering properties and their energy derivatives are, by construction, correct at several energies spanning the range of occupied states, and the transferability can be systematically improved by increasing the number of such energies. (4) The norm-conserving constraint is removed so that the pseudo-wave-function can be constructed in such a way as to optimize smoothness. (5) The pseudopotential itself becomes involved in the self-consistent screening process, thereby improving transferability with respect to changes in charge configuration.

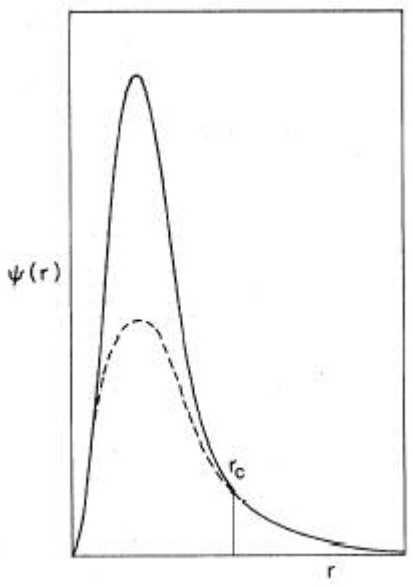

Figure 2.4 Illustration of a pseudo wave function that is strongly peaked inside the core and the modified wave function in Vanderbilt's scheme.

2.4 Computational Method

2.4.1 Steepest Descent Method and Conjugate Gradient Method

In the absence of any information about a function $F(x)$, the optimum direction to move from the point x^1 in a multidimentional space to minimize the function is just the steepest-descent direction g^1 given by

$$g^1 = -\left.\frac{\partial F}{\partial x}\right|_{x=x^1}. \qquad (2.4.1)$$

It will be assumed that direction of steepest descent at the point x^1 can be obtained from a negative of a gradient operator G acting on the vector x^1 so that

$$g^1 = -Gx^1 \qquad (2.4.2)$$

To reduce the value of the function $F(x)$ one should move from the point x^1 in the steepest-descent direction g^1 to the point $x^1 + b^1 g^1$, where the function is a minimum. This can be done by sampling the function $F(x)$ at a number of points along the line $x^1 + b^1 g^1$ in order to determine the value of b at which $F(x^1 + bg^1)$ is a minimum. It should be noted that this process minimizes only the value of the function along a particular line in the multidimentional space. To find the absolute minimum of the function $F(x)$, one must perform a series of such line minimizations. This process is illustrated schematically in the top panel of Figure 2.5.

Although each iteration of the steepest-descents algorithm moves the trial vector towards the minimum of the function, there is no guarantee that the minimum will be reached in a finite number of iterations. In many cases a very large number of steepest-

descents iterations are needed to get close to the minimum of the function. The method of steepest descents performs particularly poorly when the minimum of the function $F(x)$ lies in a ling narrow valley such as the one illustrated in Figure 2.5. The reason for the poor performance in this case is that each steepest-descent vector is orthogonal to the steepest-descent vector of the last iteration and hence many steps of movement are usually required. If the initial steepest-descent vector does not lie at right angles to the axis of the valley, successive vectors will point across rather than along the valley, so that a large number of interations will be needed to move along the valley to the minimum of the function. This problem is overcome by using the conjugate-gradient (CG) technique.

Assume we perform minimization along directions d^1 and d^2. In order for d^1 and d^2 to be independent of each other, one must require that

$$d^1 \cdot G \cdot d^2 = d^2 \cdot G \cdot d^1 = 0 \qquad (2.4.3)$$

where G is a gradient operator defined in eq. (2.4.2). This is the condition that the directions d^1 and d^2 be *conjugate* to each other and can be immediately generalized to

$$d^n \cdot G \cdot d^m = 0 \quad \text{for} \quad n \neq m \qquad (2.4.4)$$

The conjugate-gradients technique provides a simple and effective procedure for implementation of such a minimization approach. In a two-dimensional problem, it is clear that one would need only two conjugate directions, and this would be sufficient to span the space and arrive at the minimum in just two steps, as shown at the bottom of Figure 2.5. It can be proven[30] that, in a multidimentional space, directions generated in this manner are indeed conjugate.

Since conjugate-gradients technique provides an efficient method for locating the minimum of a general function, it should be a good technique for locating the minimum of the Kohn-Sham energy functional in the space spanned by plane waves. It is important, however, to implement the conjugate-gradients technique in such a way as to maximize computational speed, so that each iteration of the method is not significantly more expensive than alternative techniques, and to minimize the memory requirement so that calculations are not limited by the available memory.

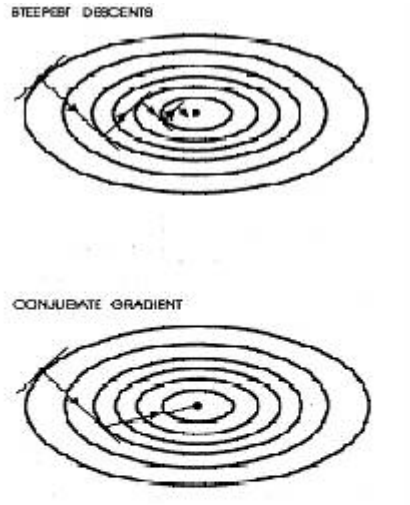

Figure 2.5 Schematic illustration of two methods of convergence to the center of an anisotropic harmonic potential. Top: steepest-descent (SD) method requires many steps to converge. Bottom: conjugate-gradient (CG) method allows convergence in two steps.

2.*4.2 VASP Software Package*

VASP (Vienna *Ab-initio* Simulation Package) is a package for performing *ab-initio* quantum-mechanical molecular dynamics using pseudopotentials and a plane wave basis set. The approach implemented in VASP utilizes density-functional theory under the generalized gradient approximation for the exchange-correlation functional

and an exact evaluation of the instantaneous electronic ground state at a given atomic configuration using preconditioned band-by-band CG scheme. The interaction between ions and electrons is described using Vanderbilt ultra soft pseudopotentials. This technique allows a considerable reduction of the necessary number of plane-waves per atom for transition metals and first-row elements. Forces and stress can be easily calculated with VASP and used to relax atoms into their local minimum. For GGA functional, we use PW91 scheme, and we neglect the spin degree of freedom, as we are interested here in the general trend of formation energies.

CHAPTER 3

COMPUTATIONAL DETAILS

Modern first-principle calculations have had a major impact on the understanding of defects and impurities in semiconductors. With the capability to calculate total energies, it is possible to investigate the atomic structure of the defect, i.e., the stable position in the host lattice, the relaxation of the surrounding atoms, as well as the energy along a migration path. [31-33]

3.1 Periodic supercells

The most common approach for performing calculations for impurities and defects is supercell geometry. The defect is surrounded by a finite number of semiconductor atoms, and that whole structure is periodically repeated.[34-36] This geometry allows the use of various techniques, such as fast Fourier Transformation (FFT) which require translational periodicity of the system. Within the supercell, relaxation of several shells of host atoms around the impurities or defect is always included.

Density Function Theory (DFT) transforms a many-body problem into a single-particle problem. In solids, a periodic condition requires that the single particle wave function be Bloch's wave function $\psi_k(r)$, as $\vec{k}$ is a good quantum number. Each electronic wave function can be expanded by a complete basis functions. Plane waves are a natural choice for the basis:

$$\psi_{\vec{k}}(r) = \sum_{\vec{G}} c_{\vec{k}+\vec{G}} \exp\left[i(\vec{k}+\vec{G})\cdot\vec{r}\right] \qquad (3.1)$$

where $\vec{G}$ is the wave vector in reciprocal space satisfying periodic boundary conditions, and the sum, in principle, is over the whole $\vec{G}$ vectors.

3.2 The k-point sampling

The first Brillouin zone can be mapped out by a continuous set of points, {k= (k_x, k_y, k_z)}, throughout the region of reciprocal space (k-space). The occupied states at each k-point contribute to the electronic potential of bulk solid. Since the set {k} is dense, there are an infinite number of k-points in the Brillouin zone at which the wavefunctions must be calculated. Therefore if a continuum of wave basis sets were required, the basis set for any calculation would still be infinite, no matter how small the plane wave energy cut-off was chosen.

While studying crystal properties, one often encounters Brillouin-zone integrals such as

$$I = \int_{BZ} f(\vec{k}) d^3k = \frac{(2\pi)^3}{\Omega} \bar{f}\,, \qquad (3.1)$$

where the integrand $f(\vec{k})$ is a periodic function of wave vector and Ω is the primitive cell volume. As shown in eq. 3.1, this integral can be expressed as the Brillouin-zone volume times the average value of $f(\vec{k})$.[37] One or a set of k points can be used to evaluate eq. (3.1).

3.2.1 Γ-point

In practice, selecting only a primary cell can not provide enough information about the object we are focusing on, such as defects. Therefore, we need to choose a larger unit cell called *supercell*, which contains more atoms. Since the size of the cell is inversely proportional to the first Brillouin zone, enlarging the cell means shrinking the size of the first Brillouin zone. In practice, when the size of the cell approaches infinitely large, all k points will shrink into one special point, the Γ point ($\vec{k}$ =0). The Γ point is significant from the computational point of view. Generally, the Bloch wave function $\psi_i(r) = e^{ik\cdot r}u_i(r)$ is a complex function, but in the case of $\vec{k} = 0$ (Γ point), it could be a real wave function, which can decrease the computation load greatly, which refers to computer memory and CPU time. Regarding how large a supercell is enough to adopt Γ- point sampling, a test is usually needed.

3.2.2 Multiple k-point sampling

For more accurate sampling, Brillouin-zone integrations (see eq. 3.1) are carried out in the present work using the Monkhorst-Pack scheme[38] with a regular spaced mesh of l×m×n points in the reciprocal unit cell shifted from the origin (to avoid picking up the Γ point as one of the sampling points). Symmetry reduces this set of pints in the irreducible part of the Brillouin zone.

When we describe a defect in a supercell approach, defect-defect interactions between defects in neighboring supercells lead to dispersion of the defect-induced levels in the band gap. A truly isolated defect (corresponding to the limit of an infinitely large supercell) would lead to a flat, dispersionless level. The use of special points

actually provides a way of averaging over the defect band that leads to a result which should be very close to the level of the isolated defect.

Γ-point sampling and multiple k-points sampling are both carried out in our calculation. We use Γ-point sampling to relax atoms since the interatomic forces are less sensitive to k-point. Then we use multi-k-point sampling to calculate the total energy of the relaxed systems. Details will be discussed in the following chapters.

3.3 Structure and basic properties of GaN

Structurally, GaN has two polytypes[39]: wurtzite 2H polytype, where the molecules follow the hcp ABABAB…stacking sequence, and the zinc-blende 3C polytype, where the molecules follow the fcc ABCABC…stacking sequence, as shown respectively in fig. 3.1. Since the wurtzite structure is the equilibrium state of GaN, and most of the developed devices such as light-emitting diode (LED) and laser diode (LD) are fabricated from this polytype, the following study hereby concentrates on wurtzite GaN. The physical parameters of GaN can be found in many review papers and books.[39-41] Some important parameters are listed in Table 3.1.

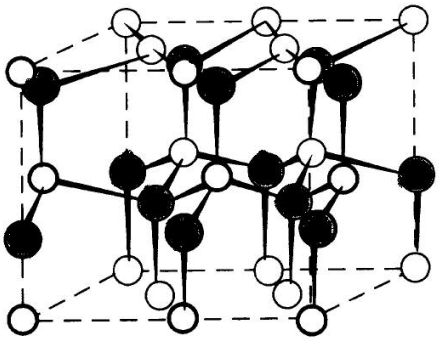

Figure 3.1 Crystal Structure of Wurtzite GaN. The big black balls are Ga atoms, and the small balls are N atoms.

Table 3.1 The fundamental properties and parameters of GaN

Crystal structure	Wurtzite GaN	
Band-gap energy	E_g(200K)=3.39eV	E_g(1.6K)=3.50eV
Temperature coefficient	$dE_g/dT=-6.0\times10^{-4}$eV/K	
Pressure coefficient	$dE_g/dP=4.2\times10^{-3}$eV/kbar	
Lattice constants	a=3.189Å	c=5.185Å
Thermal expansion	$\Delta a/a=2.59\times10^{-6}$K	$\Delta c/c=3.17\times10^{-6}$K
Thermal conductivity	K =1.3W/cm K	
Index of refraction	n(1eV)=2.33	N(3.38eV)=2.67
Melting temperature	>1700℃	
Breakdown field	3.3×10^{6}V/cm	

From the table, one can find that the most attractive feature is the wide energy band gap, which makes it suitable for visible light applications and high-temperature high-power electronics.

3.4 Chemical potential

The chemical potentials of atoms in the bulk material GaN depend on the experimental growth conditions, which can be Ga-rich or N-rich (or anything in between). They should therefore be explicitly regarded as variable in our formalism. However, it is possible to place firm bounds on the chemical potentials, and these bounds will prove very useful in interpreting the results.

3.4.1 Chemical potential of Ga

The Ga chemical potential, μ_{Ga}, is subject to an upper bound: under extreme Ga-rich conditions, $\mu_{Ga}=\mu_{Ga[bulk]}$. Indeed, in thermodynamic equilibrium the Ga chemical potential cannot be higher than the energy of bulk Ga metal. If we tried to push it higher, then we would no longer be growing GaN, but instead precipitating bulk Ga. This is indeed what is experimentally observed: MBE growth of GaN is often carried out under Ga-rich conditions, and precipitation of Ga on the growing surface can be explicitly observed.[42,43]

The ordinary gallium crystal is orthorhombic with two axes of its pseudotetragonal unit almost identical in length. The assigned structure, containing eight atoms per cell, makes these atoms equivalent and places them in special positions of the space group V_h^{18} (Bmab):

(8f) $\pm$ (0 , u, v; $\frac{1}{2}$, u+$\frac{1}{2}$, $\bar{v}$; $\frac{1}{2}$,u, v+$\frac{1}{2}$;0,u+$\frac{1}{2}$,$\frac{1}{2}$ -v)

These coordinates apply to a unit prism with axes in the sequence:

a_0= 4.5107Å; b_0= 4.5167Å; c_0= 7.6448Å.

The established parameters, u= 0.0785, v=0.1525, give an arrangement in which a gallium atom has seven neighbors. One of these is especially near, at 2.44Å, while the others are in three sets of two each, at distances between 2.71Å and 2.80Å.

Our calculated cohesive energy of the 8-atom cell with 8×8×8 k-point sampling is -23.310eV. So μ_{Ga}= -23.310/8= -2.914eV.

3.4.2 Chemical potential of N

Similarly, extreme N-rich conditions place an upper limit on $\mu_N = \mu_{N\ [N2]}$, i.e. the energy of N in a N_2 molecule.

If we assume Ga-rich condition, the chemical potential of N is calculated from the following expression at equilibium:

$$\mu_{Ga} + \mu_N = E_{tot}[GaN], \qquad (3.2)$$

where $E_{tot}[GaN]$ is the total energy of a two-atom unit of bulk GaN. The upper limit on μ_{Ga} then results in a lower limit on μ_N

$$\mu_{Nmin} = E_{tot}[GaN] - \mu_{Ga[bulk]} \qquad (3.3)$$

$E_{tot}[GaN]$ is calculated from a 72-atom supercell divided by 36. Details will be shown in the following section.

3.5 Total energy of a two-atom unit of bulk GaN

3.5.1 3d-electrons in Ga

An analysis of GaN defect and bulk calculations showed that the Ga 3d electrons are not chemically inert but play an important role for the chemical bonding.[32,33,44] 3d electrons are located close to the Fermi level, p-d repulsion then cause the Fermi level to be pushed up. Thus, in general the Ga 3d electrons can not be simply treated as core electrons although they form a closed subshell in Gas tom. They have to be explicitly treated as valence electrons. The localized nature of the Ga 3d states significantly increases the computational demand, requiring an energy cutoff of at 26 Ry in the plane-wave expansions. Our calculations showed that the cohesive energy of GaN was decreased nearly 5% when Ga 3d electrons are included.

3.5.2 Lattice constants

Defects and impurity calculations should be carried out at the theoretical lattice constant, in order to avoid a spurious elastic interaction with defects or impurities in neighboring supercells. Well-converged calculations with good-quality pseudopotentials should produce lattice parameters within a few percent of the experimental value. We tried to vary both lattice constants a_0 and c_0 in order to optimize the crystal energy. Since the ratio of c_0/a_0 is fixed at 1.633, we only need to change a_0. Nine points are chosen with a_0 changing in the range of 3.162Å to 3.220Å.

Table 3.2 Total energy of 4-atom unit-cell of GaN with lattice constant changing

a_0(Å)	c_0(Å)	Total energy(4 atoms) (eV)
3.162	5.164	-24.483
3.180	5.193	-24.507
3.192	5.213	-24.517
3.196	5.219	-24.521
3.200	5.226	-24.523
3.206	5.236	-24.521
3.212	5.245	-24.520
3.220	5.258	-24.517

If we make a plot of all the 8 points, the lowest energy rises clearly.

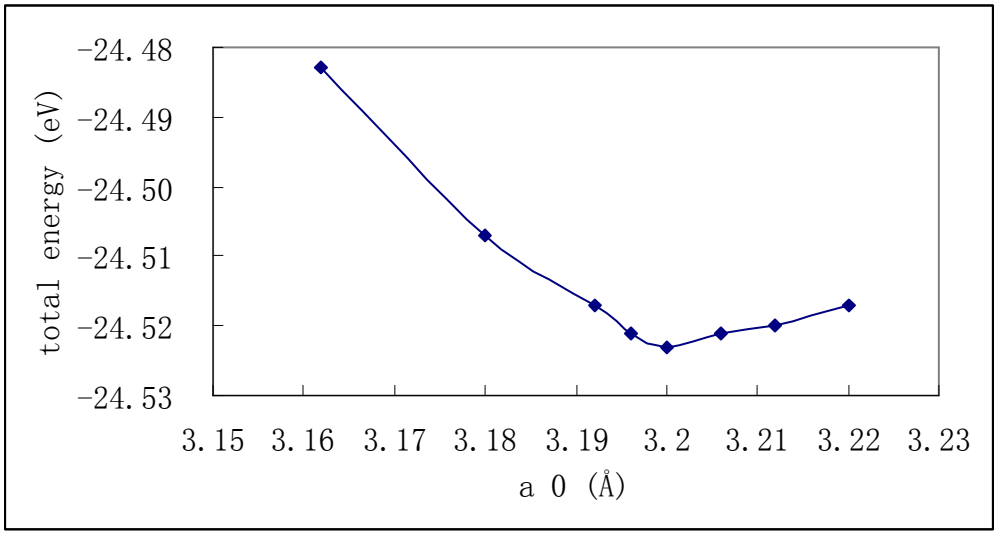

Figure 3.2 Total energy of 4-atom unit-cell vs. a_0

The lowest energy corresponds to the point where $a_0 = 3.20$Å. Compared with the experimental value 3.189 Å (from Table 3.1), the agreement is excellent. For the later calculations, we use this lattice constant.

3.5.3 72-atom supercell

To reduce the artificial interaction of defects due to the period boundary condition, we use a large supercell. To calculate the total energy of bulk GaN with more accuracy, we should use a large supercell. The 72-atom supercell is made up of 36 molecules, i.e. 36 Ga atoms and 36 N atoms, in a hexagonal packing (wurtzite structure) GaN. The whole picture of the 72-atom GaN in the supercell is shown in Figure 3.3.

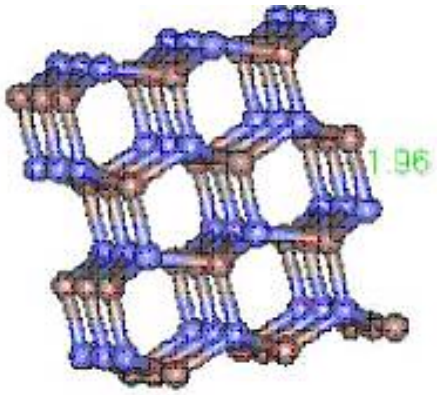

Figure 3.3 72-atom supercell of GaN. The brown balls are Ga atoms, and the blue balls are N atoms.

The bond length of Ga-N is 1.96Å. The total energy of 72-atom GaN is -442.394eV with 4×4×4 k-point sampling. Thus E_{tot}[GaN] mentioned in the previous section is -442.394/36= -12.289eV, and $\mu_N^{min} = E_{tot}[GaN] - \mu_{Ga[bulk]}$= -12.289-(-2.914)= -9.375eV under Ga-rich condition.

CHAPTER 4

POINT DEFECTS IN GALLIUM NITRIDE

Point defects are very common in crystals at finite temperature. There are three basic mechanisms for introducing a point defect in crystals, as shown in Figure 4.1. When a particle is missing at one or more lattice sites we get a vacancy. When a particle forces its way into a hole between lattice sites, we get an interstitial impurity. Substitutional impurities result from replacing the particle that should occupy a lattice site with a different particle. In our calculation, we only consider two types: substitutional impurities and interstitial impurities.

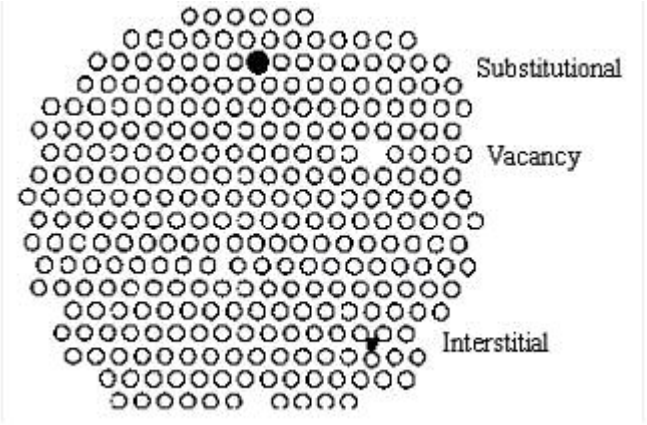

Figure 4.1 Schematic representation of three mechanisms for point defects.

The concentration (D) of a defect at thermal equilibrium temperature T, is approximately determined by

$$D = gN_s \exp[S_f / k]\exp[-E_f / kT], \qquad (4.0)$$

where *g* is the degeneracy factor representing the number of possible configurations for a defect on the same site, N_s is the number of available sites that the defect can be incorporated on, *k* is the Boltzmann constant, and S_f is the formation entropy. An *ab initio* calculation of the formation entropy is currently prohibitively expensive. In low temperature, or in case entropy contributions tend to cancel to some extentwhen comparing relatively free energies, the entropic contributions are small enough to be neglected. Hence D is approximately proportional to $\exp[-E_f / kT]$, and the lower the E_f, the easier the defect can be formed.

4.1 Formation energies

4.1.1 Definition of formation energy

The formation energy of a defect or impurity X in charge state q is defined as

$$E_f[X^q] = E_{tot}[X^q] - E_{tot}[\text{GaN, bulk}] - \Sigma n_i\mu_i + q[\mu_e + E_v], \qquad (4.1)$$

where $E_{tot}[X]$ is the total energy derived from a supercell calculation with one impurity or defect X in the cell, E_{tot}[GaN, bulk] is the total energy for the equivalent supercell containing only perfect bulk GaN, n_i indicates the number of atoms of type i (host atoms or impurity atoms) that have been added to ($n_i > 0$) or removed from ($n_i < 0$) the supercell when the defect or impurity is created, the μ_i are the corresponding chemical potentials of these species, and q represents the charge state. Chemical potentials are discussed in detail in Sec. 3.4. For now, it suffices to know that these chemical potentials represent the energy of the reservoirs with which atoms are being exchanged. μ_e is the Fermi level, referenced to the valence-band maximum (VBM) in the bulk. Due to the not

well defined zero point in the method, we need to explicitly put in the energy of the bulk VBM, E_v, in our expressions for formation energies of charged states. So our Fermi level zero is measured from VBM.

4.1.2 Charge State

In Sec 4.1.1, charge state is already mentioned. Most point defects and impurities can occur in multiple charge states by losing or acquiring extra electrons. The Fermi level works as a reservoir of electrons. Depending on the Fermi level position in the energy band, electrons will be transferred to (from) the defect to lower the E_f. Formation energies have to be calculated for each relevant charge state. The stable charge state is the one which has the lowest formation energy for a given Fermi level. Equation (4.1) shows that the formation energy of charged defects takes into account that electrons are exchanged with the Fermi level. The Fermi level μ_e is referenced with respect to the VBM in the bulk, i.e., μ_e =0 at the top of the valence band (E_v) in bulk GaN.

4.2 Neutral point defects in GaN

Gallium nitride, together with other wide gap nitrides, holds substantial promise for electronic applications.[45] As grown undoped samples of GaN are almost always n type, with the concentration of conduction electrons ranging typically from 10^{17} to 10^{20} cm^{-3}. These values are much higher than concentrations of detected impurities.[46,47] This strongly suggests that the doping is due to native defects. The residual donor was tentatively identified with the nitrogen vacancy.[46-50] When Be is doped in GaN, it can act as an acceptor or a donor. In this chapter, we present the results of formation energy

calculations for a number of Be-related centers in GaN. Firstly, all the neutral Be-dopings are studied and evaluated.

4.2.1 Be substitutials on Ga-site (Be_{Ga})

Be is a IIA element, and Ga is a IIIA element. When Be substitutes Ga, it acts as an acceptor, thus it's a p-type doping. The consensus is that Be_{Ga} is shallower than Mg_{Ga}. For example, depending on the choice of parameters and approximations used, a previous calculation predicts the difference is about 10-15 meV.[51] Another model that depends on the difference in electronegativity between a dopant and the atom it replaces predicts an even lower ionization energy for Be_{Ga}.[52]

When a defect is introduced, the atoms in the supercell need to be relaxed according to their Hellmann-Feynman forces. We use Γ-point sampling in our calculation to relax atoms. The structure of Be_{Ga} is shown in Figure 4.2.

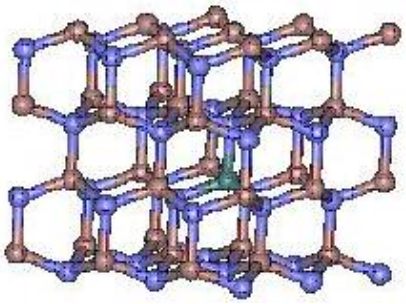

Figure 4.2 Configuration of Be_{Ga}. The brown balls are Ga atoms, the blue balls are N atoms, and the green ball is Be atom.

With the 3-d graphical tool Viewerpro, we cannot see any appreciable difference between the two structures before relaxation and after relaxation because the substituting Be still stays at Ga-site. Before relaxation, the distance between

neighboring Ga atom and N atom is 1.96 Å, while after relaxation the distance between substituting Be atom and N atom decreases to 1.85 Å. The decrease of the distance is due to the attraction between the positive hole Be_{Ga} and the orbiting electrons of N atom. The Ga-N distance at other positions remains almost the same at 1.96 Å. The formation energy is calculated according to eq. (4.1) as 0.70eV. The low formation energy makes Be a stable substitution.

4.2.2 Be substitutials on N-site (Be_N)

N is a VA element. When Be substitutes N, it obviously emerges as an acceptor as it does in Be_{Ga}, but not necessarily a shallow level.

The structure of Be_N is shown in Figure 4.3.

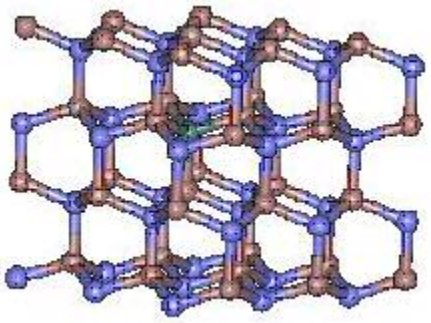

Figure 4.3 Configuration of Be_N. The brown balls are Ga atoms, the blue balls are N atoms, and the green ball is Be atom.

This time the distance between neighboring Be and Ga is increased by 4%. That can be easily expected because Be and Ga are both metallic elements between which a metallic bond is weaker than a covalent bond of Ga-N. Thus the energy cost of relaxation can also be expected to be high. The formation energy is calculated as 4.96eV. The formation energy is so high that Be_N looks very unlikely to form.

4.2.3 Be interstitials on T-site (Be_{It})

In the wurtzite structure, there are two distinct types of interstitial open spaces with high symmetry. T is the tetrahedral interstitial site (or cage site), as shown in Figure 4.4. This site is equidistant from four Ga and four N atoms.

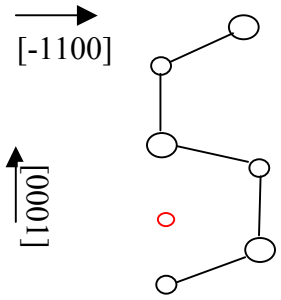

Figure 4.4 Schematic representation of atomic interstitial positions in wurtzite GaN. The large circles represent Ga atoms, medium circles represent N atoms, and the red circle represents the interstitial T-site.

The configuration of Be interstitials at T-site before and after atomic relaxations are also shown with ViewerPro in Figure 4.5.

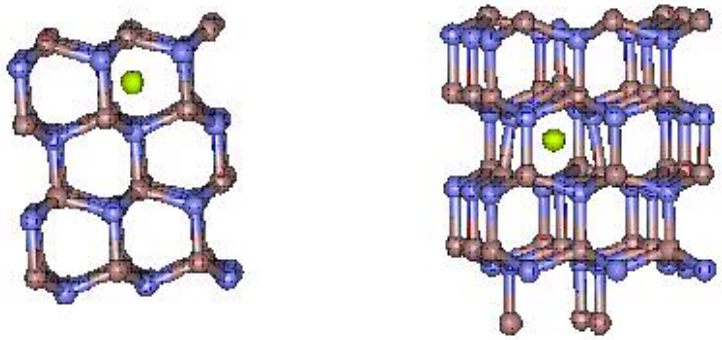

Figure 4.5 Configuration of Be_{It}. The one before relaxation is on the left, and after relaxation is on the right. The brown balls are Ga atoms, the blue balls are N atoms, and the grassy ball is Be atom.

Different from the Be_{Ga} and Be_N configurations, there is a clear distortion of Ga-N bond near interstitial Be. The interstitial Be atom is also a little away from its original tetrahedral interstitial position. The two symmetrically distorted N atoms move toward the interstitial Be atom, and oppositely move away from neighboring Ga atoms. The distance measurement shows Ga-N distance increases from 1.96 Å to 2.08 Å, 2.09 Å and 2.12 Å. Be-Ga distance increases from 1.63 Å to 2.12 Å, while Be-N distance increases a little from 1.63 Å to 1.69 Å. The formation energy is calculated as 2.87eV.

But T-site interstitial is just one possible configuration, and we need to compare it with the other interstitial at O-site.

4.2.4 Be interstitials on O-site (Be_{Io})

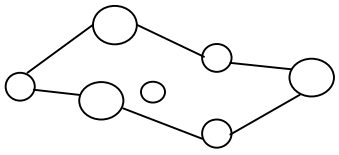

Figure 4.6 Schematic representation of atomic interstitial positions in wurtzite GaN. The large circles represent Ga atoms, the medium circles represent N atoms. The circle in the center is on O-site.

O-site is the octahedral interstitial site. Figure 4.6 shows a O-ring in wurtzite GaN. The interstitial Be lies in the center of the ring, i.e., in the center of the line connecting any two diagonal Ga-N atoms. The O-site is equidistant from six Ga and six N atoms. Both of the T-site and O-site are obvious candidates for local minima of an interstitial defect. The calculations reported in Refs. J. Neugebauer (1994)[53] and J. Neugebauer (1996)[54] found the octahedral site to be most stable for Ga interstitial, while

it was argued in Ref. P. Boguslawski (1995)[55] that the tetrahedral site was slightly more stable than the O site. Our calculations show that it is octahedral site which is more stable.

There is no distinct difference between structures before relaxation and after relaxation for this configuration, so we only show the picture before relaxation in Figure 4.7. With the measurements of distances of Ga-N, Be-Ga and Be-N pairs, we find the distortion also exists as in T-site. However, the interstitial Be atom still remains in the O-ring. The formation energy is calculated as 2.06eV.

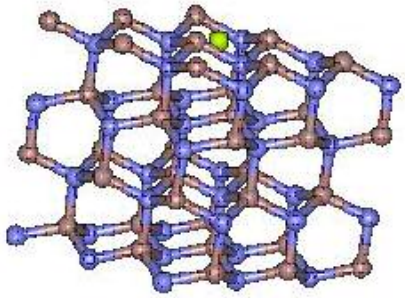

Figure 4.7 Configuration of Be_{Io}. The brown balls are Ga atoms, the blue balls are N atoms, and the grassy ball is Be atom.

Compared with the result calculated on T-site, the formation energy on O-site is substantially lowered. Obviously we are going to take O-site for later more complicate doping complexes.

4.3 Charge state of point defects in GaN

Defect states usually lie in the band gap. Depending on the Fermi level, which serves as an electron reservoir, a defect may gain or lose electron(s) to lower the formation energy. Therefore, in the defects we discussed earlier, charge states are introduced to make the structures more stable. One more electron is added in Be_{Ga}-

substitutial system, and two electrons are removed from Be-interstitial system. We didn't consider Be_N-substitutial system for charge states because its formation energy in neutral state is much higher than other doping configurations.

The formation energies are calculated and listed in Table 4.1:

Combined with neutral states, we plot all the formation energies as a function of μ_e as in eq. (4.1). Since the interstitial-O-site is more representative, only O-site is taken for plotting.

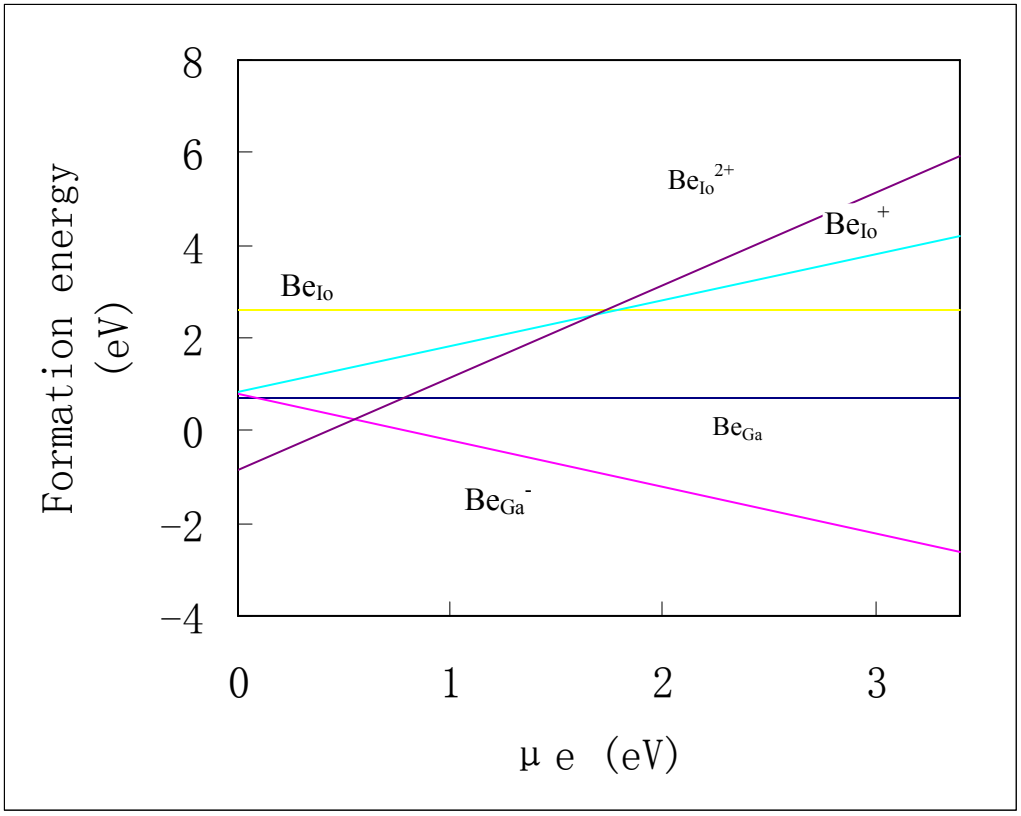

Figure 4.8 Formation energy $E_f(q)$ plotted as a function of μ_e for Be_{Ga}, Be_{Ga}^-, Be_{Io}, Be_{Io}^+ and Be_{Io}^{2+} in GaN under Ga-rich condition.

4.4 Γ-point vs. multiple-k-point calculations

In the first-principles calculation, the atomic force is usually less sensitive to k-point sampling. But energy may be sensitive, depending on the system. In our GaN supercell, this is true. To make the formation energy more accurate, we did the calculations in Sec. 4.2 again with 2×2×2 k-point sampling. Since the interatomic forces are less sensitive to k-point, we use the Γ-relaxed atomic positions for multi-k calculation. Table 4.1 shows the change of the two with different k point set.

Table 4.1 Comparison of formation energy E_f (q) under Ga-rich condition

system	Γ-point	2×2×2 k-point
Be_{Ga}	0.70	0.92
Be_{Ga}^{-}	0.78-μ_e	0.80-μ_e
Be_{It}	2.87	5.03
Be_{It}^{+}	1.10+μ_e	2.08+μ_e
Be_{It}^{++}	-0.60+2μ_e	-0.53+2μ_e
Be_{Io}	2.61	5.05
Be_{Io}^{+}	0.82+μ_e	1.86+μ_e
Be_{Io}^{++}	-0.87+2μ_e	-0.78+2μ_e

More accurate multiple-k-point calculation indicates clearly that the Be_{Ga}^{-} and BeI_o^{2+} are the most preferred Be configurations at p-type GaN (μ_e =0). 2×2×2 sampling changed E_f's significantly in some configurations. Further test on 4×4×4 shows only a few percentage change from 2×2×2's results, so 2×2×2 sampling will be good enough.

4.5 Nitrogen-rich condition

As mentioned in chapter 3, the chemical potentials depend on the experimental growth conditions, which can be Ga-rich or N-rich. Ga-rich condition is discussed in previous sections, and N-rich condition is needed to be evaluated equally.

Extreme N-rich conditions place an upper limit on μ_N given by $\mu_N=\mu_{N[N2]}$= -8.246eV, i.e. the energy of N in a N_2 molecule. $\mu_{Ga}^{min}= E_{tot}[GaN]- \mu_{N[N2]}$= -4.049eV. All the formation energies have to be recalculated when the chemical potentials are changed.

Table 4.2 Comparison of formation energy E_f (q) under N-rich condition

system	Γ-point	2×2×2 k-point
Be_{Ga}	-0.44	-0.02
Be_{Ga}^{-}	-0.36-μ_e	-0.13-μ_e
BeI_t	2.88	5.23
BeI_t^{+}	1.10+μ_e	2.28+μ_e
BeI_t^{++}	-0.60+2μ_e	-0.33+2μ_e
BeI_o	2.06	4.78
BeI_o^{+}	0.29+μ_e	1.62+μ_e
BeI_o^{++}	-1.39+2μ_e	-1.01+2μ_e

The formation energy order under N-rich condition is almost consistent with that in Ga-rich condition. Be_{Ga}^{-} and Be_{Io}^{2+} configurations are more stable than under Ga-rich condition. The complex doping of Be_{Ga} and Be_{Io} are evaluated in the next chapter.

CHAPTER 5

BERYLLIUM PAIR DOPING AND TRIPLET DOPING

Be_{Ga} acts as a shallow acceptor impurity in GaN and the theoretical value of its ionization energy is as low as 0.06eV in wurtzite GaN[7], so that it is a major p-type impurity in GaN. Experiments reported that light Be^{+} ions can be implanted deeper into GaN for a given implantation energy, and they cause less damage in the lattice than Mg^{+} ions.[5] However, it was also reported that high concentration of Be_{Ga}^{-} were difficult to achieve in GaN. To understand the cause of the low concentration of Be in GaN, we have focused on the self-compensation of $Be_{Ga}^{-}$and Be_{Io}^{2+}. If the formation energies of pair doping and triplet doping are lower than that of Be_{Ga}^{-}, the self-compensation should be in charge of low solubility of Be_{Ga}^{-}.

5.1 Beryllium pair doping

It has been proposed that an interstitial Be_I migrate to Be_{Ga} and form $(Be\text{-}Be)_{Ga}$ donors.[56] In chapter 4, Be_{Ga}^{-} performs as a stable acceptor. Meanwhile, Be_I performs as a donor matching with Be_{Ga}^{-}. They can compensate each other to form pairs in GaN. Since Be interstitial T-site provides slightly higher formation energy, it is dropped for later calculations and only O-site is taken for consideration of high concentration doping. Considering the valence electrons, the favorable charge state should be (Be_{Io}-Be_{Ga}^{+}). To find the most stable positions of the two Be atoms, three kinds of

configurations for (Be_{Ga}-Be_{Io}^{+})are taken into account, i.e., far, middle and near, which are defined by their interatomic distance as listed in Table 5.1.

5.1.1 Comparison of formation energies

Table 5.1 Formation energies for pair doping (Be_{Io}-Be_{Ga}^{+}) with different seperations

system	Be-Be distance	2×2×2 points (eV)
Be_{Ga}-Be_{Io}^{+} (far)	6.85 Å	0.56+μ_e
Be_{Ga}-Be_{Io}^{+} (middle)	4.91 Å	0.23+μ_e
Be_{Ga}-Be_{Io}^{+} (near)	1.90 Å	-0.51+μ_e

As we described in chapter 4, to make the calculation more accurate, we use the Γ-point sampling to relax the original atomic positions and the 2×2×2 k-point sampling to calculate the formation energy. Table 5.1 shows the formation energies of a positively charged Be-pair at three different separations. Each configuration is stable. Obviously the formation energy of the configuration with the nearest separation is the lowest, and hence this configuration is the most stable one.

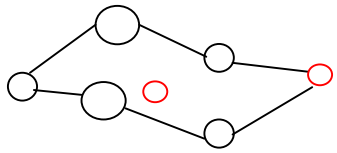

Figure 5.1 Configuration of pair (Be_{Ga}-Be_{Io}^{+}) of the nearest atomic distance of Be-Be. The largest circles are Ga atoms, medium circles are N atoms, and the red circles are Be atoms.

After measuring the Ga-N, Be-Be, and Be-N distances after relaxing, we find the interstitial Be atom almost stays at the same position as in the point defect Be_{Io}

configuration, and the substitutional Be atom almost stays at the same position as the substituted Ga-site in the point defect Be_{Ga} configuration. It demonstrates the proposition mentioned at the beginning of this chapter that Be_I compensate with Be_{Ga} and form $(Be\text{-}Be)_{Ga}$ donors. Compared with substitutial Be_{Ga}^{-}, the formation energy of pair is lowered for 1.31eV with 2×2×2 k-point sampling, which means the pair is much more stable than Be_{Ga}^{-}. The presence of self-compensation explains the failure to achieve p-type electrical activity in GaN doped with Be.

5.2 Beryllium triplet doping

We further assume increasing Be's concentration in GaN. A configuration with high Be concentration is a triplet, i.e., Be_{Ga}-Be_{Io}-Be_{Ga}. It is like the combination of two configurations, Be_{Ga}^{-} and (Be_{Io}-Be_{Ga}^{+}). The interstitial Be atom provides two valence electrons to the two acceptors Be_{Ga}^{-}, and hence it's a self-compensation system as well. The whole complex is electrically neutral.

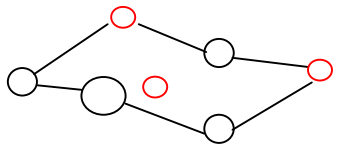

Figure 5.2 Configuration of triplet Be_{Ga}-Be_{Io}-Be_{Ga}. The largest circle represents Ga atom, the medium circles represent N atoms, and the red circles represent Be atoms.

The structure of the triplet Be configuration looks similar as the pair Be. The distances between the interstitial Be atom and neighboring N atoms are increased a little bit than in pair doping. As the same reason in an isolated Be_{Ga}^{-}, the positively charged

Be tends to attract neighboring N atoms because of Coulomb forces. The formation energies of Be pair and triplet are listed in Table 5.2.

Table 5.2 Formation energies for Be pair and triplet doping under Ga-rich condition, in comparison with Be substitutional.

system	E_f (eV)
$[Be_{Ga}^{-}]$	$0.80-\mu_e$
$[Be_{Io}-Be_{Ga}^{+}]$	$-0.51+\mu_e$
$[Be_{Ga}-Be_{Io}-Be_{Ga}]$	-0.35

Obviously, the formation energy of Be complexes are much lower than Be isolated defect, which indicates self-compensation is likely to form. Neither of (Be_{Io}-Be_{Ga}^{+}) or (Be_{Ga}-Be_{Io}-Be_{Ga}) works as an acceptor, so high concentration of Be p-type doping is difficult to achieve.

5.3 Nitrogen-rich condition

As we discussed in Chapter 4, the formation energies depend on the chemical potentials of atoms, namely environment. Formation energies are calculated when Ga-rich condition is changed to N-rich condition. The chemical potential of N (μ_N) is derived from a N_2 molecule,i.e., $\mu_N = E_{tot}[N_2]/2$. The chemical potential of Ga is then calculated from $\mu_{Ga} = \mu_{GaN} - \mu_N$ at equilibrium. Our calculated μ_N = -8.25 eV. Hence μ_{Ga} = -4.05eV. The formation energies of Be pair and triplet under N-rich condition are listed in Table 5.3.

Table 5.3 Formation energies for Be pair and triplet doping under N-rich condition, compared with Be substitutional

system	2×2×2 points (eV)
$[Be_{Ga}^{-}]$	$-0.13-\mu_e$
$[Be_{Io}-Be_{Ga}]^{+}$	$-2.56+\mu_e$
$[Be_{Ga}-Be_{Io}-Be_{Ga}]$	-3.62

The formation energies are much lower than under Ga-rich condition. This is understandable. Under N-rich conditions, we should expect a high concentration of Ga vacancy. That will provide more room for substitutional Be_{Ga}, leading to an easier formation of the triplet.

CHAPTER 6

OXYGEN CO-DOPING IN GALLIUM NITRIDE

6.1 Introduction of oxygen doping

6.1.1 Oxygen in semiconductors

Oxygen is a common impurity in many semiconductor materials and has provided a longstanding challenge for investigators in trying to capture the relevant details of its versatile role. The extent of the oxygen incorporation depends strongly on the growth method and the materials involved. To our concern, oxygen can exist in kinds of nitride semiconductors.[57] Czhochralski-grown GaAs is generally known to contain 10^{15} cm^{-3} oxygen.[58] The degree of oxygen content in GaN is not yet well documented, but secondary ion-mass spectroscopy (SIMS) measurements indicate that GaN can contain oxygen at least in the range 10^{16} -10^{17} cm^{-3}.[59] AlN seems to present an extreme case in this respect, because of the possibility of several percent of oxygen incorporation.[60]

6.1.2 Oxygen in GaN

As introduced in the previous chapters, GaN tends to exhibit excessive concentration of electrons, so it's difficult to make p-type doping. A team at the Lawrence Berkely National Laboratory (LBNL), Materials Science Division, recently pinpointed p-type doping for nitride-type alloys. By placing gallium nitride under very large hydrostatic pressure and studying it with Raman Spectroscopy，the LBNL team

monitored the change in the free electron concentration as a function of pressure. Films containing residual oxygen impurities showed a rapid decrease in the free electron concentration above a certain pressure. Films doped with the shallow donor silicon did not show a decrease in electron concentration. The team determined that the residual oxygen in the gallium nitride is the source of the high concentration of electrons in p-type regions.

Some experimental physicist reported that the activation energy level of Be acceptor level was found to decrease from ~240 to ~163 meV by the implantation of additional O atom.[5] The implantation of additional O atoms into GaN were expected to increase the probability that Be atoms will occupy a Ga-lattice site, which results in improvement of the p-type doping characteristics.

6.1.3 Comparison with other donors in GaN

Calculations for extrinsic donors have been performed for silicon, germanium, and oxygen.[61-64] Silicon is the most widely used intentional n-type dopant, while oxygen is the most likely candidate for unintentional doping.

Figure 6.1 summarizes the first-principles results for native defects and impurities relevant for n-type doping. The figure incorporates information from Refs. J. Neugebauer, 1996[54] and S. Limpijumnong, 2004[65]. We can see that nitrogen vacancies (V_N) have high energies in n-type GaN, and are thus unlikely to occur in significant concentrations. This finding allows us to conclude that nitrogen vacancies are not responsible for n-type conductivity in GaN. In contrast, Fig. 6.1 shows that oxygen and silicon have relatively low formation energies in n-type GaN, and can thus be readily

incorporated. Both oxygen and silicon form shallow donors in GaN. The slope of the lines in Fig 6.1 indicates the charge state of the defect or impurity: Si_{Ga}, O_N, and V_N all appear with slope +1, indicating they are single donors. Since Gallium-site is reserved for Be_{Ga} p-type doping, Si_{Ga} donor is out of our consideration for p-n junction diode.

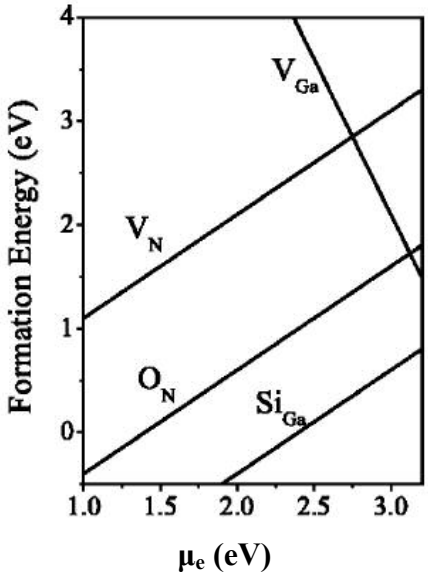

Figure 6.1 Formation energy vs. Fermi energy for native defects (nitrogen and gallium vacancies, V_N and V_{Ga}) and donors (substitutional oxygen and silicon) in GaN. The zero of Fermi energy is located at the top of the valence band. Ga-rich condition and equilibrium with Ga_2O_3 and Si_3N_4 are assumed.

The suggestion that oxygen can be responsible for n-type conductivity in GaN was made by Seifert *et al.*[66] and by Chung and Gershenzon[67]. Still, the prevailing conventional wisdom, attributing the n-type behavior to nitrogen vacancies, proved hard to overcome. After first-principles calculations showed that nitrogen vacancies could not explain the observed n-type conductivity,[53] more detailed experiments were performed, which confirmed that unintentionally doped n-type GaN samples contained concentrations of extrinsic donors (particularly oxygen) high enough to explain the

electron concentrations. The n-type conductivity of bulk GaN can therefore be attributed to unintentional oxygen.

6.2 Oxygen substitutials nitrogen (O_N) point defect

6.2.1 Chemical potential of oxygen

In principle, μ_O could be associated with the elemental value (O_2 dimer), as the chemical potential of N from N_2 molecule,

$$\mu_O = E_{tot}[O_2]/2. \quad (6.1)$$

O_2 is a diatomic molecule with a total spin of 1. So the ground state is a triplet since 2s+1=3. Our calculations show that the chemical potential of oxygen is decreased by 10% from a singlet state, leading to μ_O = -4.91eV.

However, in our 72-atom GaN supercell, when oxygen is introduced, under Ga-rich condition, a stricter criterion for the oxygen chemical potential is set by the formation of the oxides with the group-III elements. That's because the O and Ga reservoirs are coexistent in the system, so the upper limit of the chemical potential of O is determined by the formation of gallium oxide, i.e., Ga_2O_3. The most stringent condition may arise from the formation of Ga_2O_3. The structure of Ga_2O_3 is shown in figure 6.2. The beta-Ga_2O_3 is monoclinic with a tetramolecular cell having the dimensions: a_0= 12.23 Å, b_0 = 3.04 Å, c_0 = 5.80 Å; ß =103. $E_{tot}[Ga_2O_3]$ is calculated as -30.22 eV with 2×8×4 k-point sampling. Therefore μ_O, for the maximum oxygen concentration, has to be determined from the relation

$$3\mu_{Ga} + 2\mu_O = E_{tot}[Ga_2O_3], \quad (6.2)$$

i.e., $\mu_O = (E_{tot}[Ga_2O_3] - 3\times\mu_{Ga})/2 = -8.13$eV from this structure. It is greatly decreased from O_2 gas. And this should be taken for later calculations of formation energies of the complexes with O atom involved.

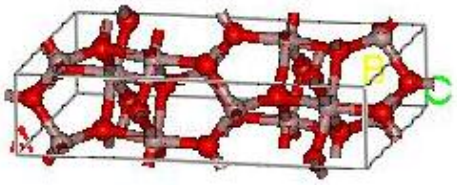

Figure 6.2 Structure of Ga_2O_3 unit-cell. The red balls are O atoms, and the brown balls are Ga atoms.

Under the Nitrogen-rich condition, Ga_2O_3 no longer exists since there is no excess Ga and μ_O is directly from O_2 dimer.

6.2.2 O_N and O_N^+ point defect in GaN under Ga-rich condition

Figure 6.1 shows that O_N exists as a stable donor in GaN. O_N and O_N^+ are associated with our calculations. The change of the structures of O_N configuration is minor. O atom doesn't move away from its original position. The only change is the O-Ga distance. It increased from 1.96Å to 2.05Å. It can be easily understood that the ionic radius of O is a little bit larger than N. Another parameter Ga-O-Ga bond angle almost remains the same at 109 degrees. The configuration of O_N before relaxation is shown in Figure 6.3.

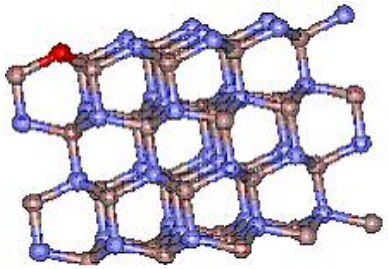

Figure 6.3 Configuration of O_N before relaxation. The brown balls are Ga atoms. the blue balls are N atoms. and the red ball is O atom.

By using the same definition of the symbols as in Chapter 3, the formation energy for O_N is calculated from Ga-rich condition.

$$E_f[O_N]= E_{tot}[O_N]-E_{tot}[GaN, bulk]- \mu_O+ \mu_N \qquad (6.3)$$

The result is -3.48 eV. It's quite low and hence stable. The more favorable charge state O_N^+ for a wide range of Fermi level is also calculated from

$$E^f[O_N^+]= E_{tot}[O_N^+]-E_{tot}[GaN, bulk]- \mu_O+ \mu_N + q[\mu_e+E_v] \qquad (6.4)$$

as -5.24eV $+\mu_e$. The total energy of the defect configuration is calculated with 2×2×2 k-point sampling, using the optimized geometry from Γ- point sampling.

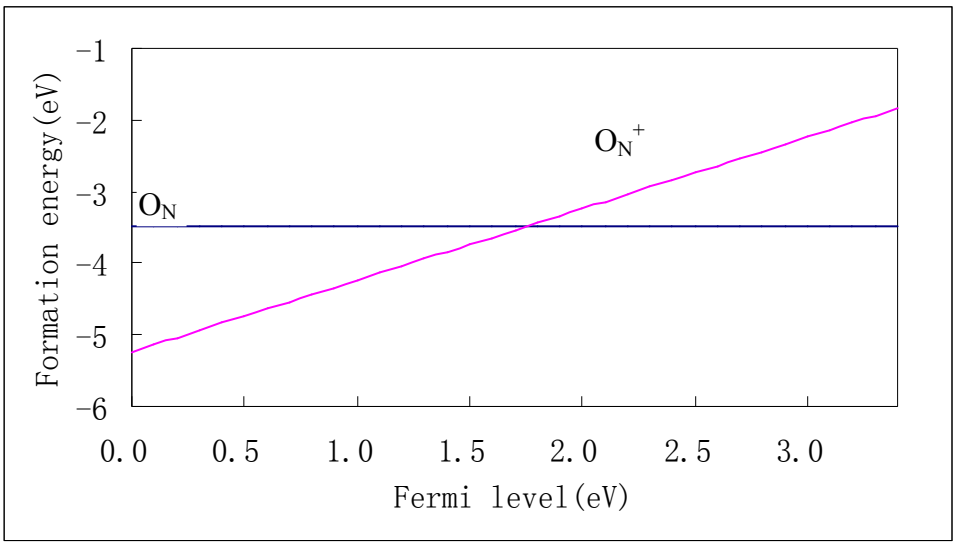

Figure 6.4 Formation energies vs. Fermi energy for the studied defects O_N and O_N^+ in GaN. The band-gap value 3.4 eV is used to give the upper limit for the electron chemical potential.

Our calculation confirms that O_N acts as a stable donor in GaN. It suggests that large oxygen concentrations are likely to occur in GaN, which is in agreement with the SIMS measurements revealing large O contamination.[59] Our results therefore support

the suggestion by Neugenauer and Van de Walle[53] that O_N is likely to act as one of the dominant donors in GaN, causing the unintentional n-type conductivity in as-grown samples, which is in agreement with the experimental observations.[59, 67, 68]

6.3 Oxygen co-doping with Be under Ga-rich condition

Although our results show the high concentration of Be atoms due to the much lower formation energies of Be complexes in GaN, p-type doping is not achieved by them. The triplet is electrically neutral. And the pair acts as a donor. To increase the p-type doping, or the population of Be_{Ga}^{-}, oxygen co-doping is worth to be investigated, since O_N has high solubility in GaN. We expect oxygen would help to enhance the hole population of Be-doped GaN when O and Be are simultaneously incorporated. If only one Be_{Ga}^{-} to combine with O_N^{+}, it forms a neutral complex and compensate each other. But if two Be_{Ga}^{-}'s combine with a O_N^{+} to form a complex, it is still an acceptor in the configuration of (Be_{Ga}^{-}- O_N^{+}- Be_{Ga}^{-}). If the formation energy of this complex is lower than Be pair and triplet, co-doping should provide a solution to the self-compensation of Be in GaN.

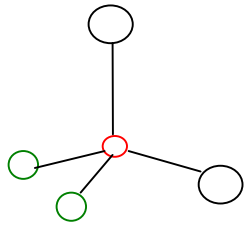

Figure 6.5 Configuration of (Be_{Ga}^{-}- O_N^{+}- Be_{Ga}^{-}). The black circles are Ga atoms, the green circles are Be atoms lying on Ga-site, and the red ball is O atom lying on N-site.

The formation energy of this configuration is calculated as

$$E_f[Be_{Ga}^{-}- O_N^{+}- Be_{Ga}^{-}] = E_{tot}[Be_{Ga}^{-}- O_N^{+}- Be_{Ga}^{-}]- E_{tot}[GaN, bulk]- 2\mu_{Be}+ 2\mu_{Ga}- \mu_O + \mu_N- q(E_v+\mu_e), \quad (6.5)$$

where E_f is 3.66eV $-\mu_e$. At Ga-rich condition, assuming the formation of Ga_2O_3 at the surface, we obtain $E_f[Be_{Ga}^{-}- O_N^{+}- Be_{Ga}^{-}] = 4.81eV - \mu_e$.

The formation energies are high at p-type doping ($\mu_e = 0$), and are against to our speculation. To have a complete picture we have also looked into the nitrogen-rich condition in the next section.

6.4 Oxygen co-doping with Be under N-rich condition

Under N-rich condition, when oxygen and Be are co-doped, both BeO and Be_3N_2 can exist. The chemical potential of Be and O are therefore adjusted from the equilibrium of those two chemicals.

6.4.1 Structure of Be_3N_2

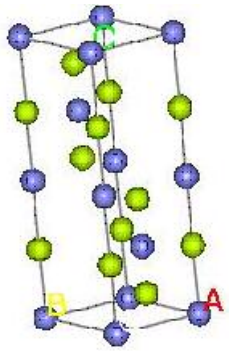

Figure 6.6 Structure of Be_3N_2 unit cell. The blue balls are N atoms and the grassy balls are Be atoms.

The atoms of beta-Be_3N_2 are in the bimolecular hexagonal unit with edge lengths $a_0 = 2.8413$Å and $c_0 = 9.693$ Å. Atoms are in the following positions of D_{6h}^4:

N(1): 0,0,0; 0,0,1/2

N(2): ± (1/3, 2/3, 1/4)

Be(1): ± (0,0,1/4)

Be(2): ± (1/3, 2/3, u; 2/3, 1/3, u+1/2)

with u=0.075.

With 2×2×4 k-point sampling, the formation energy of Be_3N_2 is calculated as -32.08eV. Then $\mu_N = E_{tot}[N_2]/2$ = -8.25 eV, $\mu_{Be} = (E_{tot}[Be_3N_2]-2\times\mu_N)/3$= -5.19 eV.

6.4.2 Structure of BeO

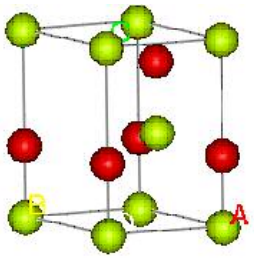

Figure 6.7 Structure of BeO unit cell. The grassy balls are Be atoms and the red balls are O atoms.

The atoms of BeO are in the two-molecule hexagonal unit of the zincite, arranged in the positions of C_{6v}^4:

Be: 0,0,0; 1/3, 2/3, 1/2,

O: 0,0,u; 1/3, 2/3, u+1/2.

With 2×2×2 k-point sampling, the formation energy of BeO is calculated as -13.99eV. So $\mu_O = E_{tot}[BeO]- \mu_{Be}$= -8.80 eV.

6.4.3 The formation energy of Be_{Ga}^-- O_N^+- Be_{Ga}^-

The formation energy of configuration (Be_{Ga}^-- O_N^+- Be_{Ga}^-) is listed in Table 6.1.

Table 6.1 Comparison of formation energies

system	Ga-rich (eV)	N-rich (eV)
[Be_{Ga}^{-}- O_N^{+}- Be_{Ga}^{-}]	4.81-μ_e	1.92-μ_e
[Be_{Ga}-Be_{Io}-Be_{Ga}]	-0.35	-3.62
[Be_{Ga}-Be_{Io}^{+}]	-0.51+μ_e	-2.56+μ_e
[Be_{Ga}^{-}]	0.80-μ_e	-0.13-μ_e

Generally, the formation energies under N-rich are lower than those under Ga-rich condition. However, neither of the two conditions shows that oxygen co-doping could enhance Be_{Ga} p-type doping in GaN as speculated before.

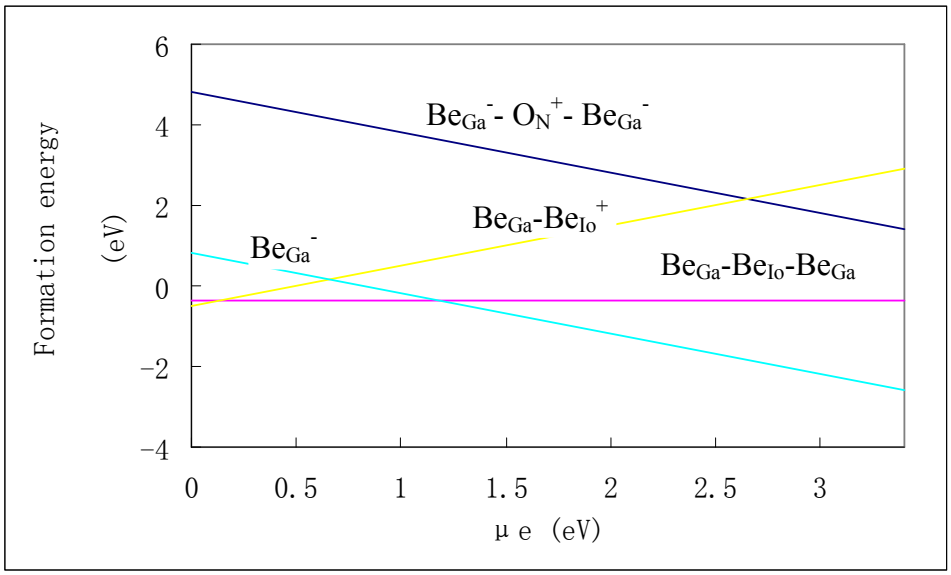

Figure 6.8 Formation energy vs. μ_e under Ga-rich condition with μ_e changing from 0 to 3.4 eV.

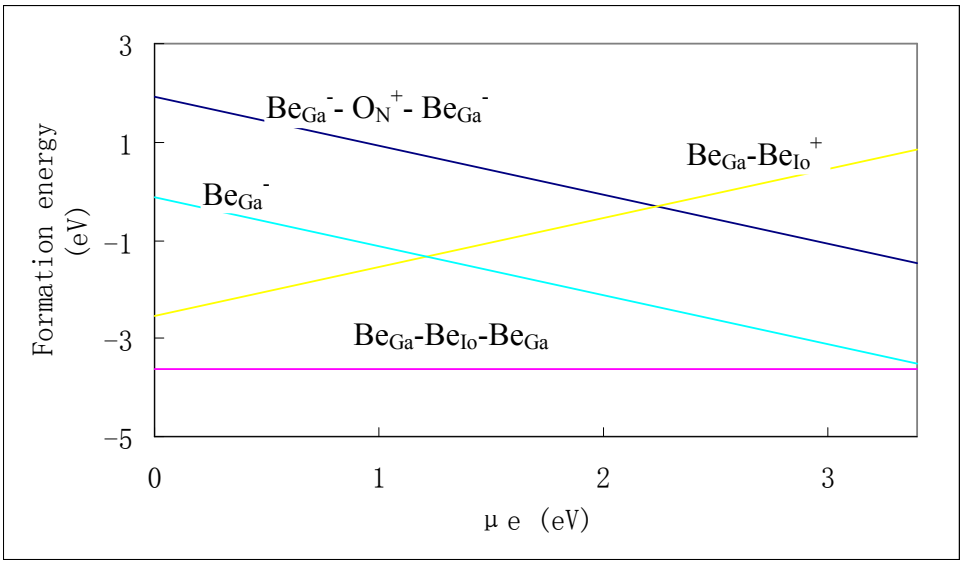

Figure 6.9 Formation energy vs. μ_e under N-rich condition with μ_e changing from 0 to 3.4 eV.

Figures 6.8 and 6.9 are two plots with formation energies changing as a function of μ_e, the Fermi level, under Ga-rich and N-rich conditions. The left side (μ_e=0) is the p-type, and the right side (μ_e=3.4eV) is the n-type. The formation energy of configuration (Be_{Ga}^{-}- O_N^{+}- Be_{Ga}^{-}) is lower than that of (Be_{Ga}-Be_{Io}^{+}) only when μ_e is greater than 2.3 eV under N-rich condition (or 2.7 eV under Ga-rich condition) above the valence band. And it can never be lower than the formation energy of (Be_{Ga}-Be_{Io}-Be_{Ga}), which shows that oxygen co-doping cannot overcome the self-compensation problem in GaN.

CHAPTER 7

CONCLUSION

In conclusion, we have studied the whole problem of p-type doping in Gallium Nitride semiconductor. GaN is a base material for green-to-UV opto-electronics, and piezoelectric and high-power devices. For these applications, controllable doping is an obvious necessity. P-type doping has long been a problem.

By means of density-functional-theory (DFT) calculations of total energies and forces, we show that Be_{Ga}^{-} configuration is thermally stable in GaN compared with other point defect configurations like Be_{Io}^{2+} and Be_{It}^{2+}. However, self-compensation due to the formation of complexes with interstitial atoms is responsible for the poor solubility of Be as a p-type dopant. Our calculations show that the formation energies for pair doping (Be_{Io}-Be_{Ga}^{+}) and triplet doping (Be_{Ga}-Be_{Io}-Be_{Ga}) are both lower than an isolated Be substitutial.

Meanwhile, O_N arises as a good donor in GaN. It motivates us to combine oxygen and beryllium doping together to enhance the solubility of the p-type dopants. Under Ga-rich condition, due to the presence of sufficient oxygen, a film of Ga_2O_3 can be formed on the surface of GaN, and the chemical potential of O is no longer determined by O_2 but by Ga_2O_3. As a result, the formation energy of p-type (Be_{Ga}^{-}- O_N^{+}-Be_{Ga}^{-}) is higher than those of Be complexes, which is against our original speculation. Similarly, under N-rich condition, Be incorporation is limited by the formations of

Be_3N_2 and BeO. The chemical potentials of Be and O are from the equilibrium of Be_3N_2 and BeO. They lead to higher formation energy of (Be_{Ga}^{-}- O_N^{+}- Be_{Ga}^{-}) as well, though a little bit lower than under Ga-rich condition. In summery, our finding of Be+O co-doping in GaN is not working in the enhancement of p-type doping.

For future works, we think this idea may be applied to other impurity than oxygen to co-doping with beryllium in searching the enhancement of p-type doping in GaN. Alternatively, finding another IIA element like magnesium (Mg) to replace Be as a shallow acceptor could also be taken into account for p-type doping in GaN.

REFERENCES

[1] Chiou, Y. Z., Su, Y. K., Chang, S. J. and Gong, G., *Journal of Electronic Materials* (May, 2003).

[2] Yan, W. P., Liu, S., Wei, K. and Zhang, G. Y., *Proceedings of SPIE*, 4414 (2001).

[3] Amano, H., Kito, M., Hiramatsu, K. and Akasaki, I., *J. Appl. Phys.*, Part 2 28, L2112 (1989).

[4] Nakamura, S., Iwasa, N., Senoh, M. and Mukai, T., *Jpn. J. Appy. Phys.*, Part 1 31, 1258 (1992).

[5] Nakano, Y., Kachi, T. and Jimbo. T., *Appl. Phys. Lett.* **82**, 2082 (2003).

[6] Van de Walle, C. G. and Neugebauer., J., *Appl. Phys. Rev.* **95**, 3852 (2004).

[7] Bernardini, F., Fiorentini, V. and Bosin, A., *Appl. Phys. Lett.* **70**(22), 1997.

[8] Perdew, J. and Zunger, A., *Phys. Rev. B* **23**, 5048 (1981).

[9] Kohn, W. and Sham, L. J., *Phys. Rev.* **140** A 1133 (1965).

[10] Parr, R. G.and Wang, W. *Density-functional Theory of Atoms and Molecules.* Oxford. (1989)

[11] Burke, K., Perdew, J. P. and Levy, M., in *Modern Density Functional Theory: A Tool for Chemistry*, Seminario, J. M. and Politzer, P., Eds. (Elsevier, Amsterdam, 1995).

[12] Langreth D.C. and Mehl, M. J., *Phys. Rev. B* **28**, 1809 (1983).

[13] Perdew, J. P., *Phys. Rev. B* **33**, 8822 (1986); **34**, 7406(1986) (E).

[14] Perdew J. P. and Wang, Y., *Phys. Rev. B* **33**, 8800(1986); **40**, 3399(1989) (E).Baldereschi, A., *Phys. Rev. B* **7** (12), 5212 (1973).

[15] Perdew, J. P., in *Electronic Structure of Solids* '91, Ziesche. P., and Eschrig, H., Eds. (Akademic Verlag, Berlin, 1991).

[16] Perdew, J. P., Chevary, J. A., Vosko, S. H., Jackson, K. A., Pederson, M. R., Singh, D. J., and Fiolhais, C., *Phys. Rev. B* **46**, 6671 1992）； **48** 4978 (1993) (E)

[17] Becke, A. D., *Phys. Rev. A* **38**, 3098(1988).

[18] Lee, C., Yang, W. and Parr, R. G., *Phys. Rev. B* **37**, 785(1988).

[19] Zunger, A. and Jaffe, J. *Phys. Rev. Lett.* **51** 662 (1983).

[20] Bernard, J. and Zunger, A. *Phys. Rev.* **36** 3199 (1987).

[21] Wei, S. H. and Zunger. A., *Phys. Rev.* B **43** 1662 (1991).

[22] Mader, K. and Zunger. A., *Phys. Rev. B* **51** 10462 (1995).

[23] Yaguchi, H., Miyoshi, S., Biwa, G., Kibune, M., Onabe, K., Shiraki, Y. and Ito, R., J. *Cryst. Growth* **170** 353 (1997).

[24] Liu, X., Bishop, S. G., Baillargeon, J. N. and Cheng, K. Y., *Appl. Phys. Lett.* **63** (1993).

[25] Zunger, A., *Solid State Physics* vol 39 ed H Ehrenreich, D Turnbull and Seitz (Boston: Academic) p 275 (1986).

[26] Xin, H. P. and Tu, C. W., *Appl. Phys. Lett.* **76** 1267 (2000).

[27] Schwabe, R., Seifert, W., Bugge, F., Bindemann, R., Agekyan, V. F. and Pogarev, S. V., *Solid State Commun*. **55** 167 (1985).

[28] Makimoto, T., Saito, H., Nishida, T. and Kobayashi, N., *Appl. Phys. Lett.* **70** 2984 (1997).

[29] Vanderbilt, D., *Phys. Rev. B* **41**, 1892 (1990).

[30] Payne, M. C., Teter, M. P., Allan, D. C., Arias, T. A. and Joannopoulos, J. D. *Review of Modern Physics* **64**, 1045 (1992)

[31] Baraff, G. A. and Schluter, M., *Phys. Rev. B* **28**, 2296 (1983).

[32] Bar-Yam, Y. and Joannopoulos, J. D., *Phys. Rev. Lett.* **52**, 1129 (1984).

[33] Car, R., Kelly, P. J., Oshiyama, A., and Pantelides, S. T., *Phys. Rev. Lett.* **54**, 360 9 (1985).

[34] Messmer, R. P. and Watkins., G. D., in *Radiation Damage and Defects in Semiconductors* (Institute of Physics and Physical Society, London, 1972), No. **16**, p.255.

[35] Pickett, W. E., Cohen, M. L. and Kittel, C., *Phys. Rev. B* **20**, 5050 (1979).

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

[37] Baldereschi, A., *Phys. Rev. B* **7** (12), 5212 (1973)

[38] Monkhorst, H. J. and Pack, J. D., *Phys. Rev. B* **13**, 5188 (1976).

[39] Morkoc, H., Strite, S., Gao, G. B., Lin, M. E., Sverdlov, B. and Burns, M., *J. Appl. Phys.* (3), 1363 (1994).

[40] Mohammad, S. N., Salvador, Arnel A. and Morkoc, Hadis, *Proceeding of the IEEE*, vol. **83**, No. 10, 1306 (1995).

[41] Pearton, S. J., Zolper, J. C., Shul, R. J. and Ren, F., *J. Appl. Phys.* **86**, 1(1999).

[42] Tarsa, E. J., Heying, B., Wu, X. H., Fini, P. and Speck, S. P., *J. Appl. Phys.* **82**, 5472 (1997).

[43] Brandt, O., Muralidharan, R., Waltereit, P., Thamm, A., Trampert, A., Von Kiedrowski, H. and Ploog, K. H., *Appl. Phys. Lett.* **75**, 4019 (1999).

[44] Baraff, G. A. and Schluter, M., *Phys. Rev. Lett.* **67**, 2339 (1991).

[45] Davis, R.F., Physics B 185, 1 (1993); Strite, S., Morkoc, H. and Vac. J., *Sci. Technol. B* **10**, 1237 (1992); Morkoc, H., *et al*., *J. Appl. Phys.* **76**, 1363 (1994).

[46] Maruska H. P., and Tietjen, J. J., *Appl. Phys. Lett.* **15**, 327 (1969); Monemar, B. and Lagerstedt, O., *J. Appl. Phys.* **50**, 6480 (1979).

[47] Ilegems, M. and Montgomery, M. C., *J. Phys. Chem. Solids* **34**, 885 (1973).

[48] Tansley T. L. and Egan, R. J., *Phys. Rev. B* **45**, 10 942 (1992), and references therein.

[49] Glaser E. R., *et al.*, *Appl. Phys. Lett.* **63**, 2673 (1993).

[50] Perlin, P., *et al*., in *22nd International Conference on the Physics of Semiconductors*, edited by D. J. Lockwood (World Scientific, Singapore, 1995), p. 2383.Boguslawski, P., Briggs, E. L. and Bernholc, J., *Phys. Rev. B* **51**, 17255 (1995).

[51] Mireles, F. and Ulloa, S. E., *Phys. Rev. B* **58**, 3879 (1998).

[52] Podor, B., *Semicond. Sci. technol.* **11**, L827 (1996).

[53] Neugebauer, J. and Van de Walle, C. G., *Phys. Rev. B* **50**, 8067 (1994).

[54] Neugebauer, J. and Van de Walle, C. G., in *Festkorperprobleme/Advances in Solid State Physics*, Vol. **35**, edited by R. Helbig (Vieweg, Braunschweig/Wiesbaden, 1996), p. 25.

[55] Boguslawski, P., Briggs, E.L. and Bernholc, J., *Phys. Rev. B* **51**, 17255 (1995).

[56] Van de Walle, C. G., Limpijumnong, S. and Neugebauer, J., *Phys. Rev. B* **63**, 245205 (2001).

[57] Mattila, T. and Nieminen, R. M., *Phys. Rev. B*, **54,** 23 (1996).

[58] Martin, G. M. and Makram-Ebeid, S., in *Deep Centers in Semiconductors, edited by S. T. Pantelides* (Gordon and Breach, New York, 1986).

[59] Gotz, W., *et al., Appl. Phys. Lett.* **68**, 3144 (1996).

[60] Youngman, R. A., Harris, J. H. and .Ceram. J. A., *Soc.* **73**, 3238 (1990).

[61] Mattila, T. and Nieminen, R. M., *Phys. Rev. B* **55**, 9571 (1997).

[62] Park, C. H. and Chadi, D. J., *Phys. Rev. B* **55**, 12995 (1997).

[63] Mattila, T. and Nieminen, R. M., *Phys. Rev. B* **54**, 16676 (1996).

[64] Boguslawski, P. and Bernholc, J., *Phys. Rev. B* **56**, 9496 (1997).

[65] Limpijumnong, S. and Van de Walle, C. G., *Phys. Status Solidi* **B** 228, 303 (2001).

[66] Seifert, W., Franzheld, R., Butter, E., Sobotta, H. and Riede, V., *Cryst. Res. Technol.* **18**, 383 (1983).

[67] Chung, B. C. and Gershenzon, M., *J. Appl. Phys.* **72**, 651 (1992).

BIOGRAPHICAL INFORMATION

Xiao Wang was born in Jiangsu Province, P. R. China. She received her B. S. degree in Materials Science from Nanjing University in 2003. She received her M. S. degree in Physics from The University of Texas at Arlington in 2005.